

LOCATING TRANSITION STATES ON POTENTIAL ENERGY SURFACES BY THE GENTLEST ASCENT DYNAMICS

Josep Maria Bofill

*Departament de Química Orgànica, Universitat de Barcelona and
Institut de Química Teòrica i Computacional,
Universitat de Barcelona, (IQTCUB),
Martí i Franquès, 1, 08028 Barcelona, Spain*

Wolfgang Quapp

*Mathematisches Institut, Universität Leipzig,
PF 100920, D-04009 Leipzig, Germany*

Marc Caballero

*Departament de Química Física, Universitat de Barcelona and
Institut de Química Teòrica i Computacional,
Universitat de Barcelona, (IQTCUB),
Martí i Franquès, 1, 08028 Barcelona, Spain*

Abstract

The system of ordinary differential equations for the method of the gentlest ascent dynamics (GAD) has been derived which was previously proposed [W. E and X. Zhou, *Nonlinearity* 24, 1831 (2011)]. For this purpose we use diverse projection operators to a given initial direction. Using simple examples we explain the two possibilities of a GAD curve: it can directly find the transition state by a gentlest ascent, or it can go the roundabout way over a turning point and then find the transition state going downhill along its ridge. An outlook to generalised formulas for higher order saddle-points is added.

Keywords: Potential Energy Surface, Reaction Path, Minimum Energy Path, Transition State Search.

INTRODUCTION

The concepts of the potential energy surface (PES) [1,2] and of the chemical reaction path are the basis for the theories of chemical dynamics. The PES is a continuous function with respect to the coordinates of the nuclei. It is an N -dimensional hypersurface if $N = 3n$ and n is the number of atoms. It must have continuous derivatives up to a certain order.

The PES can be seen as formally divided in catchments associated with local minima [1,3]. The first order saddle points or transition states (TSs) are located at the deepest points of the boundary of the basins. Two neighboring minima of the PES can be connected through a TS via a continuous curve in the N -dimensional coordinate space. The curve characterises a reaction path. One can define many types of curves satisfying the above requirement. The reaction path model widely used is the steepest descent (SD).

There exist a large number of proposed methods that in principle reach a TS when the minimums associated to the reactant and product are known. See reference [4] and references therein. There are also methods that find the TS when only one minimum is known. In this case, the problem is much more cumbersome because the initial data are just the geometry coordinates of the minimum, however, the direction of the search is open. As in the first case many algorithms have been developed for this type of problem [4]. A great number of these algorithms are based in a generalisation of the Levenberg-Marquardt method [5,6,7] that basically consists of a modification of the Hessian matrix to achieve both, first the correct spectra of the desired Hessian at the stationary point, and second to control the length of the displacement during the location process. The first proposed algorithm within this philosophy is due to Scheraga [8] and from than up to now the list is very large [9,10,11,12,13,14,15,16,17,18,19,20]. None of the methods are foolproof, each of them has some problems. Recently E and Zhou [20] have proposed an approach called the "gentlest ascent dynamics" method. This method can be seen as a new reformulation of the method proposed some time ago by Smith [12,13] under the name "image function". The method is based on the generation of an image function that is a function which has its minima at exactly the points where the original PES has its TSs and moreover by an application of a minimum search algorithm to this image function. The converged point

should correspond to a TS in the actual PES. Helgaker [14] modified the algorithm by the trust radius technique. Sun and Ruedenberg [21] analysed the method concluding that image functions do not exist for general PES so that a plain minimum search is inappropriate for them. A nonconservative field gradient of the image function exists. The global structure of the image gradient fields is considerably more complex than that of gradient fields of the original function. However, the image gradient fields appear to have considerably larger catchment basins around TSs. Besalú and Bofill [18] showed that the Smith algorithm is a special type of the Levenberg-Marquardt method.

In this article we show the connection between the Smith method [12,13] and that described by E and Zhou [20] to find TSs, and additionally, the mathematical basis of this algorithm is discussed. Finally, some examples are reported.

