Journal of Theoretical and Computational Chemistry, Vol. 2, No. 3 (2003) 0–0

World Scientific
www.worldscientific.com

REDUCED GRADIENT METHODS AND THEIR RELATION TO REACTION PATHS

WOLFGANG QUAPP*
Mathematical Institute, University of Leipzig,
Augustus-Platz, D-04109 Leipzig, Germany
quapp@rz.uni-leipzig.de

Received 27 December 2002
Accepted 14 April 2003

The reaction path is an important concept in theoretical chemistry. We discuss different definitions, their merits as well as their drawbacks: IRC (steepest descent from saddle), reduced gradient following (RGF), gradient extremals, and some others. Many properties and problems are explained by two-dimensional figures. This paper is both a review and a pointer to future research. The branching points of RGF curves are valley-ridge inflection (VRI) points of the potential energy surface. These points may serve as indicators for bifurcations of the reaction path. The VRI points are calculated with the help of Branin's method. All the important features of the potential energy surface are independent of the coordinate system. Besides the theoretical definitions, we also discuss the numerical use of the methods.

Keywords: Potential energy surface; reaction path following; minimum and saddle point; projected gradient; gradient extremal; valley-ridge inflection point; reaction path bifurcation; Branin method; turning point.

1. Introduction

The concept of the minimum energy path (MEP) or reaction path (RP) of an adiabatic potential energy surface (PES) is the usual approach to the theoretical kinetics of larger chemical systems[1–5] (see also Refs. 6–9). It is roughly defined as the line in the coordinate space, which connects two minima by passing the saddle point (SP) (the transition structure) of a PES following the valley. It is able to describe pathways of conformational rearrangements too. The energy of the SP is assumed to be the highest value tracing along the RP. It is the minimal energy a reaction needs to take place (see Fig. 1). The PES has a maximum in one and only one direction. Reaction theories are based either implicitly (transition state theory[1]), or explicitly (variational transition state theory[5]) on the knowledge of the RP. These theories require only local information about the PES along the RP. They circumvent the dimension problem: it is impossible to calculate fully the PES which remains

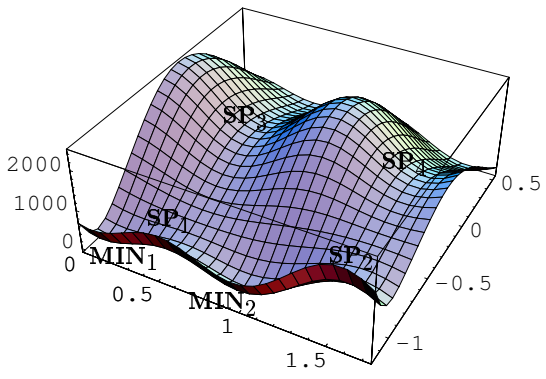

Fig. 1. 2D model potential energy surface.

*Telephone: [49] {0}341-97 32153; fax: [49] {0}341-97 32199; web: http://www.mathe.uni-leipzig.de/∼quapp/

to be *a terra incognita*. The SP and the minima form stationary points of the PES where the gradient vanishes. Roughly speaking, it is only of secondary interest, how a reaction path ascends to the SP. This looseness makes a great variety of RP definitions possible. But the mathematical description of an MEP turned out to be more difficult than expected:[6] the definition of the MEP seems to be a never-ending story!

In Fig. 1 the double minimum, MIN_1/MIN_2, is connected by the RP over SP_1, where the RP over SP_3 may leave the depicted region. Other ways out can lead straight forward over SP_2, or piecewisely combined over SP_2 and SP_4.

The fundamental problem in handling an n-dimensional hypersurface is the dimension. Usually, hypersurfaces are to calculate at discrete grid points, say m, along every one of the n degrees of freedom. A molecular PES is $n = (3N_{atoms} - 6)$-dimensional and needs m^n grid points. It is an overwhelming number for more than $N = 4$ atoms. The RP concept is a promising way out, because it reduces the problem to finding an algorithm for one-dimensional curves — without knowledge of the whole PES. Any parametrization s of the RP $\mathbf{x}(s)=(x^1(s),\ldots,x^n(s))^T$ is called reaction coordinate. Any algorithm which allows one to determine this pathway in a suitable approximation should be tested. The search for valley pathways especially is an important part of the PES analysis; to date it still offers no satisfying concept for all aspects of the problem. One underlying problem seems to be a more principal problem, it concerns many workers in chemistry: they assume that there is one MEP, and all of our different methods are methods to calculate this one MEP. That is not the case. The different methods like

(i) IRC, or steepest descent from SP,
(ii) distinguished or driven coordinate method,[10] or in other modern form RGF,[7,8]
(iii) gradient extremal (GE),[11–15]
(iv) exhaust randomly the pathway,[16]
(v) the variational path,[17]
(vi) or any other method of the future,

define and calculate different curves which may well reflect different aspects of the idea of the MEP, but they also have different drawbacks. These differences are the reason for treating other methods than (i) once and for all. So, the search for an appropriate MEP is not equivalent to the finding of the steepest descent (SD) pathway from the SP.[6,18,19] Also, the curves which follow a driven coordinate, or the projected gradient (RGF) can be used only in certain cases for the minimum path.[7,8] The GE[4,11–15] appeared to represent a suitable ansatz for a minimum path, but, with its many additional solution curves and turning points[14,20,21] this concept in its general form is not suited to be used as a routine program for the calculation of such paths. Additionally, older procedures for the calculation of the GE needed third derivatives of the PES. Interesting here is the term streambed[22] used for the valley-floor GE of the surface, which follows the direction of the eigenvector to the smallest positive eigenvalue. The term "streambed leading downwards to a minimum" is used to characterize the reaction path in two-dimensional model surfaces,[22] but it is used synonymously in any dimension. This GE leaves the minimum uphill with the gentlest ascent.

Here, we will explain in more depth that the combination of the GE concept with the RGF[7,8,21] opens a manageable way to follow a valley of the surface, uphill or downhill. RGF finds a curve where a selected search direction is equal to the gradient of the PES, at every curve point. The RGF method needs gradient and (updates of) the Hessian matrix of second derivatives of the hypersurface. There are curves, which pass all stationary points in most cases. Thus, RGF is an interesting procedure in order to determine by way of trial all types of stationary points.[7] We modify additionally the RGF method to intrinsically search the minimum path. This concept is a proposal of a practicable algorithm for searching minima or SPs of complicated, rugged hypersurfaces using explicitely their valley structure. The valley structure may be of interest by itself as it is the case in spectroscopy or for the selective choice of a reactive channel in chemistry. There, it is assumed that a molecular vibration takes place along the valley of the potential energy hypersurface, and if such a vibration is further excited, it may lead to a chemical reaction.[23]

All these detailed activities for a simple and exact calculation of RPs are prerequisites for a number of dynamical theories to come into operation, including the famous Reaction Path Hamiltonian.[2] The primary

interest in RPs in chemistry is from the viewpoint of mechanisms and dynamics. Further, the methods of direct dynamics[5,24,25] need an exact and physical sensible description of the reaction path.[26] One point of particular interest is that the knowledge of the reaction pathways may give tools for the interpretation of infrared spectra of molecules with high excited vibrations[9] and for the prediction of the conditions of mode selective reactions. This way, the MEP is the leading line characterizing the reaction channel in which the trajectories, or, in terms of quantum mechanics, the wave packets, should move.[23]

Frequently, RP branching does occur — the assumed model curve of the RP bifurcates into two or more branches. The corresponding points on the RP are the so-called bifurcation or branching points (BP). The mathematical description of RP branching is one of those questions which now requires closer consideration in theoretical chemistry. Many procedures — although mostly developed in mathematics — are not yet adapted for use in quantum chemistry though there are a number of recent studies dealing with aspects of the definition of RPs and their bifurcation.[8,13,14,22,27] The choice of a path for a chemical reaction is a complex issue. It is important to state that a BP is a component of the particular RP definition and, therefore, a bifurcation of the RP will be generally found by the BP of those particular curves selected to calculate the MEP. Bifurcations of the path may be caused by symmetry breaking.[28] Then, two or more equivalent pathways may lead over equivalent transition structures to two or more equivalent (or chiral) products. An explanation is already given for the emergence of BP in formaldehyde-like molecules by the second-order Jahn–Teller effect.[29]

It is helpful to consider that RP branching is more often than not connected with the emergence of a special class of points of the PES, the valley-ridge inflection (VRI) points.[30,31] Usually, VRI points represent non-stationary points of the PES. The traditional definition is that a VRI point is that point in the configuration space where, being orthogonal to the gradient, at least one main curvature of the PES becomes zero. The geometrical imagination is clear (but misleading under the dimension aspect): the valley-ridge inflection is that place where an eigenvalue of the Hessian orthogonal to the gradient direction changes from "+" to "–" through zero, or vice versa. A valley inflects into a ridge. The occurrence of such a VRI point is a sufficient condition for a region containing an RP bifurcation (if defined in a coordinate independent way, see below). VRI points can be defined independently of an RP definition. They are generally not identical with BPs of different RP definitions, although both are often adjacent points. "Adjacent" means that no other point of mathematical interest lies in between. A particular aspect is the computation of RP branching using calculations of symmetric VRI points by the Branin method.[8,9,32,33] The VRI points may form a manifold in the configuration space of the molecule.[8] This manifold can have the dimension $n-2$, if the configuration space of the PES has the dimension n.[34,35]

A mathematical simple RP definition is the steepest descent from SP in mass-weighted coordinates, resulting in the well-known intrinsic reaction coordinate (IRC) of Fukui.[6,18,19,36,37] This pathway is defined by an autonomous system of differential equations for a tangent vector along the curve searched for. Its solution is unique. Therefore, if starting at any initial point outside an SP, no bifurcation can occur before reaching the next stationary point. Hence, no branching of PES valleys will be truly described by following the IRC, see the discussion in Refs. 38 and 39. However, following an IRC on a symmetric PES, we may test the curvatures orthogonal to the path, thus orthogonal to the gradient of the potential — and so detect a VRI region.[40]

Gradient extremals (GE) form a second approach for RP following.[4,11–15,20,21] They are more complicated than the IRC, but better fitted to solve the valley branching problem by the determination of a GE bifurcation.[15] However, other problems arise due to the occurrence of pairs of turning points (TP) instead of the BP. Such turning points may interrupt the pathway between minimum and SP. The GE curves often show some kind of avoided crossing.[4,12,13,38,39] The BP indicated by a valley–GE is that point where the valley usually branches into three valleys[21,41] — and, usually, it does not branch into two valleys and a ridge in between, as it is assumed at a VRI point. (Three valleys often mean two "true" valleys to the left and the right, and a "cirque" in the center. A cirque is a valley where the main direction of the ascent

along the valley belongs to the second, or to a higher eigenvalue. Sometimes this cirque ends at a VRI point (see an example in Ref. 21) Because there are cases of missing the BP due to TPs, GE bifurcation is a sufficient but not a necessary condition of the occurrence of branching along an RP. Nevertheless, the whole TP region of two related GE curves has to be considered as a branching region of a valley (or of a ridge). The BP condition for GEs contains third derivatives of the PES, and it does not require a singular Hessian.[15] This indicates that in the general case the BPs of GEs are not the VRI points of the surface, and GE bifurcation can occur without a nearby VRI point.[8]

A third approach to the problem of finding the reaction path branching is quite different. It uses RGF curves. It deals not only with the direct location of the VRI points: the zero eigenvalue of the Hessian, orthogonal to a valley direction, is relatively easy to detect, at least in symmetric PES regions. There are some proposals dealing with this task.[29,40–42] But VRI points are fortunately the BPs of RGF curves. RGF is much simpler to realize in comparison with the GE following.[15,43] RGF needs gradient and (updates of) the Hessian of the PES. Thus, it is more expensive than the IRC.

The independence of different definitions of the RP on the coordinate system was shown for the IRC[6] and for the GE as well.[11,44] Since that time, progress in the understanding of invariance properties has been rather slow. It should also be noted that the calculation of VRI points is developed in a coordinate independent definition.[45,46]

The review is organized as follows: Sec. 2 repeats IRC and the mathematical fundamentals of the RGF method[7–9] and the definition of GEs as well as other methods. We define a modified RGF by the iterative method of the "tangent search". It has a close connection to the streambed GE. Further subsections deal with VRI points. We repeat the so-called global Newton method, or the Branin method. We discuss the numerical execution of the methods. Throughout, the success is demonstrated by some 2D examples. In Sec. 3, we demonstrate the independence of most of the used definitions on the coordinate system. We finally add a discussion. Some of the methods are implemented as independent modular programs. The programs can be obtained on request or retrieved.[a]

2. Theoretical Methods of PES Analysis

There is an arsenal of different methods to follow a geometrically-defined pathway which may serve as a reaction path. Geometrically-defined means that only properties of the PES are taken into account, but that no dynamical behavior of the molecule is observed. The first pathway is the well-known IRC,[36] see also Ref. 6 The favored method used in our laboratory is the so-called "reduced gradient following", RGF,[7–9] a very effective revival of the old distinguished coordinate method.[10] Equivalent curves to RGF are obtained by the global Newton method (Branin curves[32]). Branin's method is additionally well adapted to calculate exactly symmetric VRI points.[8] A recent development from this lab is the TASC method[21,47,48] which allows one to calculate valley floor GEs by second-order methods only.

2.1. *IRC*

The steepest descent (SD) from the SP in mass-weighted Cartesian coordinates[49,50] is the simplest definition of a reaction path, which is well-known as the intrinsic reaction coordinate (IRC),[36,37] but its definition may go back to Euler. Using the arc-length s for the curve parameter, a general *steepest descent* curve $\mathbf{x}(s)$ is defined by the system of vector equations in n dimensions[b]

$$\boxed{\frac{d\mathbf{x}(s)}{ds} = -\frac{\mathbf{g}(\mathbf{x}(s))}{|\mathbf{g}(\mathbf{x}(s))|}} \qquad (1)$$

where $\nabla \mathrm{E}(\mathbf{x}) = \mathbf{g}(\mathbf{x})$ is the gradient of the PES. The SD system is a system of differential equations of the first order allowing an integration constant. Thus, its solution can start at an arbitrary initial point (where the gradient is not zero). The path (1) is given by an autonomous system of differential equations using

[a]E-mail: quapp@rz.uni-leipzig.de
Web: http://www.mathe.uni-leipzig.de/∼quapp/mrgf.html

[b] *We shall denote by boldface lower-case letters geometrical vectors in configuration space or column matrices of their Cartesian coordinates; by upper-case letters second order tensors or square matrices of their components. Scalars are often denoted by lower-case Greek letters.*

the negative normalized gradient of the PES for the tangent vector of the curve. But the gradient is the zero vector at stationary points. With the exception of the stationary points the solution of the differential equation of the IRC is unique. So, the IRC cannot bifurcate[39] and consequently the IRC method is not well-adapted to tackle the problem of reaction path branching. Numerically, the mass-weighted SD is usually started near an SP of index one a step in the direction of the decomposition vector. It is the eigenvector of the Hessian matrix with negative eigenvalue. A variety of numerical methods is available for solving ordinary differential equations. The SD along the gradient, $-\mathbf{g}$, is calculated by discretizing the corresponding differential equation to

$$\mathbf{x_{m+1}} = \mathbf{x_m} - \; l \; \frac{\mathbf{g_m}}{|\mathbf{g_m}|} \tag{2}$$

where m is the step number and l is a steplength parameter used to damp or accelerate the step, but see also the more elaborated methods in Refs. 52–54. Figure 2 shows the IRC between MIN1 and MIN2 for the PES $\{1\}$[51] of torsional and wagging vibration modes of methylamine, $H_3C–NH_2$, with $l = 0.01$. (The use of torsion follows the meaning in spectroscopy.[51])

The test surface $\{1\}$ is

$$\begin{aligned} E(x,y) = {} & 44730.4129 - 66786.5363\cos(y) \\ & + 26352.6908\cos(2y) - 3117.3613\cos(4y) \\ & + 659.3217\cos(6y) \\ & + 621.9640\sin(3x)\sin(y) \\ & - 138.3050\sin(3x)\sin(2y) \\ & - 111.5488\cos(8y) + 41.8227\sin(3x)\sin(4y) \\ & - 7.7979\sin(3x)\sin(6y) + 9.9258\cos(6x) \\ & - 19.0681\cos(6x)\cos(y) \\ & + 8.7063\cos(6x)\cos(2y)\,. \end{aligned}$$

MIN1 corresponds to the torsion angle (x) of 30° and a wagging (y) of −54.54° where MIN2 corresponds to 90° and +54.54°, respectively. The indicated SP is the inversion barrier at the torsion angle of 60° and no wagging; it is depicted in Fig. 3. The use of model PES $\{1\}$ is a reduced model of the full molecular PES. The dimension of 15 internal degrees of freedom is reduced to two coordinates. All couplings between these two coordinates and the other coordinates are neglected.

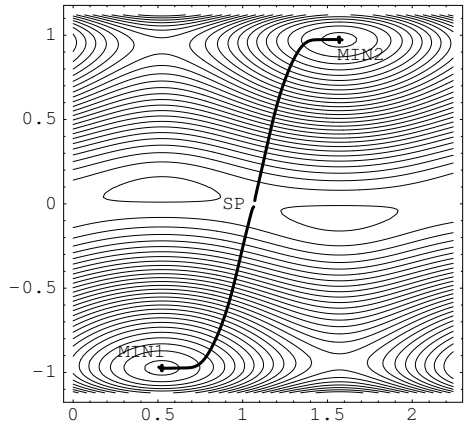

Fig. 2. IRC on test surface $\{1\}$[51].

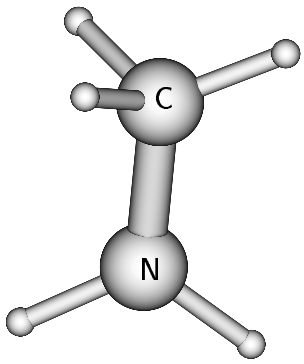

Fig. 3. Methylamine SP: inversion of torsion and wagging.

Usually, we do not propose using a corrector for the IRC method, because of the self-correcting effect that any SD finds its pathway downhill. The SD works well for steep slopes. However, along the bottom of the valley, it shows a second, a numerical disadvantage, the so-called zigzagging across the valley ground line.[19,55] This happens at the valley-floor if the "true" IRC follows asymptotically a GE to the minimum. If working with first-order methods only, it is impossible to suppress the zigzagging. Equation (1) is a so-called stiff differential equation and explicit steps (2) cannot circumvent the problem.[56,57] One may avoid the numerical problems by a quadratic SD. An efficient second-order algorithm was developed by Sun and Ruedenberg[15,58] for integrating Euler's steepest descent differential equation, dispelling the stigma of the latter as a first-order method. (It is implemented[59,60]

in the MOLPRO code.) If using this second-order proposal for the IRC,[26] one needs the Hessian, and with such an effort, one should immediately use an original second-order method like RGF/TASC, see below, or the Newton-Raphson step.

The IRC is used frequently as synonym for the MEP of the PES. But it has a third serious imperfection which is of particular interest here: using Eq. (2), we cannot switch the direction of the search in order to go uphill from the minimum to SP by positive gradient steps.[61,62] This does not work! Any small numerical deviation from the path causes a breakout of the respective steepest ascent and a failing of the SP. If looking around in a small neighborhood of the minimum using only the positive gradient — we can never know where the next SP of index one may be. This is shown using Fig. 4 where in contrast to the former case of Fig. 2, only the SP region is changed artifically, but the region of the two minima is not changed. (For **model {2}**, the last summand in the formula of the PES[51] is modified to $33.333\cos(6x - \pi/4)\cos(2y)$.) Clearly, there are again MEPs along the valleys at $y \approx 1$, and at $y \approx -1$. Nobody can know where at these paths the corresponding IRC from the given SP confluents asymptotically into the MEP. The IRC calculation always needs the knowledge of the corresponding SP. (However, see below the subsections 2.7 and 2.15 for a "gradient-only" uphill search of SPs in special cases.) The indicated SPs of Figs. 2 and 4 are transition states of the "side-on" type in a classification by F. Jensen.[63] They are in contrast to the "end of the valley" type of an SP, or the "top of a valley" type, which we find on the $y \approx 1$, and on the $y \approx -1$ pathways.

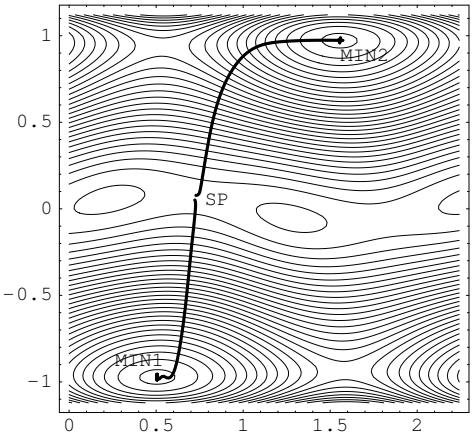

Fig. 4. IRC on test surface $\{2\}$.

In the case of two connected SPs of index one,[4,64] usually the IRC is not the tool to find the connection. This is reported in a paper dealing with H_2NO.[65] If the pathway is not constrained by symmetries, the IRC does not follow the way down from the upper SP to the lower SP, but deviates from the path. Most steepest descent paths reach a minimum tangentially to the one eigenvector of the Hessian with the lowest vibrational frequency.[3,66,67]

In internal curvilinear coordinates, we have to use either the co- or the contravariant versions of gradient and coordinate steps in Eq. (1), which includes the metric of the curvilinearity (see subsection 3.1 for a short description). The use of such coordinates makes the definition of an SD "coordinate invariant".

2.2. *Reduced gradient curves (RGF)*

Some twenty years ago, a proposal was made to choose a driving coordinate along the valley of the minimum, to go a step in this direction, and to perform an energy optimization of the residual coordinates.[54,68–73] A combination of the distinguished coordinate method starting at the SP and steepest descent was also used.[74] However, the methodology of this distinguished coordinate method was truly criticized by some workers, for example by Müller[75] and Brown[54], Williams and Maggiora,[10] and Cioslowski *et al.*[76] They found concrete samples, where the distinguished coordinate method fails: it cannot follow the path over a turning point. Recently, the method was transformed into a new mathematical form.[7] It is remarked that the former limitations do not stem from failures of the approximation of a defined curve but are manifestations of the ill-defined nature of the concept of minimization orthogonal to the distinguished direction.

The chemically most important features of the PES are the reactant and the product minimum and the SP in between. These stationary points of the PES are characterized by the condition

$$\nabla E(\mathbf{x}) = \mathbf{0}\,, \tag{3}$$

when $E(\mathbf{x})$ is the function of the PES, and $\nabla E(\mathbf{x}) = \mathbf{g}$ is its gradient vector in the configuration space, R^n, defined by the coordinates $\mathbf{x}$ of the molecule where $n = 3N$ (N = number of atoms) if Cartesian coordinates are used, or $n = 3N-6$ for internal coordinates. Thus, n indicates the dimension of the problem: $\mathbf{x}$ and $\mathbf{g}$ are vectors of the dimension n. Equation (3) is valid at stationary points of the PES. But single components of the gradient can also vanish in other regions of the PES. Using this property, a curve of points $\mathbf{x}$ is followed which fulfills the $n-1$ equations

$$\frac{\partial E(\mathbf{x})}{\partial x^i} = 0\,, \quad i = 1, \ldots, k-1, k+1, \ldots, n \qquad (4)$$

omitting the kth equation.[7,10] This gives the $(n-1)$-dimensional zero vector of the *reduced gradient*; the method is subsequently called reduced gradient following (RGF). Equation (4) means that the gradient points into the direction of the pure x^k coordinate. The concept may be generalized by the challenge that any selected gradient direction is fixed

$$\mathbf{g}(\mathbf{x})/|\mathbf{g}(\mathbf{x})| = \mathbf{r}\,, \qquad (5)$$

where $\mathbf{r}$ is the selected unit vector of the search direction. The search direction usually corresponds to an assumed start direction of a chemical reaction, for example, to the direction between the two minima of reactant and product. The "reduction" is realized by a projection of the gradient onto the $(n-1)$-dimensional subspace which is orthogonal to the one-dimensional subspace spanned by the search direction $\mathbf{r}$. A curve belongs to the search direction $\mathbf{r}$, if the gradient of the PES always remains parallel to the direction of $\mathbf{r}$ at every point along the curve $\mathbf{x}(s)$

$$\boxed{\mathbf{P_r}\,\mathbf{g}(\mathbf{x}(s)) = \mathbf{0}} \qquad (6)$$

where $\mathbf{P_r}$ projects with the search direction $\mathbf{r}$. This means $\mathbf{P_r}\,\mathbf{r} = \mathbf{0}$. Employing such a projector, instead of Eq. (5), one refrains from the use of the very uncomfortable differentiation of the absolute value in the denominator (in Ref. 77). The parameter s is a suitable parameter, e.g. arc-length; though in practical implementations it is often more convenient to use related parametrizations which are easier to incorporate: the parameter is usually not explicitly used. Equation (4) is a special case of Eq. (6) with the corresponding $(n-1)\times n$ projector matrix

$$\mathbf{P_r} = \begin{array}{c} \\ \\ \\ \\ \\ \\ \textit{column}: \end{array} \begin{array}{c} \begin{array}{ccccccc} 1 & .. & 0 & 0 & 0 & .. & 0 \end{array} \\ \begin{array}{ccccccc} . & .. & . & . & . & .. & . \end{array} \\ \begin{array}{ccccccc} 0 & .. & 1 & 0 & 0 & .. & 0 \end{array} \\ \begin{array}{ccccccc} 0 & .. & 0 & 0 & 1 & .. & 0 \end{array} \\ \begin{array}{ccccccc} . & .. & . & . & . & .. & . \end{array} \\ \begin{array}{ccccccc} 0 & .. & 0 & 0 & 0 & .. & 1 \end{array} \\ \begin{array}{ccccc} 1 & k-1 & k & k+1 & n \end{array} \end{array} \begin{array}{c} \textit{row}: \\ . \\ k-1 \\ k \\ . \\ n-1 \end{array}$$

thus, $\mathbf{P_r}$ is built by the unit vectors orthogonally to the search direction, where here again the kth unit vector is missing. Another possibility to define $\mathbf{P_r}$ is to use the dyadic product in

$$\mathbf{P_r} = \mathbf{I}_n - \mathbf{r}\,\mathbf{r}^T\,, \qquad (7)$$

where $\mathbf{I}_n$ is the unit matrix. This $\mathbf{P_r}$ is an $n\times n$ matrix of rank $n-1$, because $\mathbf{r}$ is a column vector, $\mathbf{r}^T$ is a row vector, and their dyadic product is a matrix.

Based on the explicit definition, we can follow this curve along its tangential vector. This is the RGF method. In contrast to the conventional distinguished coordinate method,[10] a reduced gradient curve passes possible TPs without jumps.[62] RGF uses the derivation of Eq. (6) to obtain the tangent $\mathbf{x}'$ to the curve

$$\begin{aligned} \frac{d}{d\,s}[\mathbf{P_r}\ \mathbf{g}(\mathbf{x}(s))] &= \mathbf{P_r}\frac{d\,\mathbf{g}(\mathbf{x}(s))}{d\,s} \\ &= \mathbf{P_r}\,\mathbf{H}(\mathbf{x}(s))\,\mathbf{x}'(s) = 0\,. \end{aligned} \qquad (8)$$

The matrix $\mathbf{H}$ is the Hessian. The projector $\mathbf{P_r}$ does not depend on the coordinates $\mathbf{x}$ or on the curve parameter. In the general case, the search direction, $\mathbf{r}$, and the tangent, $\mathbf{x}'$, are different. The algorithm is realized by the predictor-corrector method.[8,78] RGF consequently continues the former method of a distinguished coordinate (or driving coordinate) and replaces the energy optimization of the residual coordinates by the solution of the reduced gradient system. RGF is a simple but effective procedure in order to determine all types of stationary points.[7] Unlike the usual SD path from a saddle, the reduced gradient searching for a fixed direction $\mathbf{r}$ locally has an explicit analytical definition. (In Refs. 79 and 80 other predictor strategies are proposed, while in Ref. 81 two alternate corrector schemes are applied to RGF. The

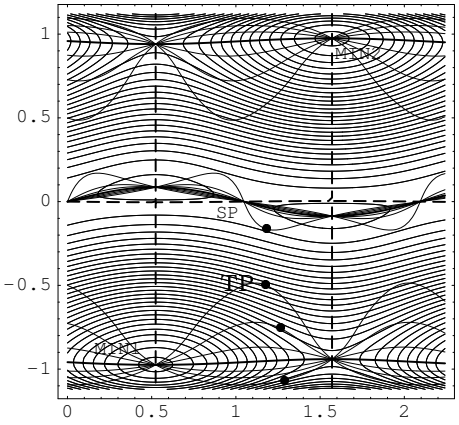

Fig. 5. RGF curves for the torsional transition on surface {1} with turning points (•).

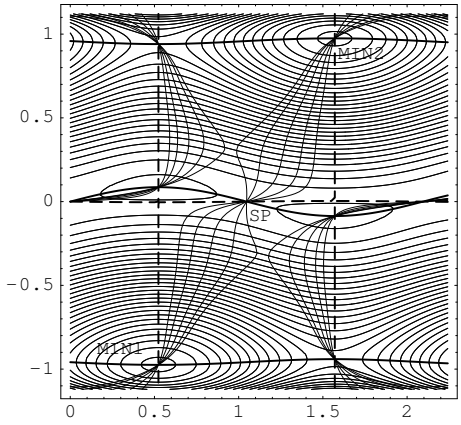

Fig. 6. RGF curves over SP on PES {1}, the inversion barrier of a combined wagging and torsion.

approach of an "activation and relaxation path"[77] resembles the RGF method.)

By the choice of k between 1 and n in Eq. (4) we obtain n different RGF curves where, in the general good-natured case, each of them passes each stationary point. (Possibly, there are different branches which are not connected.) In Fig. 5 the bold-faced curves are those with $E_y = 0$ while the bold-faced dotted curves are $E_x = 0$. Using the more general definition by a projector, Eq. (6), gives an infinite family of RGF curves,[62] where again, in the general good-natured case, each of them passes most of the stationary points. Besides the fat curves in Fig. 5 there is a family of RGF curves to different gradient directions which connect the stationary points along the two horizontal valleys of this PES, at $y \approx -1$, and at $y \approx 1$, as well as the ridge region along $y \approx 0$ where SPs of index one and two are connected. We must recall that the curves of this family usually are not MEPs (whatever this means). RGF curves are defined by a constant gradient vector, and they connect the extrema (see also Fig. 6). Nevertheless, some of the curves may follow a reaction valley in favourable cases, at least qualitatively: here two of the $E_y = 0$ curves in Fig. 5. The possibility of the MEP calculation then depends on a clever definition of the search direction.[82] In Fig. 6 an alternate family of RGF curves is chosen to explore the pathway over the SP of inversion. To find the indicated SP, one of the gradient directions of the family has to be chosen as the search direction for an RGF search, and the respective curve has to follow from minimum to SP.

The turning point case[10,54,62] divides RGF curves into those which can serve as loose RPs, and others: if the RGF curve does not contain a TP at the pathway from minimum up to the SP, it may be used as an RP. However, if a TP emerges, it will have an energy higher than the energy of the respective SP and, hence, this path does not meet the meaning of an RP. In Fig. 5, the MEP from MIN1 to the saddle at (1.57, –0.95) is accompanied by curvilinear RGF curves. The two most curvilinear RGF curves are such curves with a TP. There the tangent of the curve emerges orthogonally to the search direction $\mathbf{r}$. Behind the TP, the respective RGF curve follows a ridge structure, not a valley. That is the reason why a minimization orthogonal to $\mathbf{r}$ fails.

2.3. *Short description of the RGF algorithm*

In order to get the system of equations for RGF, we have to define the projector $\mathbf{P_r}$. In our procedure, mrgf,[a] we calculate $(n-1)$ orthonormal direction vectors being also orthogonal to the selected search direction $\mathbf{r}$ by using the modified Gram-Schmidt algorithm.[83] Then, the projector $\mathbf{P_r}$ is the matrix of these $(n-1)$ rows. Equation (8) for RGF becomes $\mathbf{P_r} \cdot \mathbf{H} \cdot \mathbf{t} = \mathbf{0}$, a linear equation for the tangent vector $\mathbf{t}$. It makes up the predictor step. The system is solved by QR decomposition.[78] The reduced Hessian $\mathbf{P_r} \cdot \mathbf{H}$ is augmented by the tangent vector to an $(n \times n)$

matrix which is the so-called **K** matrix.[78] The corrector step is applied if the norm of the reduced gradient $\mathbf{P_r} \cdot \mathbf{g}_I$ is greater than a threshold ϵ which is given as a parameter. The subsequent Newton-Raphson step of the corrector orthogonally to the tangent is realized by solving a linear equation where the **K** matrix forms the left-hand side, and the right-hand side is given by the reduced gradient augmented by zero in the nth row. Either the predictor step, or the corrector step are added to the current **z** matrix values of the internal coordinates, and the next loop of the algorithm is begun.

The RGF is now tested to be an effective tool in determining the next SP,[62,79–81,84–86] on a PES if starting at a minimum. (It is implemented[85,87] in the COLUMBUS code, also in the TURBOMOL code,[88] as well as in the MOLPRO code.[59])

We may simply test the passing of a bifurcation point by comparing the tangent vector of the predictor step with the previous one. If the tangents point into opposite directions, then a bifurcation point is passed.[78] The test of a TP is the comparison of the tangent vector of the predictor step with the direction of the first step **r**.

Recently, the predictor-corrector method for following a reduced gradient was further accelerated by a modification allowing an implied corrector step per predictor but almost without additional costs.[89] The stability and robustness of the RGF method was improved and the new RGF *version 3* in addition reduces the number of gradient and Hessian calculations. We will explain predictor and corrector steps of RGF with this most developed version of the algorithm.

The predictor step of RGF is done along the tangent $\mathbf{t} = \mathbf{x}'(s)/|\mathbf{x}'(s)|$, thus a solution of (8) and, orthogonal to this direction, Newton-Raphson-like steps of the corrector for a point near to curve (6) are calculated. If the predictor step is done from $\mathbf{x}_1$ to $\mathbf{x}_2$ by a step $p\,\mathbf{t}$, the corrector method by Newton-Raphson-like steps[78] starts with the $(n-1)\times n$ matrix equation, for **c**, to be the step $\mathbf{x}_2 \rightarrow \mathbf{x}_3$

$$\mathbf{P_r}\ \mathbf{H}(\mathbf{x}_2)\ \mathbf{c} = -\mathbf{P_r}\ \mathbf{g}(\mathbf{x}_2)\,, \tag{9}$$

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

$$\mathbf{t}(\mathbf{x}_2)^T\ \mathbf{c} = 0\,, \tag{10}$$

enforcing that the corrector step is orthogonal to the tangent of an RGF curve at $\mathbf{x}_2$. Diener and Schaback proposed[35] to avoid the conventional predictor step, $p\,\mathbf{t}$ with steplength p, and instead to determine a combined step **d** starting with **t** at $\mathbf{x}_1$ by solving Eq. (10) not orthogonally to **t**, however skewly to **t** with the scalar product

$$\mathbf{t}(\mathbf{x}_1)^T\mathbf{d} = p$$

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

$$\begin{aligned} \mathbf{P_r}\ \mathbf{H}(\mathbf{x}_1)\ \mathbf{d} &= -\mathbf{P_r}\ \mathbf{g}(\mathbf{x}_1) \\ \mathbf{t}(\mathbf{x}_1)^T\ \mathbf{d} &= p\,. \end{aligned} \tag{11}$$

The step vector **d** is a combined predictor-corrector step with the component steplength p in the direction of **t** which should give a point $\mathbf{x}_4$ near $\mathbf{x}_3$, being consequently near the searched curve (RGF, or TASC — see below) fulfilling also Eq. (6) by

$$|\mathbf{P_r}\ \mathbf{g}(\mathbf{x}_4)| < \epsilon\,, \tag{12}$$

with the threshold ϵ. In general, only if Eq. (12) is unsatisfied, do we need to take further corrector steps as defined by (9) and (10); but generally, the use of steps **d** avoids these separate corrector steps **c**. It holds especially for a large scaling of ϵ. The former proposal in the original RGF was to set $\epsilon \approx (0.01 \text{ to } 0.1) \times p$. Now, ϵ may be as large as p, or larger. For the original RGF, the inclusion of the corrector via ϵ of Eq. (12) is not to avoid 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 an automatic improvement of the predictor direction. The method shows quite good results by a low effort: usually around 15 steps of the predictor are needed to find the SP of a medium molecule. In the general case, one Hessian must be calculated at the early beginning, and along the path, updates are used (see Ref. 89).

If RGF is used to find stationary points, then the coordinate system does not matter so much; however, RGF curves can also be defined independent of coordinates. In internal curvilinear coordinates, we have the covariant version of the gradient in Eq. (5) and have to use for the selected direction **r** also a covariant vector. It seems intuitively plain that the projector (7)

can be defined as a mixed co- and contravariant tensor by a covariant $\mathbf{r}_{cov}$ and a contravariant $\mathbf{r}^T_{con}$. Then the calculation of condition (6) and steps (8) including the metric of the curvilinearity is described in subsections 3.4 below. The use of such a tensor for internal coordinates again can make the definition of an RGF "coordinate invariant".

2.4. *Branin's method to calculate symmetric VRI points*

Natura non facit saltus.
Titus Lucretius Carus

However, nature can make bifurcations. The BP of a reaction path is an interesting issue in theoretical chemistry.[28–31,38–42,90–101] Because the gradient directions of the PES are uniquely determined, curves calculated by RGF to different directions $\mathbf{r}$ cross if and only if the gradient vanishes at the crosspoint, i.e. the crosspoint has to be a stationary point. However, different branches of the solution of the *same* reduced gradient curve with respect to $\mathbf{r}$ may also cross each other. These points are characterized as the branching points of the reduced gradient curve being the VRI points of the surface. Thus, the branching of an RGF curve (symmetric or unsymmetric) is connected with the emergence of special points of the PES. Whenever a reduced gradient curve reaches a VRI point, the curve branches, and at every VRI point of the PES the solution of a corresponding reduced gradient curve branches;[8] however, not every RGF curve has a BP. The path following of those RGF curves which have a BP allows one to find VRI points. Usually those curves should have a special symmetry in configuration space. In Fig. 7 there are four VRIs indicated by a bullet. The bold-faced, dotted RGF lines in y direction (curves with $E_x = 0$) show symmetric VRI points at the crossing with the $y \approx 0$ branch. The other kind, the nonsymmetric VRI points, are obtained by the central RGF curves of the thin bundles of curves. The characteristic attribute of every VRI point is the extra zero eigenvalue of the Hessian. At least one eigenvalue changes its sign when going along the gradient, where the corresponding eigenvector is orthogonal to the gradient. This means, a valley changes into a ridge or vice versa. VRI points are defined independently of the RP definition. Figure 7 shows that only special RGF curves bifurcate. The two other RGF curves of the three-bundle shown connect two extremal points of the PES of an index-difference of one: that is the general case. The special bifurcating RGF curve connects four extremal points of the PES where the VRI point is the knotty point in between. If one looks at the two branches of this special RGF solution, one may observe that one branch connects two SPs of index one, and the other branch connects the minimum of index zero with an SP of index two.