BACKGROUND OF THE METHOD

Let us denote by $V(\mathbf{q})$ the PES function and by $\mathbf{q}^T = (q_1, \dots, q_N)$ the coordinates. The dimension of the \mathbf{q} vector is N . The superscript T means transposition. At every interesting point \mathbf{q} the PES function admits a local gradient vector, $\mathbf{g}(\mathbf{q}) = \nabla_{\mathbf{q}} V(\mathbf{q})$, and a Hessian matrix, $\mathbf{H}(\mathbf{q}) = \nabla_{\mathbf{q}} \nabla_{\mathbf{q}}^T V(\mathbf{q})$. The family of image functions of $V(\mathbf{q})$, labelled by $W(\mathbf{q})$, is defined by the differential equation [21]

$$\mathbf{f}(\mathbf{q}) = \mathbf{U}_{\mathbf{v}} \mathbf{g}(\mathbf{q}) = \left[\mathbf{I} - 2 \frac{\mathbf{v}(\mathbf{q}) \mathbf{v}^T(\mathbf{q})}{\mathbf{v}^T(\mathbf{q}) \mathbf{v}(\mathbf{q})} \right] \mathbf{g}(\mathbf{q}) \quad (1)$$

where $\mathbf{f}(\mathbf{q})$ is the image gradient vector, $\mathbf{U}_{\mathbf{v}}$ is the Householder orthogonal matrix constructed by an arbitrary vector $\mathbf{v}(\mathbf{q})$ being in principle a function of \mathbf{q} , and \mathbf{I} is the unit matrix. The Householder orthogonal matrix is a reflection at $\mathbf{v}(\mathbf{q})$. It has the property that $\mathbf{U}_{\mathbf{v}} = \mathbf{U}_{\mathbf{v}}^T$, and it is the result of the difference between the projectors $(\mathbf{I} - \mathbf{P}_{\mathbf{v}})$ and $\mathbf{P}_{\mathbf{v}}$, because it holds trivially $\mathbf{U}_{\mathbf{v}} = \mathbf{I} - 2 \mathbf{P}_{\mathbf{v}} = (\mathbf{I} - \mathbf{P}_{\mathbf{v}}) - \mathbf{P}_{\mathbf{v}}$, being $\mathbf{P}_{\mathbf{v}}$ the projector that projects into the subspace spanned by the \mathbf{v} -vector [22]. If the derivatives of $\mathbf{v}(\mathbf{q})$ with respect to \mathbf{q} are nonvanishing, the image Hessian matrix, $\mathbf{F}(\mathbf{q}) = \mathbf{U}_{\mathbf{v}} \mathbf{H}(\mathbf{q})$, is not obtained by the differentiation of $\mathbf{f}(\mathbf{q})$. Taking into account Eq. (1), this differentiation results in

$$\nabla_{\mathbf{q}} \mathbf{f}^T(\mathbf{q}) = \nabla_{\mathbf{q}} (\mathbf{g}^T(\mathbf{q}) \mathbf{U}_{\mathbf{v}}) = \mathbf{F}^T(\mathbf{q}) + \mathbf{g}^T(\mathbf{q}) \nabla_{\mathbf{q}} \mathbf{U}_{\mathbf{v}} = \mathbf{F}^T(\mathbf{q}) - 2\mathbf{g}^T(\mathbf{q}) \nabla_{\mathbf{q}} \left[\frac{\mathbf{v}(\mathbf{q}) \mathbf{v}^T(\mathbf{q})}{\mathbf{v}^T(\mathbf{q}) \mathbf{v}(\mathbf{q})} \right] \quad (2)$$

The term in brackets is usually not zero and not symmetric, and this non-symmetry is due to the effect of the differentiation on the $\mathbf{P}_{\mathbf{v}}$ projector. In other words,

$$\left(\nabla_{\mathbf{q}} \mathbf{f}^T(\mathbf{q}) \right)_{ij} - \left(\nabla_{\mathbf{q}} \mathbf{f}^T(\mathbf{q}) \right)_{ji} = \left(\nabla_{\mathbf{q}} \mathbf{f}^T(\mathbf{q}) \right)_{ij} - \left(\nabla_{\mathbf{q}} \mathbf{f}^T(\mathbf{q}) \right)_{ij}^T \neq 0 \quad i \neq j \quad (3)$$

The inequality of Eq. (3) implies that the image gradient field defined by Eq. (1) is not integrable to an "image PES" $W(\mathbf{q})$. More explicitly,

$$W(\mathbf{q}_1) - W(\mathbf{q}_0) \neq \int_{t_0}^{t_1} \mathbf{f}^T(\mathbf{q}) (d\mathbf{q}/dt) dt \quad (4)$$

where $d\mathbf{q}/dt$ is the tangent of an arbitrary curve joining the points $\mathbf{q}_0 = \mathbf{q}(t_0)$ and $\mathbf{q}_1 = \mathbf{q}(t_1)$. Due to Eq. (3), this gradient field vector should be considered as a nonconservative force field. From this fact it follows an image of the PES function does, in general, not exist [21,23].

From Eq. (3) it is easy to see that at the stationary points, where $\mathbf{g}(\mathbf{q}) = \mathbf{0}$, the inequality is transformed to an equality if the \mathbf{v} -vector is an eigenvector of the Hessian matrix. Note that if $\{h, \mathbf{v}/(\mathbf{v}^T \mathbf{v})^{1/2}\}$ is an eigenpair of the $\mathbf{H}(\mathbf{q})$ matrix, then $\mathbf{F}(\mathbf{q}) = \mathbf{U}_{\mathbf{v}} \mathbf{H}(\mathbf{q}) = (\mathbf{I} - 2\mathbf{v}\mathbf{v}^T / (\mathbf{v}^T \mathbf{v})) \mathbf{H}(\mathbf{q}) = \mathbf{H}(\mathbf{q}) - 2\mathbf{v}\mathbf{v}^T / (\mathbf{v}^T \mathbf{v}) h = \mathbf{H}(\mathbf{q}) - h 2\mathbf{v}\mathbf{v}^T / (\mathbf{v}^T \mathbf{v}) = \mathbf{H}(\mathbf{q}) - \mathbf{H}(\mathbf{q}) 2\mathbf{v}\mathbf{v}^T / (\mathbf{v}^T \mathbf{v}) = \mathbf{H}(\mathbf{q}) (\mathbf{I} - 2\mathbf{v}\mathbf{v}^T / (\mathbf{v}^T \mathbf{v})) = \mathbf{H}(\mathbf{q}) \mathbf{U}_{\mathbf{v}} = \mathbf{F}^T(\mathbf{q})$. As pointed out by Sun and Ruedenberg [21], the image functions do exist until the second order in the vicinity of its stationary points for any PES taking \mathbf{v} as an eigenvector of the Hessian matrix. Due to this fact the SD curves of the quadratic image function are approximations to the gradient image curves of $W(\mathbf{q})$ being the image potential of $V(\mathbf{q})$.

With the previous analysis of the general nonexistence of an image PES, we can take the image gradient field given in Eq. (1) to define the field of SD curves as,

$$\frac{d\mathbf{q}}{dt} = -\mathbf{f}(\mathbf{q}) = -\mathbf{U}_v \mathbf{g}(\mathbf{q}) = -[(\mathbf{I} - \mathbf{P}_v) - \mathbf{P}_v] \mathbf{g}(\mathbf{q}) \quad (5)$$

where t is the parameter that characterises the SD curve, $\mathbf{q}(t)$. If Eq. (5) is multiplied consecutively from the left by the set of $(N-1)$ linear independent orthogonal vectors to the \mathbf{v} -vector, we see that it corresponds to a curve which is energy descending along this set of directions on the actual PES, whereas is ascending on the \mathbf{v} -vector direction. This property makes the set of curves defined in Eq. (5) suitable for the location of a TS from a minimum. This observation is supported by the fact that Eq. (5) can be rewritten as

$$\frac{d\mathbf{q}}{dt} = -[(\mathbf{I} - \mathbf{P}_v) - \mathbf{P}_v] \mathbf{g}(\mathbf{q}) = -(\mathbf{I} - \mathbf{P}_v) \mathbf{g}(\mathbf{q}) + \mathbf{P}_v \mathbf{g}(\mathbf{q}) = -(\mathbf{I} - \mathbf{P}_v) \mathbf{g}(\mathbf{q}) + l \frac{\mathbf{v}}{(\mathbf{v}^T \mathbf{v})^{1/2}} \quad (6)$$