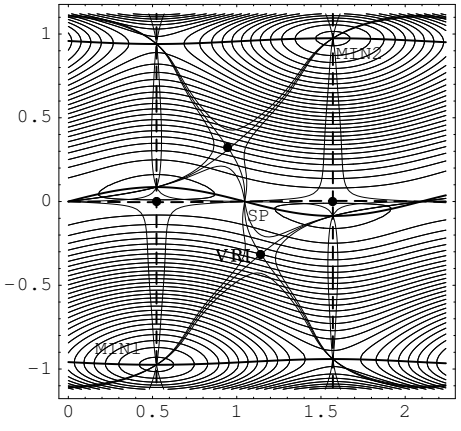

Fig. 7. RGF for PES $\{1\}$[51] with VRI points (•).

An RGF curve which reflects the symmetry of the PES, like the RGF in y direction in Fig. 7, may serve as a good model of an RP for the vibrational excitation along the second mode of the minimum MIN1, the wagging. The SP in uphill direction is an SP of index two.[102] In front of this SP, the pathway uphill bifurcates at the bullet and one branch leads to the indicated SP, a running pass (at the right-hand side, there is symmetrically also an SP at the left-hand side). In the contradictory case, it is not to assume that the special RGF line to the indicated nonsymmetric VRI point plays any special role for a reaction. It only marks the border line for the family of RGF curves which can reach the indicated SP. Those RGF curves which go uphill along a search direction with more y parts, more wagging — find the indicated SP of the inversion (see Fig. 6). The RGF curves with a higher part of x direction find the other SP of index one in the MEP valley at the line $y \approx -1$, the torsion, a low pass (see Fig. 5). We may define the "shortest" RGF

curve of the respective family to be the best model of the RP to the adjacent SP, note a quite different ansatz in subsection 2.8. Or, alternately, one may choose that special RGF curve which follows the gradient direction being equivalently to the decay vector of the SP, to represent the RP. None of these possible RPs touch at the nonsymmetric VRI point: the indicated VRI point in Fig. 7 marks the border line between two families of RGF lines leading to two different SPs. Also the IRC coming from the respective SP does not meet the nonsymmetric VRI point. If one stimulates an excitation at MIN1, where should one define the valley bifurcation for a decision to reach one of the two next SPs of index one? Is this BP at the minimum, or is it placed at the nonsymmetric VRI point? For the moment, we leave the answer open to the reader (or, you may jump to the Discussion and Fig. 13 in Sec. 4 below). *I guess that nonsymmetric VRI points will not play a central role in a theory of branching of chemical reactions.*

An inflection point of an energy profile over an RGF curve generally is not an extraneous singularity, in the sense of VRI, because $\mathbf{g}$ is not orthogonal to $\mathbf{e}_{\rm zero}$. The curvature of the energy profile along the MEP is not relevant for the VRI problem!

Difficult to imagine is a further important possibility. The VRI points may form a manifold in the configuration space of the molecule. Only for two-dimensional surfaces are these points isolated points like in Fig. 7. The PES of a molecule can have a manifold of VRI points with the dimension up to $n-2$, if the configuration space of the PES has the dimension n.[8,34,35] Because, at the VRI point, we have the two constraints only:

(i) the gradient of the PES is orthogonal to an eigenvector of the Hessian matrix of the PES,

$$\mathbf{g}^T\,\mathbf{e}_{\rm zero} = \mathbf{0}\,, \quad \text{and} \tag{13}$$

(ii) the corresponding eigenvalue of $\mathbf{e}_{\rm zero}$ is zero.

The RGF approach shows an important analogy to the mathematical theory of Branin,[32] the global Newton method.[34,35] It utilizes the adjoint matrix $\mathbf{A}$ of the Hessian matrix $\mathbf{H}$. This is defined as $((-1)^{i+j}m_{ij})^T$ where m_{ij} is the minor of $\mathbf{H}$ obtained by deletion of the ith row and the jth column from $\mathbf{H}$, and taking the determinant. If the Hessian is in two dimensions

$$\mathbf{H} = \begin{pmatrix} h_{xx} & h_{xy} \\ h_{xy} & h_{yy} \end{pmatrix}, \quad \text{then } \mathbf{A} = \begin{pmatrix} h_{yy} & -h_{xy} \\ -h_{xy} & h_{xx} \end{pmatrix}$$

is the adjoint matrix and it satisfies the important relation

$$\mathbf{HA} = \begin{pmatrix} h_{xx}h_{yy} - h_{xy}^2 & 0 \\ 0 & h_{xx}h_{yy} - h_{xy}^2 \end{pmatrix} = \mathrm{Det}(\mathbf{H})\,\mathbf{I}_2\,,$$

where $\mathrm{Det}(\mathbf{H})$ is the determinant of $\mathbf{H}$, and $\mathbf{I}_2$ is the 2D unit matrix. The adjoint matrix $\mathbf{A}$ is used to define an autonomous system of differential equations for the curve $\mathbf{x}(s)$, where s is a curve parameter:

$$\boxed{\frac{d\mathbf{x}}{ds} = \mp\,\mathbf{A}(\mathbf{x})\;\mathbf{g}(\mathbf{x})\,.} \tag{14}$$

Thus, the tangent of the curve of interest $\mathbf{x}(s)$ does not point in the direction of the gradient, as is the case when using the IRC. The tangent is the gradient $\mathbf{g}$ of the PES transformed by the adjoint matrix $\mathbf{A}$. The "+" option is used for searching stationary points with an odd index (SPs with an odd number of negative eigenvalues of the Hessian), where the "–" option searches for stationary points with an even index (minima, or SPs with an even number of negative eigenvalues of the Hessian).

The Branin method is a fine tool to find symmetric VRI points as exactly as we need them. Calculations using the Branin method can be done as follows: choose by trial and error the steplength parameter, l, and discretize Branin's differential equation (14) to

$$\mathbf{x_{m+1}} = \mathbf{x_m} \mp\; l\;\mathbf{A_m}\mathbf{g_m} \tag{15}$$

where m is the step number.[8,9,32,33,86] $\mathbf{A_m}$ is the adjoint matrix of the Hessian and $\mathbf{g_m}$ is the gradient at point $\mathbf{x_m}$. For example, we used an l value varying between 0.000 4 and 100 for the calculation of VRI points. In the case of the 4-atomic H_2CO, we used up to $l = 100$ units of the corresponding coordinate (Å, rad) for a satisfactory exploration along Branin pathways.[33] There, two-dimensional VRI manifolds are detected lying in the 3D subspace of symmetry coordinates of H_2CO. The l value depends on matrix $\mathbf{A_m}$ as well as on the gradient $\mathbf{g_m}$ at $\mathbf{x_m}$, and sometimes it has to be adapted during the calculation for a satisfactory exploration along the Branin curve. The product of adjoint matrix times gradient in (15)

becomes small near VRI points because at the VRI the gradient is an eigenvector of $\mathbf{A}$ and its eigenvalue is zero, see subsection 2.13 below. This causes smaller steps near the VRI, and if the parameter l is appropriately chosen, then a good convergence is obtained. Because of the self-correcting definition, there is no need for a corrector for the Branin method, because we are not interested in an exact curve following. In contrast, we usually search for the exact zero of the orthogonal eigenvalue of any Branin curve. Thus, the stopping criterion is the zero of the ridge eigenvalue: below 1 cm^{-1}, or something else, the algorithm may stop.

As already remarked, there is no need for an exact curve following. However, an important condition for the VRI search is the strict symmetry constraint. When starting anywhere in the coordinate space, a Branin curve may almost reach the VRI point, however, usually it turns off bypassing this point. Indeed, the point where a Branin solution turns off is often a TP of the corresponding RGF curve. The occurrence of a TP may suggest the nearby existence of a VRI point. This is shown in Figs. 2 and 7 in Ref. 8 for Branin solutions on a 3D hypersurface. To hold a special symmetry along a path, one should enforce a symmetrization of the next step in the Branin method.

The general behavior of Branin solutions is: they connect stationary points of a different index, or they end in a VRI point. Stationary points of the PES are limit points of the solution, because there $\mathbf{g} = \mathbf{0}$. However, there are further possible limit points, or fix-points, also in regions with $\mathbf{g} \neq \mathbf{0}$. These are points where

$$\mathbf{A}(\mathbf{x})\ \mathbf{g}(\mathbf{x}) = 0\,. \tag{16}$$

Because of the possible zero vector in Eq. (14), the bifurcation of solution curves can take place somewhere at the slope of the surface, where $\mathbf{g} \neq 0$. Equation (16) holds if $\mathbf{g}$ is an eigenvector with zero eigenvalue. There is sufficient evidence that the BP of a Branin curve is a VRI point.[21,47] Points which satisfy Eq. (16) are named *extraneous singularities*[34] of Eq. (14) because they are possible numerical perturbations of the search for stationary points. Because of the property of the adjoint

$$\mathbf{H}\ \mathbf{A} = \mathrm{Det}(\mathbf{H})\ \mathbf{I}_n\,, \tag{17}$$

where $\mathbf{I}_n$ is the n-dimensional unit matrix, we obtain for the nonsingular case, if $\mathbf{H}^{-1}$ exists, the system

$$\frac{d\mathbf{x}}{ds} = \mp\,\mathbf{H}^{-1}(\mathbf{x})\ \mathbf{g}(\mathbf{x})\ \mathrm{Det}(\mathbf{H}) \tag{18}$$

instead of Eq. (14). This represents a Newton step with a damping factor $\mathrm{Det}(\mathbf{H})$. Curves satisfying this expression are called Newton flows. Solution curves of Eq. (14) have a special character. Considering the behavior of the gradient $\mathbf{g}(\mathbf{x}(s))$ along a solution, $\mathbf{x}(s)$, we obtain with (14) and (17)

$$\frac{d\mathbf{g}}{ds} = \mathbf{H}\,\frac{d\mathbf{x}}{ds} = \mp\,\mathbf{H}\ \mathbf{A}\ \mathbf{g} = \mp\,\mathrm{Det}(\mathbf{H})\ \mathbf{g}\,. \tag{19}$$

Thus, the gradient $\mathbf{g}$ changes proportionally to $\mathbf{g}$. This means that the direction of $\mathbf{g}$ does not change. It is invariant along the solution. This means, the differential equation of Branin (14) has the same solution curve as the RGF Eq. (6). $\mathbf{x}'(s)$ is the tangent to the solution curve of Eq. (6).[8] On the other hand, $(n-1)$ orthogonal directions $\mathbf{e_i}$ to $\mathbf{g}$ can also be chosen constantly along a solution. Directional derivatives along these directions of the surface vanish because the surface is always orthgonal to its gradient.

$$\frac{\partial E}{\partial \mathbf{e_i}} = 0\,, \quad i = 2, \ldots, n\,. \tag{20}$$

This system of equations leads to the RGF, Eq. (4), if we use $\mathbf{g}$ and $\mathbf{e_i}$ as basis vectors of a coordinate system in Eq. (4).

From another point of view, the RGF equation gives an alternate definition of the Newton flows in comparison with Eq. (14). The two strategies: RGF and global Newton method, are slightly different with respect to their initial conditions: it is appropriate to detect unknown stationary points, for instance SPs of index one (transition states) by RGF. The method starts at a stationary point, e.g. a minimum, and follows an arbitrarily selected direction of the gradient on the PES. This may be a chemical interesting direction, a reaction path. On the other hand, the Branin differential equation, Eq. (14), may start anywhere on the PES but outside a stationary point using the gradient direction of that point. So, the Branin algorithm may easily stop anywhere, and continue using the gradient of that point.

If the search direction of a Branin curve does not exactly coincide with the direction of the gradient at

the next VRI point (which we search for), then the curve does not meet this VRI point, and we cannot grasp the VRI point.[8] Hence, we have to start at a point where the gradient has the same direction as the gradient at the VRI point, see the example of such a search direction in Ref. 8 This can be realized for manifolds of VRI points in symmetric subspaces of the internal coordinates in the configuration space. Therefore, a systematic search for VRI points is possible in symmetry hyperplanes of the PES.[8,9,33,86] In this case, along the pathway of a Branin curve, the eigenvalue of an eigenvector, being orthogonal to the gradient, converges to zero.

Figure 8 gives an example of two symmetric VRI points detected by the pure y-direction RGF line $E_x = 0$ (the bold-faced dotted curves). The surface **model {3}** is again the example of Ref. 51 being artificially modified in the last summand to $600\cos(6x)\cos(2y)$, and an additional summand of $-500\sin(3y)$ is used (see also Fig. 1). One may detect two different possibilities for the character of the VRI points:[8,103] on the left-hand col downhill from SP_3 there the valley bifurcates into two valleys leading to the two equivalent minima, MIN_1 and MIN_2. The central fork of this pitchfork bifurcation downhill is the ridge to SP_1. On the right-hand side, at VRI_2, the ridge coming uphill from SP_2 bifurcates into two ridges which lead to two saddle points of index two, to two summits. The central fork of the pitchfork bifurcation uphill is the valley to the SP_4. Consequently, on the RP between the SPs 2 and 4, there does not bifurcate any of the given RP definitions! Not every VRI point is connected with an RP bifurcation! At least from the point of view of the MEP definition by RGF curves, IRC, or GE: the latter two do not bifurcate at VRI_2, but the RGF curve bifurcates into "false" branches.

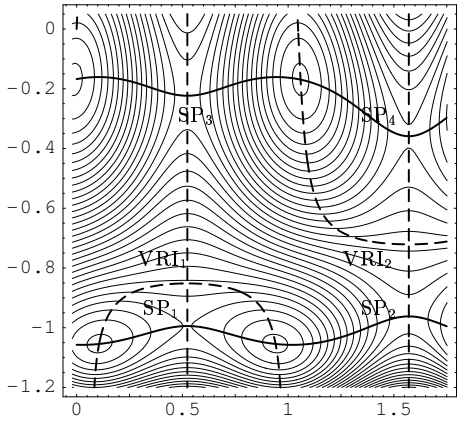

Fig. 8. RGF curves bifurcate at symmetric VRI points on surface {3}.

2.5. *Modification of RGF by following the tangent of the previous predictor step: TASC*

> Und ists Gefühl: wer weiß, wie weit es reicht
> und was es in dem reinen Raum ergiebt,
> in dem ein kleines Mehr von schwer und leicht
>
> Welten bewegt und einen Stern verschiebt.
> R. M. Rilke

We modified RGF to search for "true" MEPs.[21,47,48] Let $E(\mathbf{x})$ be the function of the PES, and $\mathbf{g}(\mathbf{x})$ its gradient vector in the configuration space, R^n. The RGF algorithm[8] uses a projection of the gradient of the PES to fulfill the system of equations

$$\mathbf{P_r}\ \mathbf{g}(\mathbf{x}(s)) = 0 \tag{21}$$

of rank $n-1$. The projector, $\mathbf{P_r}$, was chosen to be a constant matrix for RGF: one which enforces the gradient to point at every curve point, $\mathbf{x}(s)$, in the same direction $\mathbf{r}$. The tangent to this curve, $\mathbf{x}'(s)$, was obtained by a solution of the following system of linear equations (see (8) above)

$$\mathbf{P_r}\,\mathbf{H}\,\mathbf{x}' = \mathbf{0}\,. \tag{22}$$

The simplicity of the RGF method is based on the constancy of the $\mathbf{P_r}$ matrix. Now, the constant search direction $\mathbf{r}$ in Eq. (21) of the RGF method is replaced by a direction which is changed during the iteration process. To understand the idea we will look for an example. The model surface of **sample {4}** is

$$E(x,y) = x^2 - x^3 + \frac{1}{4}x^4 + \frac{1}{2}(x^2 - 1.7x + 0.6)y^2 + \frac{1}{4}y^4\,.$$

In Fig. 9 the valley (the GE — the fat curve) between the minimum and the SP is shown, as well as a family of RGF curves (dotted lines). Some of these RGF curves also connect the minimum and the SP of index one. Also shown is the x axis GE

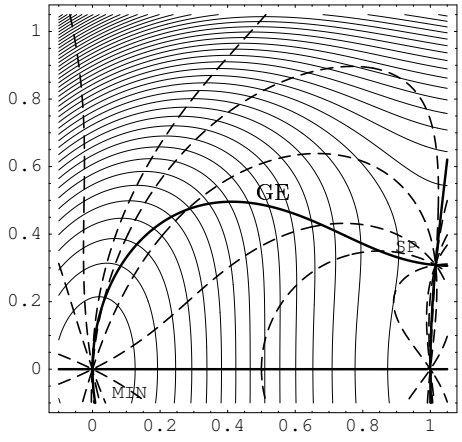

Fig. 9. RGF (dotted lines) and GE on model {4}.

going over the central hill at $(1, 0)$, an SP of index two,[102] and a ridge GE along $x \approx 1$. One may observe an interesting property of the RGF curves: their search direction, the gradient being orthogonal to the equipotential lines, does not coincide with their tangent. Usually, the tangent direction points "nearer" to the searched valley floor line (the GE) than the search direction itself. We propose to utilize the tangent of the searched curve itself as the new searched gradient direction in a modified RGF. This may sound like a vicious circle. But it is iteratively realizable because the RGF method is separated into predictor and corrector steps. We allow the projector to change after the predictor step: the tangent direction of the previous curve point iteratively becomes the search direction for the next point of the curve. The procedure is named the tangent search concept (TASC) — the task is: find the minimum path!

We define the projector with the unit vector $\mathbf{t} = \mathbf{x}'(s)/|\mathbf{x}'(s)|$, the normalized solution of (22)

$$\boxed{\mathbf{P_t} = \mathbf{I}_n - \mathbf{t}\,\mathbf{t}^T\,.} \tag{23}$$

But all calculations of the predictor-corrector method were done by Eqs. (21) and (22). In the tangent derivation for (22) we further assume a "constant" $\mathbf{P_t}$ matrix in the current step. Predictor and corrector will work locally to search for an RGF curve to the gradient direction $\mathbf{x}'(s)/|\mathbf{x}'(s)| = \mathbf{r}$. This is an approximation, but it works well and really accelerates the calculation of the streambed line. (See subsection 2.15 for a similar idea.) Usually, TASC changes the corresponding RGF curve after the predictor step. This results in a self-consistency on the valley floor GE (its definition: see subsection 2.9). Details are given in Refs. 21, 47 and 48 TASC is limited to follow the direction of the smallest eigenvalue. But the method often also works in cases where TPs of the streambed GE appear. Such regions are overcome by successive corrector steps.

In general, the TASC curve follows the valley floor. Using TASC, the diagonalization of the Hessian to calculate the lowest eigenvector is avoided. The aspect becomes computationally important for very large systems.[104] In contrast to the well-known method of eigenvector following,[105,106] the TASC method provides a locally-defined curve, found by a predictor-corrector scheme. Again, the success results from the self correction property. So, this pathway can be calculated as exactly as necessary by diminishing the steplength of the predictor and the threshold of the corrector.[48] In this manner, the path forms a fine approximation of that MEP following the smallest ascent starting from the minimum.