being $l = \mathbf{v}^T \mathbf{g}(\mathbf{q}) / (\mathbf{v}^T \mathbf{v})^{1/2}$, where the definition of \mathbf{P}_v has been used. Eq. (6) is the basic equation of the string method proposed for the location of reaction paths and TSs [24]. The \mathbf{v} -vector in this method is the current tangent of the path. Because we are interested to find TSs from minimums of the PES we can use the nonconservative property of the gradient image field to modify the \mathbf{v} -vector, during the location process. For this purpose, we first consider that at the minimum, as well as at the TS, the last term of the right hand side part of Eq. (2) is equal zero due to $\mathbf{g}(\mathbf{q}) = \mathbf{0}$. Second, at the TS, the Hessian matrix, $\mathbf{H}(\mathbf{q})$, possesses only one eigenpair with negative eigenvalue. The associated Rayleigh-Ritz quotient of this eigenpair with a negative eigenvalue is the lowest that the Hessian matrix can achieve at this point can achieved being equal to the corresponding eigenvalue [25]. The Rayleigh-Ritz quotient for a given vector \mathbf{v} and matrix \mathbf{H} is defined as, $\lambda(\mathbf{v}) = \mathbf{v}^T \mathbf{H} \mathbf{v} / (\mathbf{v}^T \mathbf{v})$. The structure of this eigenvector is unknown. To find the TS one should, however, ensure that during the research process the path walks through the PES (given until second order) such that the character of the surface becomes closer to a first order saddle point. Taking into account these two considerations we transform Eq. (3) imposing that $\mathbf{g}(\mathbf{q}) = \mathbf{0}$ at each point of the search and multiplying the resulting equation from the left by $(\mathbf{I} - \mathbf{P}_v)$ and from the right by \mathbf{P}_v ,

$$\frac{1}{2}(\mathbf{I}-\mathbf{P}_v)\left[\left(\nabla_{\mathbf{q}}\mathbf{f}^T(\mathbf{q})\right)-\left(\nabla_{\mathbf{q}}\mathbf{f}^T(\mathbf{q})\right)^T\right]\mathbf{P}_v\Big|_{\mathbf{g}(\mathbf{q})=0} = \frac{1}{2}(\mathbf{I}-\mathbf{P}_v)\left[\mathbf{H}(\mathbf{q})\mathbf{U}_v-\mathbf{U}_v\mathbf{H}(\mathbf{q})\right]\mathbf{P}_v = -(\mathbf{I}-\mathbf{P}_v)\mathbf{H}(\mathbf{q})\mathbf{P}_v \quad (7)$$

The effect of this multiplication by the projectors, $(\mathbf{I}-\mathbf{P}_v)$ and \mathbf{P}_v , is that the resulting Eq. (7), multiplied from the right by the \mathbf{v} -vector, is the gradient of the Rayleigh-Ritz quotient with respect to this vector. If the \mathbf{v} -vector is an eigenvector of the $\mathbf{H}(\mathbf{q})$ matrix then the right hand side part of Eq. (7) is equal zero because every eigenvector extremises the corresponding Rayleigh-Ritz quotient. We will denote the Rayleigh-Ritz quotient by $\lambda_q(\mathbf{v})$ to indicate its dependence on \mathbf{q} through the Hessian matrix. If a \mathbf{v} -vector makes the gradient of the Rayleigh-Ritz quotient equal zero then this vector is an eigenvector of the $\mathbf{H}(\mathbf{q})$ matrix and the value of the Rayleigh-Ritz quotient coincides with the corresponding eigenvalue of the $\mathbf{H}(\mathbf{q})$ matrix. These properties suggest that a \mathbf{v} -vector can be changed following the SD direction of the Rayleigh-Ritz quotient gradient with respect to \mathbf{v} ,

$$\frac{d\mathbf{v}}{dt} = -\frac{\mathbf{v}^T\mathbf{v}}{2}\nabla_{\mathbf{v}}\lambda_q(\mathbf{v}) = -(\mathbf{I}-\mathbf{P}_v)\mathbf{H}(\mathbf{q})\mathbf{P}_v\mathbf{v} \quad (8)$$

Eq. (8) is also a function of \mathbf{q} through the Hessian matrix. The Rayleigh-Ritz quotient of the new \mathbf{v} -vector obtained from, $\mathbf{v} \rightarrow \mathbf{v} + d\mathbf{v}/dt \Delta t$, will be lower with respect to the previous one and, in addition, Eq. (5) will give us a new energy ascent direction and a set of orthogonal $N-1$ energy descent directions. Eq. (8) searches for either the lowest positive or the single negative Rayleigh-Ritz quotient if it exists, whereas Eq. (5) determines points on the PES along the action of an increase of the energy in the \mathbf{v} -vector, and a decrease along the set of orthogonal directions to this vector. The specific action of Eq. (5) defines the type of points of the PES so that its structure until the second order on \mathbf{q} resembles a first order saddle point. In turn, Eq. (8) finds \mathbf{v} -vectors with lowest or possibly negative values of the Rayleigh-Ritz quotients of the corresponding Hessian matrix. These two coupled actions on the Eq. (5) and Eq. (8) describe a gradually or gentlest form to find the TS located on the boundary of the basin which contains the start minimum. Eqs. (5) and (8) are coupled since both depend on the \mathbf{v} -vector explicitly on \mathbf{q} through the gradient vector and the Hessian matrix

respectively. The integration of Eqs. (5) and (8) implies the solution of the differential equation

$$\begin{pmatrix} \frac{d\mathbf{q}}{dt} \\ \frac{d\mathbf{v}}{dt} \end{pmatrix} = - \begin{pmatrix} \left(\mathbf{I} - 2 \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} \right) \mathbf{g}(\mathbf{q}) \\ \left(\mathbf{I} - \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} \right) \mathbf{H}(\mathbf{q}) \mathbf{v} \end{pmatrix} \quad (9.a)$$

with the initial conditions

$$\begin{pmatrix} \mathbf{q}(t_0) \\ \mathbf{v}(t_0) \end{pmatrix} = \begin{pmatrix} \mathbf{q}_0 \\ \mathbf{g}(\mathbf{q}_0) \end{pmatrix} \quad (9.b)$$

where \mathbf{q}_0 is a slightly distorted point near the minimum. The gradient at point \mathbf{q}_0 is selected automatically as the initial \mathbf{v} -vector. Eqs. (9) are the basic expressions of the algorithm proposed by E and Zhou [20] which are reviewed in this study.

Note that the system (9) is a system of differential equations. Corresponding to the initial conditions (9a) there emerges a full family of solution curves passing every point of a given region. This is in contrast to the long known gradient extremals (GE). They also realize a shallowest ascent idea; however, they are special solution curves of the equation $\mathbf{H} \mathbf{g} = \lambda \mathbf{g}$, where λ is an eigenvalue of \mathbf{H} . Thus the GEs pass only points where the gradient itself is an eigenvector of \mathbf{H} [2,26,27,28,29,30,31]. GEs do not cover a region.

BEHAVIOUR AND ANALYSIS

The set of Eqs. (9) has been integrated using the explicit Runge-Kutta method of order 8(5,3) [32]. We have used two-dimensional PES models to analyse the behaviour of the gentlest ascent path. The first PES model used for this purpose is the Wolfe-Quapp PES [33,34]. The equation of this surface model is

$$V(x,y) = x^4 + y^4 - 2x^2 - 4y^2 + xy + 0.3x + 0.1y \quad (10)$$

We look for the gentlest ascent dynamics if we start from the minimum located at the point (1.124, -1.486) with energy in arbitrary units -6.369. We take the corresponding gradient vector there as a starting point (1.2, -1.5) and as the initial \mathbf{v} -vector, according to Eq. (9.b). The curve line is depicted in Fig. 1. It starts from the minimum. The curve ends at the TS located in (0.093, 0.174) with energy -0.644. It follows the valley joining these two stationary points. Fig. 1 also shows the behaviour of the \mathbf{v} -vector. Initially the \mathbf{v} -vector has the same direction to the tangent of the curve but along the initial sub-arc of the curve the vector points to the direction where the energy increases very fast, and finally the vector corrects its direction towards the direction of the transition vector. The transition vector is the eigenvector of the Hessian matrix at the TS with the negative eigenvalue. At the end point, namely the TS, the \mathbf{v} -vector is the vector which produces the lowest value of the Rayleigh-Ritz quotient of the Hessian matrix. It is the eigenvalue of the transition vector. This behaviour conforms to that one which is explained in the previous section. The curve is determined by the condition that the \mathbf{v} -vector minimises the $\lambda_q(\mathbf{v})$ within the limits imposed by the Mini-Max eigenvalue theorem [25].