There is compelling proof that the method converges to the GE, if appropriate conditions are fulfilled.[48] Reported examples are valley pathways for H_2O,[21] and $C_2H_5^+$,[86] as well as Lennard-Jones clusters containing up to 55 atoms.[21] In the best case there are needed 15 predictor + 31 corrector steps to climb uphill in the large 55-cluster a broad valley to the SP at the top of the valley.[21]

2.6. *Possibility of Hessian update*

To calculate the different Hessians for RGF or TASC path points, the possibility of updating this expensive matrix should be used. There are two proposals: first we have the Davidon-Fletcher-Powell (DFP) update of the Hessian matrix.[4] The update works in the case when the index of the Hessian at the minimum changes into the SP index.[4,7] Of course, this is the condition for an update to serve for a search of pathways from minimum to SP. The second possibility is to use Bofill's update[107] which is well accepted in this field of computations (see an application in Ref. 89)

2.7. *Stochastically explored pathways*

There is an alternate way to calculate the RP: by stochastic means. The presence of a huge number of

minima in a large molecule can stall the progress of any deterministic algorithm. Stochastic algorithm based techniques for locating SPs and constructing reaction paths have been proposed recently.[108] The objective function of an optimization is chosen to contain information about local gradients (and curvatures). It is $E(\mathbf{x})$ the energy function, $\mathbf{g}(\mathbf{x})$ the gradient, and λ_i is the ith eigenvalue of the Hessian matrix. Usually, stimulated annealing, genetic algorithms, or other stochastic algorithms try to search for a minimal value of $E(\mathbf{x})$. Chaudhury and Bhattacharyya[108] defined an objective function

$$F(\mathbf{x}) = [E(\mathbf{x}) - E_P]^2 + \sum_{k=1}^{n} \beta_k \, g_k(\mathbf{x})^2 + \sum_{i=1}^{n} \gamma_i Exp(p_i \lambda_i) \qquad (24)$$

which allows one to explore the MEP. Here β_k, and γ_i are penalty weight factors, and the $p_i = \mathrm{Sign}(\lambda_i)$ are phase factors, and E_P is a guessed energy of the searched minimum, or SP, or of a corresponding RP point. E_P may serve as an estimated lower bound to the energy which we provide. This locates a critical point around E_P, if it is there. For locating a minimum on the PES the curvature constraint term is dropped.

A glance at the objective function expression (24) would suggest that for large systems the techniques mentioned might become costly. If at each step the Hessian matrix has to be constructed and diagonalized, the method might not be appealing for larger systems. One may note[16] that the full expression of the curvature term $k = \sum_{i=1}^{n} \gamma_i Exp(p_i \lambda_i)$ in (24) is not needed to enforce the curvature constraint since the higher terms do not contribute much in the overall expression. So, evaluation of the first eigenvalue is enough to guide the search (after eliminating the six zero eigenvalues associated with translation and rotation which has to be done carefully). With this simplification, the curvature constraint term in the objective function would still read $k' = \eta_1 Exp(p_1 \lambda_1)$, where λ_1 is the smallest eigenvalue of the Hessian associated with the vibrational modes of the molecule. Further simplification can be achieved by noting that along an MEP to an index one SP, the eigenvector corresponding to the negative eigenvalue of the Hessian should point nearly parallel to the gradient. The eigenvector concerned can therefore be approximated by the gradient itself (see also Fig. 12).

Now since only the first term, $\eta_1 e^{p_1 \lambda_1}$, in the curvature constraint remains, λ_1 is substituted by the simple difference of gradients

$$\lambda_{1\ \mathrm{new}} \approx |\mathbf{g}(\mathbf{x_l})| - |\mathbf{g}(\mathbf{x})| \qquad (25)$$

where $\mathbf{g}(\mathbf{x_l})$ is the local gradient at a second test point obtained by taking a small step from the given point along the direction of the local gradient: $\mathbf{x_l} = \mathbf{x} + l\,\mathbf{g}(\mathbf{x})$. Near an SP on an MEP, $\lambda_{1\ \mathrm{new}}$ should be negative. The curvature constraint term at the designated point on the surface becomes

$$k'' = \eta \, \mathrm{sign}(|\mathbf{g}(\mathbf{x_l})| - |\mathbf{g}(\mathbf{x})|)\, e^{p_1 \{|g(\mathbf{x_l})| - |g(\mathbf{x})|\}} \qquad (26)$$

where now η is the positive penalty factor which should lead the method to an SP. (One cannot exclude the possibility of obtaining an SP of a higher index.) Having established the objective function one now uses the stochastic optimizer, for example a genetic algorithm for exploring the search space. Searching uphill goes on by successive enlargement of the value of E_P, starting at the minimum, by small steps. Results for the MEPs of LJ cluster with 7 to 30 atoms are obtained by usually 600 generations of the genetic algorithm by the so-called gradient-only method (see Ref. 16).

2.8. *Variationally optimized reaction paths*

Liotard and Penot proposed early in the eighties another alternate way, a "line"-variational ansatz. The idea was to lay a cord over the mountains and put it down, at both ends, in the hope that the cord will approximate the SP.[109] The method starts with an approximate path, for example a line between two minima, and then refines or relaxes the points of the path until appropriate conditions are met. Stacho and Ban describe a procedure in which the points of the initial path are relaxed along the SD step for every point, and then a redistribution of the resulting points is done to maintain equal spacing.[110] Ayala and Schlegel refined the approximate path by optimizing the highest point to the SP and the remaining points have to satisfy the SD equation,[111] (see also Ref. 112).

An important innovation was made independently with the RP proposal by Elber and Karplus[17,113] and

by Pratt.[114] The aim is, for large molecules, not to use the expensive Hessian matrix: the PES itself is a barren landscape. Of sole interest are the lowest critical points. It is proposed to optimize a line integral S along a set of curve points $\mathbf{x}(s)$ again seen as the reaction path approximation:

$$S = \frac{\int_{M1}^{M2} E(\mathbf{x}(s)) \sqrt{\sum x^{i\,\prime\,2}} ds}{\int_{M1}^{M2} \sqrt{\sum x^{i\,\prime\,2}} ds}\,, \qquad (27)$$

where $M1$ and $M2$ are reactant and product of a reaction, respectively, and S is the functional to be optimized. The arc-length element of the line integral in the denominator gives the path length L. To be able to use such an ansatz in a computer calculation, Eq. (27) is discretized by

$$S(\mathbf{x}_0, \mathbf{x}_M; \mathbf{x}(s)) \approx \frac{1}{L} \sum_{i=1}^{M} E(\mathbf{x}_i) |\Delta \mathbf{x}_i|\,. \qquad (28)$$

$\mathbf{x}_0$ and $\mathbf{x}_M$ are fixed start and end points of the reaction, and $\mathbf{x}_i$ are points of the reaction path which are searched for. The intervals $\Delta \mathbf{x}_i$ are the differences $\mathbf{x}_i - \mathbf{x}_{i-1}$. To keep the points approximately equidistant, restraints are used for the intervals. The rigid-body translation or rotation of the molecule is suppressed by a penalty function. Additional forces are added to prevent the path from becoming kinked. The standard non-linear minimization of S is done. It requires only first derivatives of the PES, which makes the method suitable for large systems. The number of points on the path, M, is assumed to be small against the dimension n. (M is in the order of 10.)

The Euler-Lagrange equations of the first variation of Eq. (27) are for every curve point of a path which minimizes S

$$|\mathbf{x}'|\,, g_i - t^i \frac{dE}{ds} - (E - S)\,\frac{d}{ds}(t^i) = 0\,, \quad i = 1, \ldots, n\,,$$

where $\mathbf{t} = \{t^1, \ldots, t^n\}$ is the unit tangent of the RP. S is the value of the functional, all other entries in this equation are local values of the path. If the RP is assumed to be an SD, then the first two summands cancel and from the third summand it would follow that the path has to be a straight line with curvature zero. Thus, in this special case, the Elber–Karplus method approximates the SD from SP, but for a curvilinear valley, this pathway will also deviate from the SD. If the SD, coming from a strong side slope, reaches the valley floor with a "sharp" curvature then the Elber–Karplus path cuts the corner.[115] The connection of this method with the other RP definitions of this review is still mathematically mostly not understood. The method was applied to a conformational transition of the protein myoglobin with 1531 atoms,[17] and it gives pretty good information concerning an MEP which connects two minima. Further applications are given in Refs.115 and 116 and in references therein, as well as in the NEB method ("Nudged Elastic Band").[117] An elastic band stretching between the two endpoint configurations is optimized until the band traces out the MEP using a discrete representation which is largely decoupled from the motion of the band.

2.9. *Gradient extremal*

Pancíř[118] and Basilevsky/Shamov[11] were the first to formulated local criteria for describing a valley floor line. Pancíř determined two conditions which he assumed to be obviously given:

(i) The energy must increase along all directions perpendicularly to the direction of the valley floor line.
(ii) The curvature of the energy surface along the direction of the valley must be less than the curvature along any other direction.

Pancíř came to the conclusion that a path satisfying (i) and (ii) should be a sequence of points where the gradient, $\mathbf{g}$, is an eigenvector of the Hessian, $\mathbf{H}$.

If the norm of the gradient forms a minimum along points of an equi-subsurface, $E(\mathbf{x}) = \text{const.}$, i.e. along all directions perpendicular to the gradient,[11,12,14,119,120] a point of gentlest ascent of a valley is found. The measure for the ascent of the function $E(\mathbf{x})$ is the norm of the gradient vector, the functional

$$\sigma(\mathbf{x}) := \frac{1}{2} |\mathbf{g}(\mathbf{x})|^2\,. \qquad (29)$$

The implicit condition $E(\mathbf{x}) = c$ may be fulfilled by the sub-hypersurface $\mathbf{x}(\mathbf{u}, c)$, where $\mathbf{u}$ may be an $(n-1)$-dimensional parameter. We treat the parametric optimization problem with the objective function

$$\sigma(\mathbf{x}) \rightarrow \underset{\mathbf{x}(\cdot, c)}{Min}\ ! \qquad (30)$$

where the nonlinear constraint is $E(\mathbf{x}) = c$. Thus, objective function and constraint are developed from the function E itself. We are interested in following a path of local minima as the parameter increases (if we do an ascent to the surface) or decreases (if we go downhill). For almost all values of c one generally might expect that a local minimum $\mathbf{x}(c)$ of problem (30) depends differentially on c, mainly by virtue of the implicit function theorem. In the general case, the determination of a solution curve of a problem like (30), in dependence of c, needs third derivatives of objective function and constraints,[34] however, compare TASC! Using the normalized gradient

$$\mathbf{w}(\mathbf{u}, c) := \mathbf{g}(\mathbf{x}(\mathbf{u}, c))/|\mathbf{g}(\mathbf{x}(\mathbf{u}, c))| \tag{31}$$

and

$$\mathbf{P}_{\mathbf{w}(\mathbf{u},c)} := \mathbf{I} - \mathbf{w}(\mathbf{u}, c)\ \mathbf{w}(\mathbf{u}, c)^T\,, \tag{32}$$

the requirement for an extremal value of σ is expressed by

$$\mathbf{P}_{\mathbf{w}(\mathbf{u},c)} \nabla\ \sigma(\mathbf{x}(\mathbf{u}, c)) = 0\,, \tag{33}$$

with c = constant. Because of $\nabla\,\sigma(\mathbf{x}) = \mathbf{H}(\mathbf{x})\,\mathbf{g}(\mathbf{x})$, and setting $\lambda := \mathbf{w^T}\,\mathbf{H}\,\mathbf{w}$, it results in the basic eigenvector relation

$$\boxed{\mathbf{H}(\mathbf{x})\ \mathbf{g}(\mathbf{x}) = \lambda(\mathbf{x})\mathbf{g}(\mathbf{x})\,.} \tag{34}$$

The proportional factor $\lambda(\mathbf{x})$ is an eigenvalue of the Hessian matrix, and the gradient is its eigenvector. Curves $\mathbf{x}(c)$ defined by (34) consisting of such points on consecutive equi-hypersurfaces for different sections of increasing or decreasing c are termed *gradient extremals*.[12] Thus, we define that a point $\mathbf{x}$ belongs to the GE if the gradient of the PES at that point $\mathbf{x}$ is an eigenvector of the Hessian matrix. However, following a curvilinear curve of consecutive GE points implies that one actually does not move in the direction of the gentlest ascent.[12,14] The tangent to the GE differs again from its "search direction" being here the gentlest ascent. The GE Eq. (34) selects points of the configuration space having an extreme value of $\sigma(\mathbf{x})$ with respect to variations on equi-hypersurfaces. So, if $\sigma(\mathbf{x})$ has a minimum, the PES may have a valley-floor GE, however, it may also have a crest of a ridge. The extrema of $\sigma(\mathbf{x})$ can also be maxima or degenerate stationary points.[4,14,15] In Fig. 10 we show the GEs for the test functions of

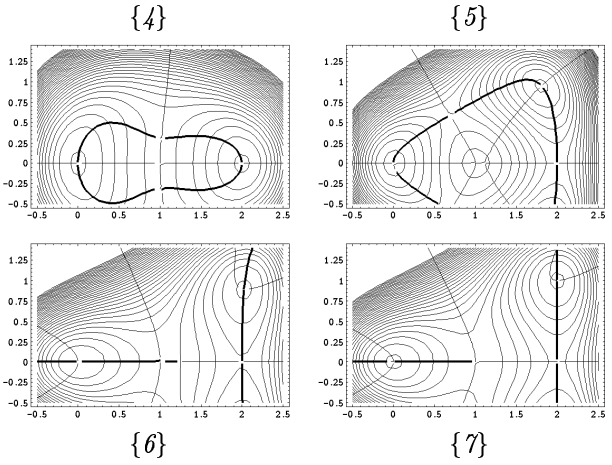

Fig. 10. Examples with valley GEs (bold-faced curves).

sample {4}, and for

$$\begin{aligned}
&\textbf{sample \{5\}}: \quad E(x,y) = \\
&x^2 - x^3 + \frac{1}{4}x^4 + \frac{1}{2}(x^2 - 3.0x + 1.25)y^2 + \frac{1}{4}y^4\,, \\
&\textbf{sample \{6\}}: \quad E(x,y) = \\
&x^2 - x^3 + \frac{1}{4}x^4 + \frac{1}{2}(x^2 - 4.2x + 3.6)y^2 + \frac{1}{4}y^4\,, \\
&\textbf{sample \{7\}}: \quad E(x,y) = \\
&x^2 - x^3 + \frac{1}{4}x^4 + \frac{1}{2}(x^2 - 4.0x + 3.0)y^2 + \frac{1}{4}y^4\,.
\end{aligned} \tag{35}$$

The valley-floor GE (fat lines) to the smallest eigenvalue follows the streambed of the surface very well, where the GE to the second eigenvalue follows a ridge or a cirque, or quickly deviates from its orthogonal direction to the valley line (in the cases {6} and {7}). Thus, GEs to higher eigenvalues are often not well adapted to follow the line of interesting mountain formations, or they are misleading. The two lower examples in Fig. 10 are exceptional in spite of the fact that a valley line goes (from the left minimum uphill) over the central SP and then finds a ridge. There is no valley GE to the two right minima directly from the central SP. By the way, there is also no direct IRC. But see subsection 2.2 for a satisfactory path definition in this case: there is an RGF curve which bifurcates at the symmetric VRI point.

In Fig. 11 the special RGF curve $E_y = 0$ is drawn for example {6}. It may serve as a model for an RP

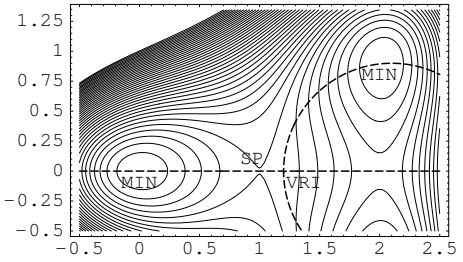

Fig. 11. RGF(dotted curve) bifurcates at the VRI.

for a reaction from the left minimum to one of the two right minima, or vice versa. The pitchfork bifurcation of the RGF curve well reflects the "quasi invisible" emergence of the side valleys at the VRI point. To take such a branching pathway for a model of an RP seems, on the other hand, still somehow questionable because a reaction will not take place "around a corner" of such a bifurcation. (One may observe that the case {7} of Eq. (35) in Fig. 10 is still more special, because the central SP itself is the VRI point of the surface. Note: it is not a monkey SP (see also the extended treatment in Ref. 31).

One can connect the GE with the steepest descent: using the arc-length s for the curve parameter, a general *steepest descent* curve $\mathbf{x}(s)$ is defined by

$$\frac{d\mathbf{x}(s)}{ds} = -\frac{\mathbf{g}(\mathbf{x}(s))}{|\mathbf{g}(\mathbf{x}(s))|} =: \mathbf{w}(s),. \tag{36}$$

Its *curvature vector* is defined by

$$\mathbf{k} := \frac{d^2\mathbf{x}}{ds^2} = \frac{d}{ds}\,\frac{d\mathbf{x}}{ds}\,. \tag{37}$$

We quickly find that GEs consist of points where SD lines have zero curvature[15,119,120] (compare also Fig. 14). Because, with $\mathbf{P}_{\mathbf{w}(s)} = \mathbf{I} - \mathbf{w}(s)\ \mathbf{w}(s)^T$ we have

$$\mathbf{k} = -\frac{\mathbf{P}_{\mathbf{w}(s)}\,\mathbf{H}\,\mathbf{w}(s)}{|\mathbf{g}|}\,. \tag{38}$$

The vector $\mathbf{k}$ is perpendicular to the gradient $\mathbf{g}$ due to the action of the projector $\mathbf{P}_{\mathbf{w}(s)}$, and its absolute value is the scalar curvature:

$$k = (\mathbf{w^T}\,\mathbf{H}\,\mathbf{P_w}\ \mathbf{H}\ \mathbf{w})^{1/2}/|\mathbf{g}|\,. \tag{39}$$

It is zero iff Eq. (34) is fulfilled.

The development gives rise to the formulation of the streambed description of the valley ground GE, which follows the smallest eigenvalue: if we are on this gradient extremal, then from the left as well as from the right-hand side, the SD lines confluent to this valley line[15] (see Discussion in Sec. 4). The GE forms an isolated curve in the configuration space. It does not form fields of curves as do the SD lines. But if the lines of two different families of SD curves confluent into the GE, from the right as well as from the left, it can serve as a model of the valley floor.