A second curve is computed on the same PES to get the same TS revealing new features of this type of GAD curves. In this case the curve starts near the minimum located at the point (-1.174, 1.477), with energy -6.762. The behavior of this curve is drawn in Fig. 2. The path follows the valley in the direction where $\lambda_q(\mathbf{v})$ takes the possibly lowest value. The path arrives a highest energy value at a turning point and starts to descend along a ridge, and it ends again at the TS located at (0.093, 0.174) with energy -0.644. As shown in Fig. 2, on a large sub-arc from the minimum the \mathbf{v} -vector has the same direction as the tangent of the curve. Of course, the large arc with the turning point at its highest energy does not represent the “gentlest ascent”, seen from the TS and the valley nearby. So, the name of the method, GAD, is a suggestive name but it is not realised in every case. However, note that the TS is in a side valley, but the method finds the TS, in general. At the point (1.849, 0.635) the curve achieves the highest energy value and starts its descents. This point is a turning point (TP) of the curve, where the next equation is satisfied

$$\frac{dV}{dt} = \mathbf{g}^T \frac{d\mathbf{q}}{dt} = -\mathbf{g}^T \mathbf{U}_v \mathbf{g} = -\mathbf{g}^T (\mathbf{I} - \mathbf{P}_v) \mathbf{g} + \mathbf{g}^T \mathbf{P}_v \mathbf{g} = 0 \quad (11)$$

where Eq. (5) has been used. In general a TP appears on the gentlest ascent path when the square of the gradient norm in the subspace spanned by the \mathbf{v} -vector is equal to the square of the gradient norm in the complementary subspace spanned by the set of linear independent vectors orthogonal to the \mathbf{v} -vector. This result is identical to the equality $\mathbf{g}^T \mathbf{g} / 2 = \mathbf{g}^T \mathbf{P}_v \mathbf{g}$, and with the definition of \mathbf{P}_v we can rewrite it as, $1/2^{1/2} = \mathbf{g}^T \mathbf{v} / (\mathbf{g}^T \mathbf{g} \mathbf{v}^T \mathbf{v})^{1/2}$, concluding that the GAD curve has a TP at the point where the \mathbf{v} -vector and the gradient of the actual PES form an angle of $\pi/4$ radians. The curve leaves the valley where the starting minimum is located, at the point TP (1.842, 0.768) and it enters to a ridge region. Because this point satisfies the next equations

$$\frac{\mathbf{g}^T \mathbf{A}(\mathbf{q}) \mathbf{g}}{\mathbf{g}^T \mathbf{g}} = 0 \quad (12.a)$$

$$\frac{\mathbf{A}(\mathbf{q}) \mathbf{g}}{(\mathbf{g}^T \mathbf{g})^{1/2}} \neq \mathbf{0} \quad (12.b)$$

where $\mathbf{A}(\mathbf{q})$ is the adjoint matrix of the Hessian $\mathbf{H}(\mathbf{q})$ at the \mathbf{q} point, we conclude that this point is a valley-ridge transition point (VRT) rather than a valley-ridge inflexion point (VRI) [35,36,37]. In a VRI point both equations are equal zero. From the minimum to the VRT point the curve shows positive values of the Rayleigh quotient, $\mathbf{g}^T \mathbf{A} \mathbf{g} / (\mathbf{g}^T \mathbf{g})$, whereas they are negative from the VRT point to the TS. This quotient is taken as a convexity criterion of a PES region where a point is located. Positivity of the quotient means that the point is in a valley, otherwise it is on a ridge [38]. The result supports the above comment, namely, that one sub-arc of the curve is located in the deep valley from the minimum to the VRT point. At this point the curve turns into the direction of the ridge following the lowest value of $\lambda_q(\mathbf{v})$. But this corresponds to the \mathbf{v} -vector orthogonal to both, the ascent of the valley and the descending ridge directions. On the sub-arc located on the ridge at each point the gradient vector increases its orthogonality with respect to the current \mathbf{v} -vector. The curve evolves along a direction of decreasing energy. We can write $dV/dt = \mathbf{g}^T d\mathbf{q}/dt = -\mathbf{g}^T \mathbf{U}_v \mathbf{g} = 2 \mathbf{g}^T \mathbf{g} ((\mathbf{g}^T \mathbf{v})^2 / (\mathbf{g}^T \mathbf{g} \mathbf{v}^T \mathbf{v}))$

$-1/2) \approx -\mathbf{g}^T \mathbf{g} < 0$ since $\mathbf{v}^T \mathbf{g} \approx 0$. In general the curve evolves so that the energy decreases at each step if the gradient vector \mathbf{g} and the \mathbf{v} -vector form an angle within the open domains $(\pi/4, 3\pi/4)$ or $(5\pi/4, 7\pi/4)$ radians.

In general we can mark the following features of a GAD curve. When along the sub-arc of the gentlest ascent curve its gradient vectors define an angle with the corresponding \mathbf{v} -vectors within the open domains $(\pi/4, 3\pi/4)$ or $(5\pi/4, 7\pi/4)$ radians, then the sub-arc shows an energy decrease. If the angle is within the open domains $(3\pi/4, 5\pi/4)$ or $(7\pi/4, \pi/4)$ radians it shows an energy growth. When the angle is equal to $\pi/4$, $3\pi/4$, $5\pi/4$, or $7\pi/4$ radians the curve is stationary with respect to the variation of the energy, the TP situation.

The same type of calculations with the PES proposed by Neria et al. [39] and modified by Hirsch and Quapp [38] confirms the same facts. The equation of this surface model is

$$V(x, y) = 0.06(x^2 + y^2)^2 + xy - 9 \exp\left(-(x-3)^2 - y^2\right) - 9 \exp\left(-(x+3)^2 - y^2\right). \quad (13)$$

The central point is the TS located at $(0, 0)$ with energy -0.002 . The minima are located at $(2.71, -0.15)$ and $(-2.71, 0.15)$ with energy -5.24 . The surface has gorges around the points $(2, -2)$ and $(-2, 2)$ [38]. We start the gentlest ascent path at the point $(2.6, -0.2)$ near to a minimum. Its evolution goes over the gorge, it leaves the valley where the minimum is located and it enters the ridge, but then it enters again to the valley region of the minimum. The ascending energy behaviour ends at the point $(0.871, -1.428)$ being a TP. On the sub-arc between the points TP and TS a VRT point exists at $(0.002, -1.28)$ corresponding to the descending energy. At this point the GAD curve enters a ridge that touches the TS. Fig. 3 depicts this behaviour, the evolution of the \mathbf{v} -vector is also shown.

On the Müller-Brown PES [40], the GAD method exhausts its possibilities. The global minimum is located at a very deep valley. The coordinates are $(-0.56, 1.44)$. We use two different starting points. The first starting point of the GAD method is at the point $(-0.54, 1.4)$ near to the minimum. The GAD trajectory leads to the next TS, see Fig. 4(a) where it is the dotted line. It goes over a TP but that is normal. In Fig. 4(a) we show the behavior of the gradient extremal curve (GE) [2,26,27,28,29,30,31] for comparison. The GEs are the bold dashed lines. A part of the GAD on the ridge

coincides with the GE. However, the coincidence is for different reasons. For the GAD case, in the ridge region the \mathbf{v} -vector is orthogonal to the gradient and due to this fact the trajectory shows a decrease in energy. In other words, dV/dt is reduced, it is, $dV/dt = -\mathbf{g}^T \mathbf{g} < 0$ since $\mathbf{v}^T \mathbf{g} = 0$. Recall that the \mathbf{v} -vector evolves in this process to reach the eigenvector of the Hessian matrix of lowest eigenvalue. Analyzing the GE in the region of coincidence, the gradient at each point of this curve is by construction an eigenvector of the Hessian corresponding to a positive eigenvalue. The eigenvalue is positive because its eigenvector comes from the evolution of the Hessian at the TS labeled as SP_1 , in direction orthogonal to the col. For the GE the variation energy is given by, $dV/dt = (\mathbf{g}^T \mathbf{g})^{1/2} e_g$, being e_g a component of the GE tangent equation [28,31] and depending of its sign, the curve increases or decreases in energy. One can conclude that in the region of coincidence the \mathbf{v} -vector is orthogonal to the gradient vector, an eigenvector of the Hessian matrix. Thus, the \mathbf{v} -vector is in this region an eigenvector of the Hessian.