2.10. *The Sun-Ruedenberg approximation of the GE*[15]

A point showing the shallowest ascent of a valley is defined by the condition that the gradient is the eigenvector of the Hessian,[11,12] Eq. (34). To follow the GE is a task which is different from the task for an SD: the latter always runs downhill, in any point. But in a general point, there does not exist a GE. To follow a GE is only useful if we are on such a curve and will go a next step further. The method of choice is the predictor-corrector strategy. If $\mathbf{x}(s)$ is the GE with parameter s, then a differentiation of the GE Eq. (34) to the parameter s gives

$$\left(\frac{\partial\mathbf{H}}{\partial\mathbf{x}}\,\mathbf{x}'\right)\cdot\mathbf{g} + \mathbf{H}\cdot\left(\frac{\partial\mathbf{g}}{\partial\mathbf{x}}\,\mathbf{x}'\right) = \left(\frac{d\lambda}{ds}\right)\,\mathbf{g} + \lambda\left(\frac{\partial\mathbf{g}}{\partial\mathbf{x}}\,\mathbf{x}'\right).$$

Using the matrix $\partial\mathbf{H}/\partial\mathbf{x} = \mathbf{T}$ of third derivatives, the relation is to obtain

$$(\mathbf{T}\,\mathbf{x}')\,\mathbf{g} + \mathbf{H}^2\,\mathbf{x}' = \frac{d\lambda}{ds}\,\mathbf{g} + \lambda\mathbf{H}\mathbf{x}'$$

for the tangent $\mathbf{x}'$ of the GE equation. If the third derivative term is not zero, the tangent cannot be an eigenvector of the Hessian. The tangent differs from the gradient by a term depending on the third derivatives.[120] Any attempt to follow the GE by simply stepping along the gradient is incorrect even in the first-order term. The algorithm of Sun and Ruedenberg[15] constructs the tangent to the GE and takes a step along the tangent to produce a predictor geometry. Quite analogous to the RGF, or TASC methods above, at the predictor geometry a corrector step is calculated which brings the point back to the GE. However, the condition for a GE Eq. (34) is

already an equation which contains second derivatives of the PES. Thus, the derivatives for a tangent require knowledge (of at least one component) of third derivatives of the PES. This is very expensive! It may be calculated from two Hessians at two slightly displaced geometries.

We put $\mathbf{w} = \mathbf{g}/|\mathbf{g}|$, and write the GE equation

$$\mathbf{0} = (\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathbf{T}})\mathbf{H}\ \mathbf{w}/|\mathbf{g}| =: \mathbf{f}/\mathbf{g}^{\mathbf{T}}\mathbf{g} \qquad (40)$$

using the setting

$$\mathbf{f} = (\mathbf{I} - \mathbf{w}\mathbf{w}^{\mathbf{T}})\mathbf{H}\ \mathbf{g} = \mathbf{P}_{\mathbf{w}}\ \mathbf{H}\ \mathbf{g}\,.$$

Equation (40) is a special case for the value of the curvature k of SD lines where $k = \mathbf{f}/\mathbf{g}^{\mathbf{T}}\mathbf{g}$. The loci with $k = 0$ are crossed by GE curves. The tangent to $\mathbf{f} = \mathbf{0}$ is given by

$$\mathbf{df} = \sum_{i=1}^{n} \left(\frac{\partial \mathbf{f}}{\partial x^i}\right) dx^i = l \sum_{i=1} \left(\frac{\partial \mathbf{f}}{\partial x^i}\right) t^i = 0 \qquad (41)$$

with $\mathbf{dx} = l\,\mathbf{t}$, and $\mathbf{t}$ is the unit vector of the tangent at the GE. If one assumes a new coordinate system, denoted by $\mathbf{y}$, where the first coordinate is the gradient direction, and the other form the $n-1$ remaining eigenvectors of the Hessian (with the diagonal entries $\epsilon_i = \lambda_i$ being the eigenvalues if the point is on the GE). Let $\mathbf{y}_0$ be a point on the GE. The tangent in $\mathbf{y}_0$ can be obtained by the solution of the system of equations:

$$\sum_{i=2}^{n} Q_{ki} t^i = B_k t^1\,, \quad k = 2, \ldots, n \qquad (42)$$

with the new symbolic coefficients

$$Q_{ki} = \frac{\partial H_{ki}}{\partial\, y_1}|\mathbf{g}| + \delta_{ki}\,\epsilon_k(\epsilon_k - \epsilon_1)\,, \qquad (43)$$

and

$$B_k = -Q_{k1} = -\frac{\partial H_{k1}}{\partial\, y_1}|\mathbf{g}|\,. \qquad (44)$$

The H_{ki} and ϵ_i are to calculate in the $\mathbf{y}$ coordinate system. The value t^1 in (42) may initially be put to one, and it is determined by normalization if a solution of (42) was obtained. Going along the tangent to point $\mathbf{y_p}$, one may leave the GE.

$$\mathbf{y_p} = \mathbf{y_0} + l\,\mathbf{t}\,.$$

Here, l is the steplength. The corrector step $\mathbf{c}$ back to the GE is

$$c_i = \sum_{k=2}^{n} [\mathbf{S}^T(\mathbf{S}\,\mathbf{S}^T)^{-1}]_{ik}\ H_{k1}\ |\mathbf{g}|\,, \quad i = 1, \ldots, n$$

with the $\mathbf{S}$ matrix

$$S_{ki} = \frac{\partial H_{ki}}{\partial\, y_1}|\mathbf{g}| + (\mathbf{H}^2)_{ki} - H_{11}\,H_{ki}\,, \quad k = 2, \ldots, n\,.$$

On the GE the S_{ki} entries reduce to Q_{ki}. There are $(n-1)$ equations for n variables $t^1, \ldots, t^n$ which determine the direction of $\mathbf{t}$. If matrix $\mathbf{Q}$ from (43) is not singular, one has the tangent vector to the GE

$$\mathbf{t} = \{1,\, (\mathbf{Q^{-1}B})_2, \ldots, (\mathbf{Q^{-1}B})_n\}^T\, t^1\,, \qquad (45)$$

and t^1 again is the factor for the normalization.

Remark. It is clear that the tangent to the GE and the tangent to an SD path (= $\mathbf{g}$) coincide if

$$B_k = \left(\frac{\partial^3 E_k}{\partial y_1^2 \partial y_k}\right) = 0\,, \quad \text{for all } k\,. \qquad (46)$$

From the calculation of the matrices $\mathbf{Q}$ and $\mathbf{B}$ it is clear that $\mathbf{t}$ needs third derivatives of the PES. However, they are "only" of the kind

$$\frac{\partial^3 E_k}{\partial y_k \partial y_i \partial y_1}\,, \quad i, k = 2, \ldots, n\,, \qquad (47)$$

thus only $n(n+1)/2$ such derivatives are necessary. Per definition, y_1 should point to $\mathbf{g}$. One can reduce in that way the derivatives to one further calculation of a Hessian, if one searches (H_{ki}) at a point on y_1, that means on $\mathbf{w}$

$$\tilde{\mathbf{x}}_\mathbf{0} = \mathbf{x_0} + l \cdot \mathbf{w} \qquad (48)$$

Applications and comparisons of the method are reported by F. Jensen using Ar_8 as an LJ cluster, and H_2CO.[20,43,63]

2.11. *GEs and VRI*

There is a difference when GE Eq. (34) is compared to the VRI orthogonality requirement of Eq. (13). The VRI point satisfies the necessary condition of orthogonality of the two vectors included in its definition Eq. (13) for an eigenvector $\mathbf{e_{zero}}$ of $\mathbf{H}$ with zero eigenvalue and for the gradient $\mathbf{g}$. A VRI point is simultaneously the bifurcation point of the two

branches of an RGF curve. In the GE definition (34) the gradient has to be an eigenvector itself, while in Eq. (13) it only has to be orthogonal to a special other eigenvector. Thus, the gradient need not be an eigenvector at a VRI point (in more than 2 dimensions).

Equation (34) forms a system of n equations of rank $n-1$ with n unknown variables $x^1, \ldots, x^n$. For curve tracing, we need therefore $n-1$ independent equations. From the extremal definition of GE, it follows that the $n-1$ directional derivatives have to vanish:

$$GE_i(\mathbf{x}) := \frac{\partial \sigma}{\partial \mathbf{e_i}} = 0\,, \quad i = 2, \ldots, n\,. \tag{49}$$

In Eq. (49), analogous i directional derivatives orthogonal to the gradient of E($\mathbf{x}$) are used, as well as in Eq. (20), but now the gradient is an eigenvector, and the $\mathbf{e_i}$ also may be eigenvectors.

Despite the computational problems, GEs are often used to describe valley ground pathways and their branching. However, BPs of a GE are in general not identical with the VRI points of the PES.[8,14] On the other hand, if a GE meets a symmetric VRI point, then the gradient is an eigenvector of $\mathbf{H}$ by condition (34) of the GE. If the side branches of the corresponding RGF curve describe valleys, then they may serve seriously for the RP branches which emerge from the VRI point (see Figs. 8 and 11).

2.12. *GEs and TASC*

If we change the projector of RGF after every predictor step with the tangent direction of the previous curve point, from Eq. (22), so that $\mathbf{t}$ iteratively becomes the search direction (23) then the procedure is the TAngent Search Concept, TASC. An RGF curve crosses the GE (see Fig. 9) if its tangent $\mathbf{t}$ is parallel to the gradient of the PES. On the GE, on the other hand, the gradient is an eigenvector of the Hessian. If the tangent of an RGF curve coincides with the gradient, then Eq. (22) is fulfilled. Thus, in the limiting case for "infinitely small" predictor steps the numerical TASC procedure should lead to a GE. To verify this, we define a variable projector by the dyadic product

$$\mathbf{P_{A\,g}} = \mathbf{I} - \frac{(\mathbf{A\,g})(\mathbf{A\,g})^{\mathbf{T}}}{|\mathbf{A\,g}|^2}\,. \tag{50}$$

A point $\mathbf{x}$ where the gradient $\mathbf{g}(\mathbf{x})$ fulfills

$$\mathbf{P_{A\,g}}\ \mathbf{g} = 0 = \mathbf{g} - \mathbf{A\,g}\frac{(\mathbf{g^T\,A\,g})}{|\mathbf{A\,g}|^2} \tag{51}$$

belongs to a gradient extremal.

Proof. Instead of RGF ansatz (51), we use the equivalent Branin Eq. (14). The projector of the original RGF was $\mathbf{P_r}$ to the search direction, $\mathbf{r}$. We replace the constant direction (for a predictor step) by the direction of the tangent, $\mathbf{x}'$ in (23). If we search for a solution curve of (21) with direction $\mathbf{A\,g}$, thus (50), this becomes Eq. (51). Multiplication by $\mathbf{H}$ from the left-hand side is with Eq. (17)

$$\mathbf{H}\ \mathbf{g} = \mathrm{Det}(\mathbf{H})\ \mathbf{I}\ \mathbf{g}\ \frac{(\mathbf{g^T\,A\,g})}{|\mathbf{A\,g}|^2}\,. \tag{52}$$

$\mathrm{Det}(\mathbf{H})\,\mathbf{I}$ is commutative with all other terms and can change its place into the product $(\mathbf{g^T\,A\,g})$, and then we can replace it back to $\mathbf{A}\ \mathbf{H}$. Thus, the expression on the right-hand side, without one $\mathbf{g}$, is a scalar. If we denote it by λ, we obtain the known eigenvector equation (34), which is the equation of the GE.[12] With $\mathbf{v} := \mathbf{A\,g}$ the eigenvalue becomes[14]

$$\lambda = \frac{\mathbf{v^T\,H\,v}}{|\mathbf{v}|^2}\,. \tag{53}$$

2.13. *The action of TASC*

If the predictor steplength is not "zero", the resulting curve of TASC approximates the valley floor. We find TASC works well along the direction of the eigenvector with the smallest (absolute) eigenvalue. Of course, the best proof is the numerical success of the method.[21] To see it theoretically with sufficient evidence, we again use the equivalent Branin differential Eq. (14) which has the same solution as the method of RGF. If $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the eigenvectors of $\mathbf{H}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ 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

$$\mathbf{H}\ \mathbf{e}_i = \lambda_i\ \mathbf{e}_i\,, \tag{54}$$

and, by multiplication with the adjoint matrix, we get

$$\mathbf{A}\ \mathbf{H}\ \mathbf{e}_i = \mathrm{Det}(\mathbf{H})\ \mathbf{e}_i = \lambda_i\ \mathbf{A}\ \mathbf{e}_i\,, \tag{55}$$

with the basic property

$$\mathrm{Det}(\mathbf{H}) = \prod_{k=1}^{n} \lambda_k \,. \tag{56}$$

If a point of the solution curve of the RGF method with the search direction $\mathbf{r}$ is reached, the gradient of Eq. (14) points in the same direction. Expressing $\mathbf{r}$ by the eigenvectors

$$\mathbf{r} = \sum_{i=1}^{n} r_i \; \mathbf{e}_i \tag{57}$$

we obtain the relation for the tangent direction of an RGF = Branin curve

$$\mathbf{x}' = \mathbf{A} \; \mathbf{r} = \sum_{i=1}^{n} r_i \; \left(\prod_{k \neq i} \lambda_k \right) \mathbf{e}_i \,. \tag{58}$$

Let λ_1 be the smallest (absolute) eigenvalue and $r_1 \neq 0$. The $\mathbf{e}_1$ component of the preceding search direction $\mathbf{r}$ is enforced, if in the next step, the new direction $\mathbf{x}'$ of (58) is used in Eqs. (21) and (22). Thus, if the new search direction is the tangent of the former RGF curve, this direction is now turned nearer to the $\mathbf{e}_1$ eigenvector direction of the Hessian. The action is the greater 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 valley floor depends on the ratio of the extremal eigenvalues of $\mathbf{H}$, but it is also dependent on the entire matrix spectrum.

Remarks. Formula (58) allows the eigenvector following along the eigenvector of the smallest eigenvalue λ_1, which is automatically realized by TASC. Under this procedure, the calculation of the first eigenvector is not necessary. In contrast, we may use the product[14,15,119]

$$\lambda_1 = \mathbf{g}^{\mathbf{T}} \, \mathbf{H} \, \mathbf{g} / |\mathbf{g}|^2 \tag{59}$$

to guess the smallest eigenvalue on a TASC pathway, and the gradient is the corresponding eigenvector; see also the equivalent formula (53). It is necessary for the use of (59) that we have calculated exactly the streambed line, by small predictor steps and a small threshold of the corrector.

If in expansion (57) the r_1 becomes exactly $r_1 = 0$ then the action of the eigenvector weighting for the smallest λ_1 does not work. (For example, if in the course of a TASC search the direction $\mathbf{e}_1$ becomes orthogonal to a symmetry plane where TASC searches.)

2.14. *Perspectives of TASC*

TASC is a direct direction method: it follows straight forward the ground line of the streambed of a hypersurface, downhill or uphill, if such a streambed exists. The workability of the TASC algorithm to follow the streambed GE is demonstrated.[21,86,89] Tests are done on highly coupled problems with strong nonlinearity. The method performs well in practice. For minimum optimization, one can start at any point in the catchment region of a minimum and follow the gradient down the slope to the streambed, the "minimum path". Then it is to follow this path in the direction of the smallest eigenvector by TASC. The procedure is a potent method for studying the streambed of a multidimensional hypersurface. The streambed is calculable as exactly as we need it. The success of TASC is based on the tracing of the minimum path, which we geometrically understand as the valley floor GE. TASC with projector (23) does not follow other normal modes along cirques or cliffs: the proof in the previous subsection shows that the valley direction to the smallest eigenvalue is followed.

There are other well-known methods of eigenvector-following for the search of SPs.[105,106] The methods are more likely to "jump" over the PES than follow a leading line. They are successful, and their success uses to some extent the accident. The success is demonstrated for high-dimensional LJ clusters with an overwhelming number of minima and SPs.[106]

2.15. *A gradient-only search for the SP: IRC uphill?*

The direct inversion of the IRC search direction, to go uphill, is (numerically) impossible[6,62] — there is no hope following the "good" path upward from either reactant or product. In this subsection we report on algorithms which allow one to go uphill along a valley bottom in the direction of the weakest ascent. If an SP is at the top of the valley then it might be found. The energy profile over the reaction path should be a "valley floor" of the PES leading through the point of highest energy, the SP of index one. There are two

methods proposed for a "gradient-only" scheme: a minimization of an equation including the gradient,[120] and a controlled zigzagging procedure.[61]

Schlegel detected[120] that leaving the GE along the gradient direction in a GE point $\mathbf{x_0}$ does not change the direction of the gradient, in first-order (see Fig. 12). If again we put $\mathbf{w} = \mathbf{g}/|\mathbf{g}|$, and do a step $\mathbf{x_1} = \mathbf{x_0} + l\,\mathbf{w_0}$, then we can develop the gradient in $\mathbf{x_1}$ to

$$\begin{aligned}\mathbf{g_1} &= \mathbf{g_0} + \mathbf{H_0}\,(\mathbf{x_1} - \mathbf{x_0}) + O(l^2) \\ &= \mathbf{g_0} + \frac{l}{|\mathbf{g_0}|}\mathbf{H_0}\,\mathbf{g_0} + O(l^2)\end{aligned}$$

and using the GE equation this gives

$$= (1 + \lambda l/|\mathbf{g_0}|)\,\mathbf{g_0} + O(l^2)\,.$$

Thus, $\mathbf{g_1}$ is parallel to $\mathbf{g_0}$. This can be rewritten in an equation

$$\mathbf{g_1} - (\mathbf{g_1}^T\mathbf{g_0})\,\mathbf{g_0}/|\mathbf{g_0}|^2 = 0\,.$$

The zero search of this equation can be done by standard methods if a steplength l is set. It seems, however, that it will require evaluations of many points which make this procedure less attractive than the tangent-based algorithms described above.

Quite an analogous idea was proposed in Ref. 61 We should realize that the direction of the scaled gradient vector of the point $\mathbf{x}_0$ pointing to $\mathbf{x}_1$ does not agree with the direction of the gradient vector in $\mathbf{x}_1$, in the general case. However, on the GE this may happen (see Fig. 12). In the points $\mathbf{x}_0$, $\mathbf{x}_{0l}$, and $\mathbf{x}_{0r}$ the gradients are drawn. In points $\mathbf{x}_1$, $\mathbf{x}_{1l}$, and $\mathbf{x}_{1r}$ the negative gradients are used. In case of the fall in either of the two directions $\mathbf{g}(\mathbf{x}_0) = -\mathbf{g}(\mathbf{x}_1)$, we assume that we are on a valley pathway. Controlling for coincidence in the two gradients is achieved by means of the scalar product of the normalized vectors. A point $\mathbf{x}$ belongs to a q-minimum energy path (qMEP) if the vector equation holds

$$\mathbf{g}(\mathbf{x})/|\mathbf{g}(\mathbf{x})| = \mathbf{g}(\mathbf{x_q})/|\mathbf{g}(\mathbf{x_q})|\,, \qquad (60)$$

where

$$\mathbf{x_q} = \mathbf{x} + q\;\mathbf{g}(\mathbf{x})/|\mathbf{g}(\mathbf{x})|\,,$$

and $q > 0$ is a steplength parameter for an uphill search. In comparison to Eq. (5), this is an intensification. (The definition is also satisfied by points on a "q-ridge".)

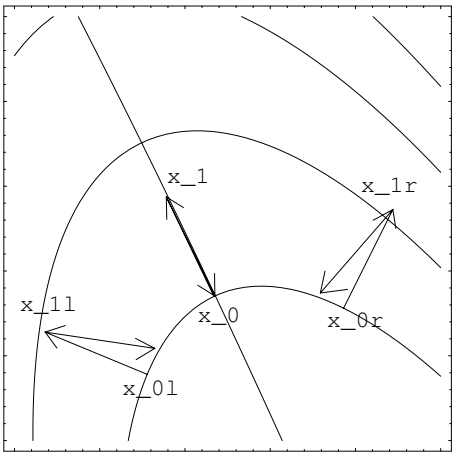

Fig. 12. Scheme of gradient vectors on and near the valley floor. $-\mathbf{g}(\mathbf{x}_{1l})$ and $-\mathbf{g}(\mathbf{x}_{1r})$ are used for corrector steps.

The definition (60) contains a local characterization of certain points $\mathbf{x}$. Thus, the definition does not need an initial condition of the pathway, as opposed to the IRC, which does need an SP.[18,36] The definition compares differences of gradient vectors. However, it does not use higher derivations, as it is necessary using the local definition of GEs.[12,15,121] If q is sufficiently small, and if the point $\mathbf{x}$ fulfills the condition of Eq. (60) then we are near the so-called asymptotic steepest descent path,[22] a line defined by the confluence of many SD lines from the right and from the left-hand side into the streambed of the valley ground. The valley floor may be curvilinear. The qMEP will follow its curvature. Points next to the qMEP are shown in Fig. 12. Gradients at these points have a characteristic pattern. If the point $\mathbf{x}_{l0}$ is to the "left" of the qMEP, then the negative gradient of $\mathbf{x}_{l1}$ points a step to the right. And vice versa, if the point $\mathbf{x}_{r0}$ is displaced to the "right" then the negative gradient of $\mathbf{x}_{r1}$ points to the left. Thus, these negative gradients may be used as corrector steps. They work in a procedure to follow the qMEP.

With two parameters: steplength $l = q$, and tolerance ϵ, the algorithm works well if ϵ is two or more powers of ten smaller than q. Results are given in Ref. 61 for small test cases; they show the theoretical possibility to work with such an algorithm, but they also show analogous problems like the SD

itself: the strategy has a strong affinity to zigzagging along the pathway. To suppress this, we have to use very small steps q and very restrictive corrector thresholds. The algorithm needs at least 1800 steps to go uphill in an LJ-cluster of Ar_4 from tetrahedron minimum to diamond SP. The procedure is a local working predictor-corrector algorithm using "controlled" zigzagging of gradient steps uphill along the qMEP. The algorithm is very simple. It works in any coordinate system, in Cartesian or internal coordinates. All degrees of freedom are automatically taken into account. The algorithm will not be disturbed by zero eigenvalues because it does not need a matrix inversion. The algorithm needs a high number of steps to find the exact localization of the SP in comparison with methods of second order. Thus, as all other gradient methods do, the proposed method suffers from slow convergence to the stationary point.[122] It is well-known that this may be greatly improved by incorporating a method of higher order.[123] However, if the Hessian of the PES is available, we do not need any first order method at all.

In the limit for $q \rightarrow$ zero, the qMEP will converge to the GE. If the GE does not form a pathway between a minimum and a nearby SP, however, if it is separated by TPs which hide the valley structure of the PES, then the difference between a qMEP with a non-zero q-value and this GE is that the qMEP can bridge the gap between the TPs and goes up to a SP of the potential.[61]

3. Invariance of Definitions from the Coordinate System

Reaction paths are a widely used concept in theoretical chemistry. Properties of such paths should be independent of the coordinate system employed. If this were not the case then we would obtain different dynamic and thermodynamic properties from different coordinates. This seems like a trivial point, and we are sorry to belabor it. It is evident that the invariance problem, which has been solved mathematically a long time ago (report[6]), again and again penetrates the discussions in this field.[45,46,124,125]

The geometrical arrangement of the atoms of a molecule in the space R^3 can be computed in a definite mathematical way and used to obtain the electronic energy of exactly this shape of the molecule. If we change the molecular structure, we will get a different energy. Thus, the potential energy surface (PES) emerges as the result of computations as a hypersurface over the configuration space of the molecule. The geometry of every molecular structure clearly corresponds to a particular molecular electronic energy, and these energies are independent of the kind of coordinates in the configuration space of the molecule. Analogically, we can go the next step: we define by a pure mathematical concept a "pathway" of changing the molecular structure from one special point of its configuration space to another point. A definite energy of the molecule belongs to any point along this hypothetical pathway. From a mathematical point of view, it is clear that we can define this pathway as being independent of the choice of coordinates in the configuration space of the molecular system.[126] Perhaps, some confusion concerning the invariance problem[124] comes from the fact that the usual concepts for defining reaction paths use the properties of the PES in a concrete coordinate system. But a change in the coordinate system by means of a definite transformation formula can always be compensated for by changing the method for the computation of the reaction path by an inverse transformation formula. What is the original pathway? In general, this is, and remains, a question of convention. It depends on the purpose of the investigation. For instance, a mass-weighted Cartesian system is well suited, if we search for chemical reaction pathways. It is an isoinertial system, and it is useful for dynamic calculations as a natural continuation of the spectroscopic treatments of vibrations and force constants.[127]

In a reaction, each atom describes its own pathway in the 3D Cartesian space, and the total movement of N atoms of the molecular system defines the migration of a point in the configuration space R^{3N}, say by coordinates y^i, $i = 1, \ldots, 3N = n + 6$. One needs coordinate invariance for the mass-weighted internal coordinates $x^k = x^k(y^i)$ throughout. They are given by a $\mathbf{z}$ matrix. The PES $E(x^1, \ldots, x^n)$ is a scalar depending on the coordinates x^i of the molecule. The coordinates are assumed as the contravariant vector $\mathbf{x} = (x^1, \ldots, x^n)^T$. Each derivation of $E(\mathbf{x})$ to x^i, with $i = 1, \ldots, n$, yields a component of a vector, termed the gradient vector of E. $\partial E/\partial x^i$ takes the ith place in the gradient $\mathbf{g} = (g_1, \ldots, g_n)^T$

which is a covariant vector.[6,37] The corresponding $\mathbf{B}$ matrix[127] is

$$\mathbf{B} = \left(\frac{\partial x^k}{\partial y^i}\right), \quad k = 1, \ldots, n\,, \quad i = 1, \ldots, n+6\,, \tag{61}$$

with n rows and $(n+6)$ columns. It serves as a linear transformation

$$\mathbf{g}_C = \mathbf{B}^T\,\mathbf{g}_I \tag{62}$$

for the change of the gradients in Cartesian or internal coordinates, and

$$\mathbf{H}_C = \mathbf{B}^T\,\mathbf{H}_I\mathbf{B} + \Gamma \tag{63}$$

for the change of the Hessians. The Γ term comes out of the chain rule to

$$(\Gamma_{ij}) = \left(\sum_k (\mathbf{g}_I)_k \frac{\partial^2 x^k}{\partial y^i \partial y^j}\right). \tag{64}$$

Explicit derivatives of the $\mathbf{B}$ matrix to the Cartesian coordinates are given in Ref. 129 The backtransformation to the gradient $\mathbf{g}_I$ of (62) can go on by first a multiplication from the left-hand side with $\mathbf{B}$ and then with the use of the inverse of the regular $(\mathbf{B}\,\mathbf{B}^T)$ matrix:

$$(\mathbf{B}\,\mathbf{B}^T)^{-1}\,\mathbf{B}\mathbf{g}_C = \mathbf{g}_I\,. \tag{65}$$

The transformation matrix on the left-hand side is the "left inverse" of $\mathbf{B}^T$; it is called the pseudoinverse matrix $\mathbf{B}^+$. It is also an $(n \times n+6)$ matrix. The contravariant metric tensor $\mathbf{G}^{-1} = (g^{ij})$ is calculated point by point along the pathway of an MEP taking $\mathbf{B}\,\mathbf{B}^T$ where the usual metric tensor $\mathbf{G} = (g_{ij})$ forms its inverse matrix. (There is a misprint in Ref. 9) The pseudoinverse is $\mathbf{B}^+ = \mathbf{G}\,\mathbf{B}$. The gradient $\mathbf{g}_C$ and the Hessian matrix $\mathbf{H}_C$ are calculated in Cartesian coordinates by most quantum chemical programs, for example we use the Gamess-UK.[128] However, at any point they are transformed to their internal version by $\mathbf{B}^+\,\mathbf{g}_C = \mathbf{g}_I$ and $\mathbf{B}^+\,\mathbf{H}_C\,(\mathbf{B}^+)^T = \mathbf{H}_I$ (where we ignore the Γ term in (63) for the Hessian, because it is usually small, cf. also the subsection 3.2).

3.1. *Path of steepest descent*

We assume a curve $\mathbf{x}(s) = (x^1(s), \ldots, x^n(s))$ in the configuration space R^n where the tangent vector of $\mathbf{x}(s)$ should point at any value of the parameter s in the direction of the negative gradient vector of E. The curve $\mathbf{x}(s)$ is described by contravariant coordinates, $x^i(s)$, being themselves simple functions of s. Thus, their derivatives to s are dx^i/ds (also contravariant coordinates). It is only permitted to compare objects of the same covariant or contravariant character. We have to use the contravariant form of the gradient vector[6] which includes the metric tensor on the right-hand side

$$\boxed{\frac{dx^i}{ds} = -\sum_{j=1}^{n} \mathrm{g}^{ij} g_j = -g^i\,, \quad i = 1, \ldots, n\,.} \tag{66}$$

Both sides of Eq. (66) are contravariant vectors and change according to the same rule. Mathematically, a solution of Eq. (66) yields a curve invariant from the actual coordinate system. When making the often used (but questionable) ansatz

$$\frac{dx^i}{ds} = -g_i\,, \quad i = 1, \ldots, n \tag{67}$$

one deals with two different kinds of vectors on both sides of the equation. Thus, we could not expect similar behavior on both sides of the equation under coordinate transformations. We can compute a solution curve of Eq. (67) by analytical formulas, if possible, or by numeric methods. However, a computation using this equation would make the curve of Eq. (67) (the SD of E in the actual coordinates) to a non-invariant line; see Refs. 3 and 4 for examples, and Ref. 6 for an extended analysis of the problem. For an example in this review see subsection 3.5. Equation (66) can be utilized to calculate an RP provided the SP of the PES is found beforehand. This "intrinsic" SD path (IRC)[36] leads to the next minimum. However, in general, it does not always follow the "valley floor" line (see Figs. 2 and 4 — compare also Fig. 4 of Ref. 14). The reason is that the valley floor often represents an asymptote of all gradient curves descending to the reaction valley from its side slopes.[27] We need a criterion that allows one to distinguish the floor line points of a valley from all other points. This may be found in the GE equation.[11,12,103,118]

3.2. *Gradient extremal*

For a valley floor, the ansatz of a defining curve results from the idea that the gradient norm of the PES, is proved along an equipotential level, should be

minimal. In Cartesian coordinates, this leads to the equation

$$\mathbf{H}_C \mathbf{g}_C = \lambda \mathbf{g}_C \,, \tag{68}$$

where $\mathbf{H}$ is the Hessian matrix, $\mathbf{g}$ is the gradient vector, and λ is an eigenvalue of the Hessian matrix. The character of the system of Eq. (68) is different from that of the system of Eq. (66). The SD system can be fitted to arbitrary starting points where the gradient is not zero. We can draw a family of gradient lines over a region of the configuration space. In contrast, Eq. (68) has isolated solution curves, in the general case. They do not form a field of neighboring lines.[121] To develop (68) in a coordinate independent shape, we first study the behavior of gradient and Hessian matrix under a transformation. The gradient $\mathbf{g} = (g_1, \ldots, g_n)^T$ is a covariant vector. Correspondingly, we get its contravariant components by

$$g^i = \sum_{j=1}^{n} \mathrm{g}^{ij} g_j \,, \tag{69}$$

(compare Eq. (66)). A derivation of $\mathbf{g}$ to any x^k gives a two-dimensional field of combinations of $\partial^2 E/\partial x^i \partial x^k$ with i and k. It is usually arranged in the so-called Hessian matrix. In general, co- or contravariant characteristics cannot be assigned to the partial derivatives of this matrix under a coordinate transformation, because there are mixed terms resulting from the application of the chain rule[44] (see Eq. (62)). The new terms are connected with the coordinate system and can be compressed in special symbols in the pure internal coordinate space. The matrix

$$(H_{ij}) = \left(\frac{\partial^2 E}{\partial x^i \partial x^j} - \sum_{k}^{n} \frac{\partial E}{\partial x^k} \Gamma_{ij}^k \right) \,, \quad i,j = 1, \ldots, n \tag{70}$$

with

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{l}^{n} \mathrm{g}^{kl} \left(\frac{\partial \mathrm{g}_{jl}}{\partial x^i} + \frac{\partial \mathrm{g}_{il}}{\partial x^j} - \frac{\partial \mathrm{g}_{ij}}{\partial x^k} \right)$$

and

$$\sum_{l}^{n} \mathrm{g}^{kl} \mathrm{g}_{lm} = \delta_m^k \,,$$

shows the character of a two-fold covariant tensor. The functions Γ_{ij}^k are the Christoffel symbols (E. B. Christoffel, 1829–1900)[130] (see textbooks in differential geometry). In the Cartesian coordinates, the metric elements, g_{ij}, are all constants, the Christoffel symbols are zero, and $\mathbf{H}$ reduces to the second-order partial derivatives only. This simple matrix can be the initial matrix for a general definition of the Hessian tensor, in the general case of curvilinear coordinates,[131,132] which we develop with Eq. (70). Note, in the internal molecular coordinates, the Christoffel symbols are connected with the Γ term of Eq. (63) by

$$(\Gamma_{ij}^k) = \mathbf{B}^+ \left(\frac{\partial^2 x^k}{\partial y^i \partial y^j} \right) (\mathbf{B}^+)^T \,, \quad k = 1, \ldots, n \,. \tag{71}$$

The $n+6$ rows of the matrix Eq. (68) become, if we assume that the used coordinates are the incipient ones,

$$\sum_j (H_{ij} - \lambda \delta_{ij}) g_j = 0 \,, \quad i = 1, \ldots, n+6 \,. \tag{72}$$

If we assume the Hessian to be of two-fold covariant character corresponding to Eq. (70), we can transform Eq. (72) from y^i coordinates onto the new coordinates x_i. We get

$$\begin{aligned} 0 &= \sum_{k,l} \sum_j \left(H'_{kl} \frac{\partial x^k}{\partial y^i} \frac{\partial x^l}{\partial y^j} - \lambda \delta_{ij} \right) \left(\sum_m g'_m \frac{\partial x^m}{\partial y^j} \right) \\ &= \sum_{l,m,k} (H'_{kl}\, \mathrm{g}'^{lm}\, g'_m - \lambda g'_k) \frac{\partial x^k}{\partial y^i} = 0 \,, \\ & \quad i = 1, \ldots, n+6 \,. \end{aligned}$$

One needs $\partial x^k / \partial y^i \neq 0$, thus the transformation should not be singular. We get the general, invariant gradient extremal equation[11]

$$\boxed{\sum_{l,m}^{n} H_{kl}\, \mathrm{g}^{lm}\, g_m = \lambda\, g_k \,, \quad k = 1, \ldots, n \,.} \tag{73}$$

Again, it follows from the tensor character of this equation that any solution curve is invariant under coordinate transformations.

3.3. *Eigenvectors of Hessian in internal coordinates*

Sometimes the Hessian is needed, in equations like (73), in the mixed character $\mathbf{H}_I \mathbf{G}^{-1}$. The Hessian may be symbolically given as $(\mathbf{G}^{-1/2})^T\, \mathbf{H}_I\, \mathbf{G}^{-1/2}$ for the frequency analysis,[9] using the Cholesky decomposition[83] of (g^{ij}) into a product of a lower and

an upper triangular matrix:

$$(\mathrm{g}^{ij}) = (\mathrm{g}^{ij})^{1/2} \cdot ((\mathrm{g}^{ij})^{1/2})^T \,.$$

The eigenvalues of the internal Hessian matrix combined with both factors $(g^{ij})^{1/2}$ correspond to the eigenvalues of the Cartesian Hessian $\mathbf{H}_C$ where, additionally, the so-called zero eigenvalues of translation and rotation emerge. To expand this a little, the ansatz (73) is used for a covariant eigenvector $\mathbf{e}$ of $\mathbf{H}_I\,\mathbf{G}^{-1}$. We may imagine the following operation: split the inverse matrix of the metric, $\mathbf{G}^{-1}$, into one lower and one upper tridiagonal matrix

$$(\mathrm{g}^{lm}) = \mathbf{L}\mathbf{L}^T = \begin{pmatrix} L^{11} & \cdot\cdot & 0 \\ \cdot\cdot & \cdot\cdot & \cdot\cdot \\ L^{n1} & \cdot\cdot & L^{nn} \end{pmatrix} \begin{pmatrix} L^{11} & \cdot\cdot & L^{1n} \\ \cdot\cdot & \cdot\cdot & \cdot\cdot \\ 0 & \cdot\cdot & L^{nn} \end{pmatrix}$$

which means for the elements

$$\mathrm{g}^{lm} = \sum_{k=1}^{n} L^{lk}\; L^{km} = \sum_{k=1}^{n} L^{lk}\; L^{mk} \,. \qquad (74)$$

The calculation of $\mathbf{L}$ starts with $L^{11} = \sqrt{g^{11}}$ and obtains by iteration the further L^{lm}.

The eigenvector equation (73) being now

$$\mathbf{H} \cdot \mathbf{G}^{-1} \cdot \mathbf{e} = (\mathbf{H} \cdot \mathbf{L}) \cdot (\mathbf{L}^T \cdot \mathbf{e}) = \lambda \mathbf{e} \qquad (75)$$

gives, if multiplied with $\mathbf{L}^T$ from the left,

$$\boxed{(\mathbf{L}^T \cdot \mathbf{H} \cdot \mathbf{L}) \cdot (\mathbf{L}^T \cdot \mathbf{e}) = \lambda\, (\mathbf{L}^T \cdot \mathbf{e})\,.} \qquad (76)$$

The matrix $(\mathbf{L}^T \cdot \mathbf{H} \cdot \mathbf{L})$ remains symmetric and has the real eigenvector $(\mathbf{L}^T \cdot \mathbf{e})$ and the real eigenvalue λ. Vice versa, it holds for the eigenvector

$$\mathbf{L} \cdot (\mathbf{L}^T \cdot \mathbf{e}) = (\mathbf{L} \cdot \mathbf{L}^T) \cdot \mathbf{e} = \left(\sum_{m=1}^{n} \mathrm{g}^{lm} e_m \right) = (e^l) \qquad (77)$$

which are the contravariant components of the "eigenvector" $(\mathbf{L}^T \cdot \mathbf{e})$, and which covariant version was the starting point in (73). Using the inverse rule to Eq. (69) one may obtain $e_m = \sum \mathrm{g}_{ml} e^l$.

3.4. *The invariance of RGF and of Branin's equation*

For the definition of RGFs using the projector (7) in Eq. (6), we have to start with a covariant vector $\mathbf{r}$ for the search direction because we will compare it with the covariant gradient. We then have to use for the transpose $\mathbf{r}^T$ the contravariant form, in the projector $\mathbf{P_r} := \mathbf{I} - \mathbf{r}\mathbf{r}^{\mathbf{T}}$. It has to be recalculated at new curve points because the metric (slowly) changes from point to point. The dyadic product makes $\mathbf{P_r}$ is a mixed tensor with one covariant and one contravariant index. $\mathbf{P_r}$ is applicable to the covariant gradient in Eq. (6), as well as to the two-fold covariant Hessian in Eq. (8), but it does not change the character of its argument after application. Then an RGF calculation is invariant under coordinate transformation. For TASC, the search direction $\mathbf{r}$ is replaced by the covariant tangent vector. The predictor-corrector method works with $\mathbf{P_r} = \mathbf{P_t}$ as long as the GE to follow is the GE with the smallest (absolute) eigenvalue of the PES.[21]

RGFs are defined by Eqs. (6) and (8). However, they are also solutions of the differential equation of Branin (14) using the adjoint matrix $\mathbf{A}$. The property holds

$$\mathbf{H}\;\mathbf{A} = \mathrm{Det}(\mathbf{H})\;\mathbf{I}\,. \qquad (78)$$

We have with Eq. (70) that $\mathbf{H}$ should be used as a two-fold covariant tensor. To fulfil Eq. (78) under any coordinate transformation: giving a "constant" matrix on the right-hand side, the matrix $\mathbf{A}$ has to be a two-fold contravariant tensor. Thus, in Eq. (14), the matrix $\mathbf{A}$ shifts the covariance of $\mathbf{g}$, and saves the contravariant character of $\mathbf{x}'$. In contrast to the SD Eq. (66), no additional use of the metric is needed.

3.5. *Coordinate independence of VRI points*

We describe an example of an independent definition of VRI points on a 2D test surface.[45,46] The (x, y) system may be a Cartesian system. The function of the surface is $f(x, y) = -x\,(1 + y^2)$. There is the matrix of second derivatives

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 0 & -2y \\ -2y & -2x \end{pmatrix}, \qquad (79)$$

thus the point (0,0) is a VRI point of the surface.[125] A nonlinear coordinate transformation $x(u, v) = u - v^2/2$, $y(u, v) = v$ giving the transformed function $f(x(u, v), y(u, v)) = F(u, v) = (v^2/2 - u)(1 + v^2)$, and the matrix of second derivatives of this function at $u = v = 0$ can be used. The components of the gradient of $f(x, y)$ are $f_x = -1 - y^2$, $f_y = -2xy$, and the

incipient SD equations may be

$$x' = -f_x\,, \quad y' = -f_y\,. \tag{80}$$

In a local understanding of the coordinate transformation, we use the linearized part of the transformation in the usual B matrix scheme[127]

$$\begin{pmatrix} du \\ dv \end{pmatrix} = \begin{pmatrix} u_x(x,y) & u_y(x,y) \\ v_x(x,y) & v_y(x,y) \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix}. \tag{81}$$

To define the SD in a coordinate invariant calculation, we use the chain rule $f_x = F_u\,u_x + F_v\,v_x$, $f_y = F_u\,u_y + F_v\,v_y$, and we use the transformed components $F_u = -1 - v^2$, $F_v = v(1 - 2u + 2v^2)$. They belong to the covariant gradient vector. We include these terms in Eqs. (80) and we multiply the Eqs. (80) either (i) by u_x and u_y, correspondingly, and add both, or (ii) by v_x and v_y and add again both. There is

$$u' = u_x\,x' + u_y\,y' = -F_u\,\mathrm{g}^{11} - F_v\,\mathrm{g}^{12} \quad \text{in case (i)}$$

$$v' = v_x\ ,x' + v_y\,y' = -F_u\,\mathrm{g}^{21} - F_v\,\mathrm{g}^{22} \quad \text{in case (ii)}$$

with the metric symbols $\mathrm{g}^{11} = u_x\,u_x + u_y\,u_y$, $\mathrm{g}^{12} = \mathrm{g}^{21} = u_x\,v_x + u_y\,v_y$, and $\mathrm{g}^{22} = v_x\,v_x + v_y\,v_y$. We obtain the components of the contravariant gradient vector, which is defined by

$$\begin{aligned} F^u &= \mathrm{g}^{11}F_u + \mathrm{g}^{12}F_v\,, \quad \text{and} \\ F^v &= \mathrm{g}^{21}F_u + \mathrm{g}^{22}F_v\,. \end{aligned} \tag{82}$$

The contravariant inverse metric tensor g^{ij} is

$$(\mathrm{g}^{ij}) = \begin{pmatrix} 1+v^2 & v \\ v & 1 \end{pmatrix}. \tag{83}$$

We find

$$\begin{aligned} F^u &= -1 + v^2(v^2 - 2u - 1)\,, \quad \text{and} \\ F^v &= v(-2u + v^2)\,. \end{aligned} \tag{84}$$

The contravariant gradient vector is the right-hand side of a coordinate invariant steepest descent equation

$$\frac{du(s)}{ds} = -\frac{F^u}{|F^u|}\,, \quad \frac{dv(s)}{ds} = -\frac{F^v}{|F^v|}\,. \tag{85}$$

It has the same solution as in the original SD equation in the Cartesian coordinates (x, y). The reason is that this gradient vector has tensor character.[6,131] The long known conclusion is: *the pathway of SD is invariant from a coordinate system* (if we calculate it in the coordinate invariant form of Eq. (85), but see in contrast Eq. (67) above).

In order to study the Hessian matrix, the next derivation of the gradient to u or v results in a combination of second derivations. The invariance problem is trivial for stationary points because the gradient is zero at those points. Invariance problems arise from a non-vanishing gradient. The additional terms are connected with the new coordinate system and can be compressed in special symbols. The matrix becomes[44] in the case of the (u, v) system

$$\begin{aligned} H_{11} &= \frac{\partial^2 F}{\partial u \partial u} - \frac{\partial F}{\partial u}\Gamma^1_{11} - \frac{\partial F}{\partial v}\Gamma^2_{11}\,, \\ H_{12} = H_{21} &= \frac{\partial^2 F}{\partial u \partial v} - \frac{\partial F}{\partial u}\Gamma^1_{12} - \frac{\partial F}{\partial v}\Gamma^2_{12}\,, \\ H_{22} &= \frac{\partial^2 F}{\partial v \partial v} - \frac{\partial F}{\partial u}\Gamma^1_{22} - \frac{\partial F}{\partial v}\Gamma^2_{22}\,. \end{aligned} \tag{86}$$

This matrix (H_{ij}) shows the character of a two-fold covariant tensor. In order to calculate the last term H_{22}, for example, we have to calculate the so-called Christoffel symbols[44,130,132]

$$\begin{aligned} \Gamma^1_{22} &= \frac{1}{2}\,\mathrm{g}^{11}\left(2\frac{\partial \mathrm{g}_{21}}{\partial v} - \frac{\partial \mathrm{g}_{22}}{\partial u}\right) + \frac{1}{2}\,\mathrm{g}^{12}\,\frac{\partial \mathrm{g}_{22}}{\partial v} = -1\,, \\ \Gamma^2_{22} &= \frac{1}{2}\,\mathrm{g}^{21}\left(2\frac{\partial \mathrm{g}_{21}}{\partial v} - \frac{\partial \mathrm{g}_{22}}{\partial u}\right) + \frac{1}{2}\,\mathrm{g}^{22}\,\frac{\partial \mathrm{g}_{22}}{\partial v} = 0\,. \end{aligned} \tag{87}$$

It is also $\Gamma^1_{11} = \Gamma^2_{11} = \Gamma^1_{12} = \Gamma^2_{12} = 0$. Here, we have to use the covariant metric tensor g_{ij} which is the inverse of the g^{ij}. It is

$$\mathbf{G} = (\mathrm{g}_{ij}) = \begin{pmatrix} 1 & -v \\ -v & 1+v^2 \end{pmatrix}. \tag{88}$$

In non-curvilinear Cartesian coordinates, the metric elements g_{ij} are constant, the Christoffel symbols are zero, and (H_{ij}) reduces to the second partial derivatives of Eq. (87) only. For a general definition of the Hessian tensor in curvilinear coordinates we fully need Eq. (87).[44,130,132] For the example holds $H_{11} = 0$, $H_{12} = H_{21} = -2v$, and $H_{22} = 5v^2 - 2u$,

and it is $H_{22} = 0$ for $u = v = 0$. The valley-ridge-inflection point has not moved in the invariantly defined Hessian. It comes out of the tensor character of this matrix that its zero attribute is invariant under coordinate transformations.

4. Discussion

There are demonstrations of the workability of the new algorithms, RGF, Branin, TASC, for following loose RP definitions, or the streambed GE as exactly as one needs it, to find a corresponding SP of index one. TASC can find an SP lying on the top of the streambed GE. One has to start at a given minimum into the direction of the "smallest" eigenvector, or into an assumed reaction direction. (Alternately, we may also start at an SP and go down.) The tangent search concept (TASC) is a modified RGF method comparable in its effort with the RGF method.[7–9] Only the projector matrix, $\mathbf{P_t}$, has to be recalculated after every predictor step. The additional numeric effort is next to nothing. The TASC procedure is a potent method for studying the streambeds of multidimensional surfaces. It allows one to go really uphill a valley floor, "*like so many salmon swimming upstream*". Its success is based on the numerical tracing of the MEP which we mathematically understand as the valley floor GE. This success results from the self correction of the modified RGF method: the tangents of RGF solutions to different directions are "contractive" in the sense that they are always a "better" guess than a constant RGF direction, to search for the streambed line. The original RGF[7,8] already forms an effective tool to find SPs where the choice of the search direction, thus of the "loose RP" is quite arbitrary. But RGF can diverge more or less from the MEP even if it starts in eigenvector direction. The choice of the actual tangent in TASC now overcomes (and restricts) the arbitrariness of the direction choice used in the RGF method.

The solution of TASC is, after some initial iterations, uniquely defined by the PES, actually a numeric approximation of the valley floor GE. But this is also its "fault": its special mathematical working mechanism restricts TASC to GE pathways where the valley direction represents the smallest (absolute) eigenvalue. The complexity of all possible GEs (compare Figs. 10 and 13) is automatically ruled out by TASC to the valley floor GEs being those GEs which are frequently of first chemical interest. Researchers often wish to find the streambed in the conformational space of dihedral angles which represent the weakest modes in a molecule.[133] For example, in the up-to-date discussed folding-unfolding problem of proteins, it is assumed that mainly the weak dihedral variables are involved.[73,134]

Next to the problem of defining a suitable curve for the reaction path of chemistry is the possibility of RP branching. The corresponding points are the so-called bifurcation or branching points (BP). Bifurcations of the path may be caused by symmetry breaking.[28] Then, two equivalent pathways may lead over equivalent transition structures to two or more equivalent (or chiral) products. We assume here that symmetric RP branching is connected with the emergence of a special class of points of the PES, the symmetric valley-ridge inflection (VRI) points.[30,31] The method of following a reduced gradient as well as the Branin method[8] have succeeded in computing such VRI points.

We find manifolds of VRI points: in higher dimensional PES applications, as is the usual case in theoretical chemistry, a whole manifold of VRI points does exist.[8,9,33] This revises the older view of the problem, which suggested obtaining an isolated, well-defined VRI point. The "manifold character" may be compared with the MEP always to be understood as a curve. The case raises the question of the significance of high-dimensional manifolds of VRI points. Which points on the VRI manifold correspond to the chemical concept of reaction path branching? To answer this question we need a criterion allowing us to decide whether a VRI point is located on an MEP or not. The IRC is not defined locally and so it is unsuitable for such a task. In contrast, if a symmetric VRI point fulfills the conditions of a gradient extremal and the eigenvalues of the Hessian indicate an end of a cirque, this VRI point is located on an MEP and so this VRI point is the branching point of that reaction path. The branches are the RGF curves of the VRI point. Usually, the GE itself does not bifurcate at the crossing of the VRI line,[21] it only changes its characterization. The IRC does not play a role in the branching task: it does not bifurcate, any where.

In 1870, James Clerk Maxwell[135] gave a discussion of the so-called lines of slope on the earth surface:

On Lines of Slope

"Lines drawn so as to be everywhere at right angles to the contour-lines are called lines of slope. At every point of such a line there is an upward and a downward direction. If we follow the upward direction we shall *in general* reach a summit, and if we follow the downward direction we shall *in general* reach a bottom. In particular cases, however, we may reach a pass or a bar."

On Hills and Dales

"Hence each point of the /../ surface has a line of slope, which begins at a certain summit and ends in a certain bottom. Districts whose lines of slope run into the same bottom are called Basins or Dales. Those whose lines of slope come from the same summit may be called, for want of a better name, hills. Hence the whole earth may be naturally divided into Basins or Dales, and also, by an independent division, into hills, each point of the surface belonging to a certain dale and also to a certain hill."

On Watersheds and Watercourses

"Dales are divided from each other by Watersheds, and Hills by Watercourses. To draw these lines, begin at a pass or a bar. Here the ground is level, so that we cannot begin to draw a line of slope; but if we draw a very small closed curve round this point, it will have highest and lowest points, the number of maxima being equal to the number of minima, and each one more than the index number of the pass or bar. From each maximum point draw a line of slope upwards till it reaches a summit. This will be a line of Watershed. From each minimum point draw a line of slope downwards till it reaches a bottom. This will be a line of Watercourse. Lines of Watershed are the only lines of slope which do not reach a bottom, and lines of Watercourse are the only lines of slope which do not reach a summit. All other lines of slope diverge from some summit and converge to some bottom, remaining throughout their course in the district belonging to that summit and that bottom, which is bounded by two watersheds and two watercourses.

In the pure theory of surfaces there is no method of determining a line of watershed or a line of watercourse, except by first finding a pass or a bar and drawing the line of slope from that point. / ... /"

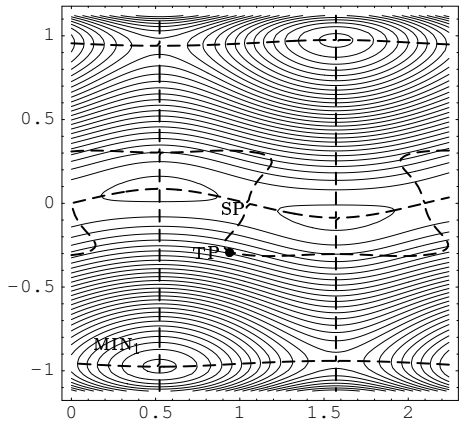

Fig. 13. PES $\{1\}$[51] with GEs and a TP of the valley GE through SP. The TP characterizes the "closure" of the SP col.

The watercourse is, in our frame, the IRC of Fukui. It is a differential equation with the initial value "near" the SP, no local description of this curve is known. The formula for the SD line from SP is still open, for the problem coming from a time two centuries ago! And it could be undetectable, at all. To compare it with RGF: there is the differential equation with the same solution like RGF, the Branin equation. The initial condition for Branin is also the gradient direction of the initial point. The differential equation of Branin is quite analogous to that of the SD. However, in contrast to the SD, there was found the second definition for Branin solutions: the simple algebraic formula (6) which allows a local decision for every point.

There is the second type of curves: the valley-floor-GE. It is a supplement to Maxwell's procedure as long as the valley-floor-GE exists. (Then it can be calculated by TASC.) But also if the valley ends, if it is a "blind" valley, the GE can give us a more accurate means to detect the end. A GE of the PES $\{1\}$ is a suitable example showing how the definition of these curves works (see Fig. 13). The glen of Min_1 is a deep, long, and straight valley. The col of the marked SP opens to this main valley and an SD path goes downhill along the slope perpendicular to the contour lines of the floor (see Fig. 2). At the valley floor, it joins the floor line. However, we cannot decide at which point the lines cross, because this is an asymptotic"junction". The GE curve, on the other hand, shows a totally

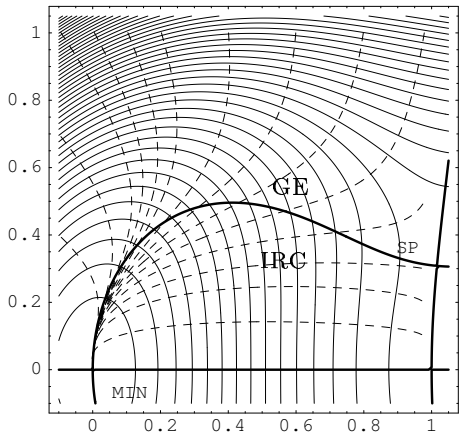

Fig. 14. PES of sample $\{4\}^{96}$ with GE. Steepest descent paths cross the GE, or is confluent to it.

different behavior. It also runs along the col of the SP. But then it turns sideways and goes uphill. The behavior is a consequence of the GE definition to indicate a valley floor line: the col of SP ends at the slope to the broad main valley of Min_1. In other words, the valley is a "blind" one. Thus, the GE has to end, too. The final point is the TP of the GE. The curve continues as a so-called flank line of the potential. One may additionally note that the TP of the GE of Fig. 13 is not the VRI point of the surface (see Fig. 7), although the VRI point (in the 2D case) has to lie on the GE. The gradient is orthogonal to the zero eigenvector. Thus, it is itself an eigenvector in a 2D configuration space. The "closure" of the col valley is not detectable by SD. It is indicated only by the GE.

In Fig. 14 one may finally look at the connection of GE and SD lines: the latter confluent asymptotically into a valley floor, however, they also cross the GE in the curvilinear case. (An instructive example is given in Fig. 9 in Ref. 22.) The SD lines may again deviate from the GE. The crossing point of GE and SD line is the point where the curvature of the SD line is zero.[15,120] The SD from SP is the IRC. GEs and RGF curves (see Fig. 9) form a tool that, in comparison to the IRC treatment only, significantly broadens the possibilities to explore a PES. However, even GEs and RGF cannot answer the ultimate question:

"What is the 'true' MEP?"

— in the general case.

Acknowledgments

The work was made possible through the financial support of the Deutsche Forschungsgemeinschaft. The author warmly thanks M. Hirsch for stimulating discussions, and Prof. D. Heidrich for critically reading the manuscript, and for a long, fruitful cooperation. The author thanks the referees for suggestions and comments.

List of Abbreviations

(Why is "abbreviation" such a long word?)

BP	– bifurcation point (of a curve)
EV	– eigenvector
GE	– gradient extremal
g	– gradient (of the PES)
H	– Hessian (of the PES)
IRC	– intrinsic reaction coordinate
LJ	– (cluster of) Lennard-Jones (particles)
MEP	– minimum energy path
MIN	– Minimum (of PES)
PES	– potential energy hypersurface
RGF	– (method of) reduced gradient following
RP	– reaction path
SD	– (method of) steepest descent
SP	– saddle point (of PES)
TASC	– (method of) tangent search concept
VRI	– valley-ridge-inflection point (of PES)

References

1. K. Laidler, *Theory of Reaction Rates* (McGraw-Hill, New York, 1969).
2. W. Miller, N.C. Handy and J.E. Adams, *J. Chem. Phys.* **72**, 99 (1980).
3. P.G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987).
4. D. Heidrich, W. Kliesch and W. Quapp, *Properties of Chemically Interesting Potential Energy Surfaces* (Springer, Berlin, 1991).
5. D. Heidrich, *The Reaction Path in Chemistry*: *Current Approaches and Perspectives* (Kluwer, Dordrecht, 1995).
6. W. Quapp and D. Heidrich, *Theor. Chim. Acta* **66**, 245 (1984).
7. W. Quapp, M. Hirsch, O. Imig and D. Heidrich, *J. Comput. Chem.* **19**, 1087 (1998).
8. W. Quapp, M. Hirsch and D. Heidrich, *Theor. Chem. Acc.* **100**, 285 (1998).
9. M. Hirsch, W. Quapp and D. Heidrich, *Phys. Chem. Chem. Phys.* **1**, 5291 (1999).
10. I.H. Williams and G.M. Maggiora, *J. Mol. Struct. (Theochem)* **89**, 365 (1982).

11. M.V. Basilevsky and A.G. Shamov, *Chem. Phys.* **60**, 347 (1981).
12. D.K. Hoffman, R.S. Nord and K. Ruedenberg, *Theor. Chim. Acta* **69**, 265 (1986).
13. W. Quapp, *Theor. Chim. Acta* **75**, 447 (1989).
14. W. Quapp, O. Imig and D. Heidrich, *The Reaction Path in Chemistry*: *Current Approaches and Perspectives*, ed. D. Heidrich (Kluwer, Dordrecht, 1995), p. 137.
15. J.-Q. Sun and K. Ruedenberg, *J. Chem. Phys.* **98**, 9707 (1993).
16. P. Chaudhury, S.P. Bhattacharyya and W. Quapp, *Chem. Phys.* **253**, 295 (2000).
17. R. Elber and M. Karplus, *Chem. Phys. Lett.* **139**, 375 (1987); *Chem. Phys. Lett.* **311**, 335 (1999).
18. K. Fukui, *J. Phys. Chem.* **74**, 4161 (1974); K. Fukui, *The World of Quantum Chemistry*, eds. R. Daudel and P. Pullman (Dordrecht, Reidel, 1974) p. 113.
19. B.L. Garret, R. Steckler, D.G. Truhlar, K.K. Baldrige, D. Bartol, M.V. Schmidt and M.S. Gordon, *J. Phys. Chem.* **92**, 1476 (1988).
20. F. Jensen, *J. Chem. Phys.* **102**, 6706 (1995).
21. W. Quapp, M. Hirsch and D. Heidrich, *Theor. Chem. Acc.* **105**, 145 (2000).
22. K. Ruedenberg and J.-Q. Sun, *J. Chem. Phys.* **100**, 5836 (1994).
23. A.H. Zewail, *J. Phys. Chem.* **104**, 5660 (2000).
24. D.G. Truhlar and B.C. Garrett, *Acc. Chem. Res.* **13**, 440 (1980); *Annu. Rev. Phys. Chem.* **35**, 159 (1984); D.G. Truhlar, B.C. Garrett and S.J. Klippenstein, *J. Phys. Chem.* **100**, 12771 (1996).
25. A.D. Isaacson, *The Reaction Path in Chemistry: Current Approaches and Perspectives*, ed. D. Heidrich (Kluwer, Dordrecht 1995), p. 191.
26. A.G. Baboul and H.B. Schlegel, *J. Chem. Phys.* **107**, 9413 (1997).
27. M.V. Basilevsky, *Chem. Phys.* **24**, 81 (1977).
28. K.G. Collard and G.G. Hall, *Int. J. Quantum Chem.* **12**, 623 (1977).
29. T. Taketsugu and T. Hirano, *J. Chem. Phys.* **99**, 9806 (1993); *J. Mol. Struct. (Theochem)* **130**, 169 (1994).
30. H. Metiu, J. Ross, R. Silbey and T.F. George, *J. Chem. Phys.* **61**, 3200 (1974).
31. P. Valtazanos and K. Ruedenberg, *Theor. Chim. Acta* **69**, 281 (1986).
32. F.H. Branin, *IBM J. Res. Develop.* **16**, 504 (1972).
33. W. Quapp and V. Melnikov, *Phys. Chem. Chem. Phys.* **3**, 2735 (2001).
34. H.T. Jongen, P. Jonker and F. Twilt, *Parametric Optimization and Related Topics*, ed. J. Guddat (Akademie-Verl., Berlin, 1987), p. 209; H.T. Jongen, *Computational Solutions of Nonlinear Systems of Equations*, eds. E.L. Allgower and K. Georg (Amer. Math. Soc., Providence, 1990), p. 317; H.T. Jongen, P. Jonker and F. Twilt, *Nonlinear Optimization in Finite Dimensions — Morse Theory, Chebychev Approximation, Transversality, Flows, Parametric Aspects* (Kluwer, Dordrecht, 2000).
35. I. Diener and R. Schaback, *J. Optimiz. Theory Appl.* **67**, 87 (1990); I. Diener, *Globale Aspekte des kontinuierlichen Newton-Verfahrens*, (Habilitation, Göttingen, 1991).
36. K. Fukui, *J. Phys. Chem.* **74**, 4161 (1970).
37. A. Tachibana and K. Fukui, *Theor. Chim. Acta* **49**, 321 (1978).
38. H.B. Schlegel, *J. Chem. Soc., Faraday Trans.* **90**, 1569 (1994).
39. W. Quapp, *J. Chem. Soc., Faraday Trans.* **90**, 1607 (1994).
40. J.Baker and P.M.W. Gill, *J. Comput. Chem.* **9**, 465 (1988).
41. M.V. Basilevsky, *Theor. Chim. Acta* **72**, 63 (1987).
42. A. Tachibana, I. Okazaki, M. Koizumi, K Hore and T. Yamabe, *J. Amer. Chem. Soc.* **107**, 1190 (1985); A. Tachibana, H. Fueno and T. Yamabe, *J. Am. Chem. Soc.* **108**, 4346 (1986).
43. K. Bondensgård and F. Jensen, *J. Chem. Phys.* **104**, 8025 (1996).
44. W. Quapp, *The Reaction Path in Chemistry*: *Current Approaches and Perspectives*, ed. D. Heidrich (Kluwer, Dordrecht 1995), p. 95.
45. D.J. Wales, *J. Chem. Phys.* **113**, 3926 (2000).
46. W. Quapp, *J. Chem. Phys.* **114**, 609 (2001).
47. W. Quapp, *Computers Math. Appl.* **41**, 407 (2001).
48. W. Quapp, submitted, *Optimization.*
49. R.A. Marcus, *J. Chem. Phys.* **45**, 4493 (1966); *J. Chem. Phys.* **49**, 2610, 2617 (1968).
50. D.G. Truhlar and A.J. Kuppermann, *J. Amer. Chem. Soc* **93**, 1840 (1971).
51. Y.G. Smeyers, M. Villa and M.L. Senet, *J. Mol. Spectrosc.* **177**, 66 (1966).
52. K. Ishida, K. Morokuma and A. Komornicki, *J. Chem. Phys.* **66**, 2153 (1977).
53. M.W. Schmidt, M.S. Gordon and M. Dupuis, *J. Amer. Chem. Soc.* **107**, 2585 (1985).
54. K. Müller and L.D. Brown, *Theor. Chim. Acta* **53**, 75 (1979).
55. R.W. Hamming, *Introduction to Applied Numerical Analysis* (McGraw-Hill, New York, 1971).
56. C.W. Gear, *Numerical Initial Value Problems in Ordinary differential Equations* (Prentic-Hall, New York, 1971).
57. V.S. Melissas, D.G. Truhlar and B.L. Garret, *J. Chem. Phys.* **96**, 5758 (1992).
58. J.-Q. Sun and K. Ruedenberg, *J. Chem. Phys.* **99**, 5269; **99**, 5276 (1993); **100**, 6101 (1994).
59. R.D. Amos, A. Bernhardsson, A. Berning, P. Celani, D.L. Cooper, M.J.O. Deegan, A.J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, P.J. Knowles, T. Korona, R. Lindh, A.W. Lloyd, S.J. McNicholas, F.R. Manby, W. Meyer, M.E. Mura, A. Nicklaß, P. Palmieri, R. Pitzer, G. Rauhut, M. Schütz, U. Schumann, H. Stoll,

A. Stone, J.R. Tarroni, T. Thorsteinsson and H.-J. Werner, *MOLPRO quantum chemistry package, v. 2002.1*, http://www.tc.bham.ac.uk/molpro
60. F. Eckert and H.-J. Werner, *Theor. Chem. Acc.* **100**, 21 (1998).
61. W. Quapp, *Chem. Phys. Lett.* **253**, 286 (1996).
62. W. Quapp, *J. Comput. Chem.* **22**, 537 (2001).
63. F. Jensen, *Theor. Chem. Acc.* **99**, 295 (1998).
64. U. Pidun and G. Frenking, *Chem. Eur. J.* **4**, 522 (1998).
65. Y. Kurosaki and T. Takayanagi, *J. Mol. Struct. (Theochem)* **507**, 119 (2000).
66. A. Tachibana and K. Fukui, *Theor. Chim. Acta* **51**, 189 (1979).
67. P. Pechukas, *J. Chem. Phys.* **64**, 1516 (1976).
68. C.W. Bauschlicher, H.F. Schaefer III and C.F. Bender, *J. Amer. Chem. Soc.* **98**, 1653 (1976).
69. M.J. Rothmann and L.L. Lohr Jr., *Chem. Phys. Lett.* **70**, 405 (1980); M.J. Rothmann, L.L. Lohr Jr., C.S. Ewig and J.R. van Wazer, *Potential Energy Surfaces and Dynamics Calulations*, ed. D.G. Truhlar (Plenum Press, New York, 1981), p. 653.
70. P. Scharfenberg, *Chem. Phys. Lett.* **79**, 115 (1981); *J. Comput. Chem.* **3**, 277 (1982).
71. U. Burkert and N.L. Allinger, *J. Comput. Chem.* **3**, 40 (1982).
72. R. Czerminski and R. Elber, *J. Chem. Phys.* **92**, 5580 (1990).
73. M. Černohorský, S. Kettou and J. Koča, *J. Chem. Inf. Comput. Sci.* **39**, 705 (1999).
74. R. Steckler and D.G. Truhlar, *J. Chem. Phys.* **93**, 6570 (1990).
75. K. Müller, *Angew. Chem.* **92**, 1 (1980).
76. J. Cioslowski, A.P. Scott and L. Radon, *Mol. Phys.* **91**, 413 (1997).
77. W. Kliesch, *J. Math. Chem.* **28**, 91 (2000).
78. E.L. Allgower and K. Georg, *Numerical Continuation Methods — An Introduction* (Springer, Berlin, 1990); and in *Handbook of Numerical Analysis V*, eds. P.G. Ciarlet and J.L. Lions (Elsevier, Amsterdam 1997), p. 3.
79. J.M. Anglada, E. Besalu, J.M. Bofill and R. Crehuet, *J. Comput. Chem.* **22**, 387 (2001).
80. J.M. Bofill and J.M. Anglada, *Theor. Chem. Acc.* **105**, 463 (2001).
81. R. Crehuet, J.M. Bofill and J.M. Anglada, *Theor. Chem. Acc.* **107**, 130 (2002).
82. D. Lauvergnat, A. Nauts, Y. Justum and X. Chapuisat, *J. Chem. Phys.* **114**, 6592 (2001).
83. A. Kiełbasiński and H. Schwetlick, *Numerische Lineare Algebra* (Deutscher Verl. Wiss., Berlin, 1988).
84. K. Schiele and R. Hemmecke, *Z. Angew. Math. Mech.* **81**, 291 (2001).
85. M. Dallos, H. Lischka, E. Ventura do Monte, M. Hirsch and W. Quapp, *J. Comput. Chem.* **23**, 576 (2002).
86. W. Quapp and D. Heidrich, *J. Mol. Struct. (Theochem)* **585**, 105 (2002).
87. H. Lischka, R. Shepard, I. Shavitt, R.M. Pitzer, M. Dallos, Th. Müller, P.G. Szalay, F.B. Brown, R. Ahlrichs, H.J. Böhm, A. Chang, D.C. Comeau, R. Gdanitz, H. Dachsel, C. Ehrhardt, M. Ernzerhof, P. Höchtl, S. Irle, G. Kedziora, T. Kovar, V. Parasuk, M.J.M. Pepper, P. Scharf, H. Schiffer, M. Schindler, M. Schüler, M. Seth, E.A. Stahlberg, J.-G. Zhao, S. Yabushita and Z. Zhang, *COLUMBUS ab initio electronic structure program, release 5.8*, http://www.itc.univie.ac.at/∼hans/ Columbus/columbus (2001).
88. R. Ahlrichs, M. Bär, H.-P. Baron, R. Bauernschmitt, S. Böcker, M. Ehrig, K. Eichkorn, S. Elliott, F. Furche, F. Haase, M. Häser, C. Hättig, H. Horn, C. Huber, U. Huniar, M. Kattannek, A. Köhn, C. Kölmel, M. Kollwitz, K. May, C. Ochsenfeld, H. Öhm, A. Schäfer, U. Schneider, O. Treutler, K. Tsereteli, B. Unterreiner, M. v.Arnim, F. Weigend, P. Weis and H. Weiss, *TURBOMOLE Program Package for ab initio Electronic Structure Calculations, Version 5.5*, http:// www.chemie.uni-karlsruhe.de/PC/ TheoChem/ (2002).
89. M. Hirsch and W. Quapp, *J. Comput. Chem.* **23**, 887 (2002).
90. W.A. Kraus and A.E. De Pristo, *Theor. Chim. Acta* **69**, 309 (1986).
91. R.M. Minyaev and M.E. Kletskii, *J. Struct. Chim.* **30**, 202 (1989).
92. B.J. Smith, D.J. Swanton, J.A. Pople, H.F. Schaefer and L. Radom, *J. Chem. Phys.* **92**, 1240 (1990).
93. R.G.A. Bone, T.W. Rowlands, N.C. Handy and A.J. Stone, *Mol. Phys.* **72**, 33 (1991); R.G.A. Bone, *Chem. Phys. Lett.* **193**, 557 (1992); R.G.A. Bone, *Chem. Phys.* **178**, 255 (1993).
94. A. Tachibana, H. Fueno, I. Okazaki and T. Yamabe, *Int. J. Quantum Chem.* **42**, 929 (1992).
95. R.M. Minyaev, *Russ. Chem. Rev.* **63**, 883 (1994); *Zh. Fiz. Khim.* **69**, 408 (1995).
96. J.L. Liao, H.L. Wang and H.W. Xin, *Chin. Sci. Bull.* **40**, 566 (1995).
97. T. Taketsugu and M.S. Gordon, *J. Chem. Phys.* **103**, 10042 (1995); Y. Yanai, T. Taketsugu and T Hirano, *J. Chem. Phys.* **107**, 1137 (1997).
98. G.V. Shustov and A. Rauk, *J. Org. Chem.* **63**, 5413 (1998).
99. S. Sakai, *J. Phys. Chem. A* **104**, 922 (2000).
100. O. Castaño, R. Palmeiro, L.M. Frutos and J.L. Andrés, *J. Comput. Chem.* **23**, 732 (2002).
101. T.H. Choi, S.T. Park and M.S. Kim, *J. Chem. Phys.* **114**, 6051 (2001).
102. D. Heidrich and W. Quapp, *Theor. Chim. Acta* **70**, 89 (1986).
103. M.V. Basilevsky, *Chem. Phys.* **67**, 337 (1982).

104. A.J. Turner, V. Moliner and I.H. Williams, *Phys. Chem. Chem. Phys.* **1**, 1323 (1999).
105. C.J. Cerjan and W.H. Miller, *J. Chem. Phys.* **75**, 2800 (1981); J. Simons, P. Jørgensen, H. Taylor and J. Ozment, *J. Chem. Phys.* **87**, 2745 (1983); D. O'Neal, H. Taylor and J. Simons, *J. Chem. Phys.* **88**, 1510 (1984); A. Banerjee, N. Adams, J. Simons and R. Shepard, *J. Phys. Chem.* **89**, 52 (1985); J. Baker, *J. Comput. Chem.* **7**, 385 (1986).
106. D.J. Wales, *J. Chem. Soc., Faraday Trans.* **89**, 1305 (1993); *J. Chem. Phys.* **101**, 3750 (1994).
107. J.M. Bofill, *J. Comput. Chem.* **15**, 1 (1994).
108. P. Chaudhury and S.P. Bhattacharyya, *Chem. Phys.* **241**, 313 (1999); *Int. J. Quantum Chem.* **76**, 161 (1999).
109. D.A. Liotard and J.P. Penot, *Study of Critical Phenomena*, ed. DellaDora (Springer, Berlin, 1981), p. 213.
110. L.L. Stacho and M.I. Ban, *Theor. Chim. Acta* **83**, 433 (1992).
111. P.Y. Ayala and H.B. Schlegel, *J. Chem. Phys.* **107**, 375 (1997).
112. S. Fischer and M. Karplus, *Chem. Phys. Lett.* **194**, 252 (1992).
113. R. Elber, *Recent Developments in Theoretical Studies of Proteins*, ed. R. Elber (World Scientific, Singapore, 1996), p. 65; R. Czerminski and R. Elber, *Int. J. Quantum Chem., Suppl.* **24**, 167 (1990).
114. L.R. Pratt, *J. Chem. Phys.* **85**, 5045 (1986).
115. S.H. Huo and J.E. Straub, *J. Chem. Phys.* **107**, 5000 (1997).
116. S.S.-L. Chiu, J.W. McDouall and I.H. Hillier, *J. Chem. Soc., Faraday Trans.* **90**, 1575 (1994).
117. G. Henkelman, B.P. Uberuaga and H. Jónsson, *J. Chem. Phys.* **113**, 9901 (2000); G. Henkelman and H. Jónsson, *J. Chem. Phys.* **113**, 9978 (2000); P. Maragakis, St. A. Andreev, Y. Brumer, D.R.Reichman and E. Kaxiras, *J. Chem. Phys.* **117**, 4651 (2002).
118. S. Pancíř, *Coll. Czech. Chem. Commun.* **40**, 1112 (1975).
119. D.J. Rowe and A. Ryman, *J. Math. Phys.* **23**, 732 (1982).
120. H.B. Schlegel, *Theor. Chim. Acta* **83**, 21 (1992).
121. W. Quapp, *Theor. Chim. Acta* **75**, 447 (1989).
122. H. Schwetlick, *Numerische Lösungen nichtlinearer Gleichungen* (Deutscher Verl. Wiss., Berlin, 1979).
123. S.F. Chekmarev, *Chem. Phys. Lett.* **227**, 354 (1994).
124. Special issue of *J. Chem. Soc., Faraday Trans.* **90**, issue 12 (1994).
125. M.-N. Ramquet, G. Dive and D. Dehareng, *J. Chem. Phys.* **112**, 4923 (2000).
126. W. Quapp, H. Dachsel and D. Heidrich, *J. Mol. Struct. (Theochem)* **205**, 245 (1990).
127. E.B. Wilson, J.C. Decius and P.C. Cross, *Molecular Vibrations* (McGraw-Hill Comp., New York, 1955).
128. M.F. Guest, P. Fantucci, R.J. Harrison, J. Kendrick, J.H. van Lenthe, K. Schoeffel and P. Sherwood, *GAMESS-UK Program, Revision C.0* (CFS Ltd Daresbury Lab, 1993).
129. V. Bakken and T. Helgaker, *J. Chem. Phys.* **117**, 9160 (2002).
130. E.B. Christoffel, *J. Reine u. Angew. Math. (Crelle)* **70**, 46 (1869).
131. A. Banerjee and N.P. Adams, *Int. J. Quantum Chem.* **43**, 855 (1992).
132. H. Dachsel and W. Quapp, *J. Math. Chem.* **6**, 77 (1991).
133. L.N. Santana, F.D. Suvire, R.D. Enriz, L.L. Torday and I.G. Csizmadia, *J. Mol. Struct. (Theochem)* **465**, 33 (1999); I. Kolossváry and W.C. Guida, *J. Comput. Chem.* **20**, 1671 (1999); D. ben-Avraham and M.M. Tirion, *Physica A* **249**, 415 (1998); H. Gotō and E. Ōsawa, *J. Mol. Struct. (Theochem)* **285**, 157 (1993).
134. M. Gruebele, *Annu. Rev. Phys. Chem.* **50**, 485 (1999); J. Durup, *J. Mol. Struct. (Theochem)* **424**, 157 (1998); A. Hansen, M.H. Jensen, K. Sneppen and G. Zocchi, *Physica A* **250**, 355 (1998); Y.Z. Chen, V. Mohan and R.H. Griffey, *Phys. Rev. E* **58**, 909 (1998); T. Lazarides and M. Karplus, *Science* **278**, 1928 (1997); P. Argos and R. Abagyan, *Comput. Chem.* **18**, 225 (1994); S. Rackowsky and H.A. Scheraga, *Acc. Chem. Res.* **17**, 209 (1984).
135. J.C. Maxwell, *Philosophical Magazine* **40**, 421 (1870); reprinted in *The Scientific Papers, Vol. II*, (1890), p. 233.