For a second trajectory we use the point (-0.58, 1.427) for the start, also near the minimum, however, we get a totally different result, see Fig. 4(b). At the beginning, the GAD trajectory, a continuous line, goes along the deep valley of the minimum, uphill like the GE, which in Fig. 4(b) is again represented by dashed lines, see also Fig. 2 of ref. [28], or Fig. 5 of ref. [41]. The initial evolution of GAD goes in the direction where the gradient and the \mathbf{v} -vector form an angle lower than $\pi/4$ radians. In fact it coincides with the GE, which in turn implies that the \mathbf{v} -vector is parallel to the gradient. In the region where both curves coincide the increasing energy along the GAD curve is $dV/dt = (\mathbf{g}^T \mathbf{v})^2 / \mathbf{v}^T \mathbf{v} > 0$. The behavior ends at a TP which emerges near (-2.8, 0) in a region which is far away from the border line of the searched valley-ridge transition. In Fig. 14 of ref. [38] the borders between valleys and ridges of the Müller-Brown PES are shown. After the TP the trajectory turns into the left large side valley of the Müller-Brown surface. There is no stationary point, so GAD cannot find any one. The region is a kind of a trap, a ‘dead’ valley. The GE passes the region and leaves it to the mountains at the left hand side, but the GAD trajectory goes on and back, from one TP to the next, and again back, and only by an accident, it later leaves the region. However, then it immediately leads to the TS2 of the Müller-Brown surface. It seems to be accidentally which TS is found. After the ‘chaotic’ evolution, the GAD trajectory finally catches a TS.

A DIFFERENT VIEW OF THE THEORY AND GENERALISATION

Mathematically the above method can be reformulated as

$$\left\{ \begin{array}{l} \text{Max}_v \text{Min}_w \left\{ U(v, \mathbf{w}) \mid v = \mathbf{v}^T \mathbf{q}; \mathbf{w} = \mathbf{Z}^T \mathbf{q} \text{ and } \mathbf{Z}^T \mathbf{v} = [\mathbf{z}_1 \mid \dots \mid \mathbf{z}_{N-1}]^T \mathbf{v} = \mathbf{0}_{N-1} \right\} \\ \text{where the } \mathbf{v}\text{-vector satisfies } \text{Min}_v - \frac{\mathbf{v}^T \mathbf{H}(\mathbf{q}) \mathbf{v}}{\mathbf{v}^T \mathbf{v}} \end{array} \right. \quad (14)$$

where $U(v, \mathbf{w})$ is the PES function, $V(\mathbf{q})$, expressed as a function of (v, \mathbf{w}) coordinates being the \mathbf{w} vector of dimension $N-1$. Finally, the \mathbf{Z} matrix is formed by a set of $N-1$ linear independent vectors orthogonal to the \mathbf{v} -vector. A suggested solution of the conditions is formulated in Eq. (9). As noted, the method is appropriate for a first-order saddle point search, and can be easily generalised to the search of higher order saddle-points. The importance of second order saddle-points is known [42,43]. If I is the order of the saddle-point, where $I \leq N$, then the set of Eqs. (9) takes the form

$$\left(\begin{array}{c} \frac{d\mathbf{q}}{dt} \\ \frac{dv_1}{dt} \\ \vdots \\ \frac{dv_I}{dt} \end{array} \right) = - \left(\begin{array}{c} \left(\mathbf{I} - 2 \sum_{i=1}^I \mathbf{P}_{v_i} \right) \mathbf{g}(\mathbf{q}) \\ \left(\mathbf{I} - \sum_{i=1}^I \mathbf{P}_{v_i} \right) \mathbf{H}(\mathbf{q}) \mathbf{v}_1 + \left(\sum_{\substack{i=1 \\ i \neq 1}}^I \mathbf{P}_{v_i} \right) \mathbf{H}(\mathbf{q}) \mathbf{v}_1 \\ \vdots \\ \left(\mathbf{I} - \sum_{i=1}^I \mathbf{P}_{v_i} \right) \mathbf{H}(\mathbf{q}) \mathbf{v}_I + \left(\sum_{\substack{i=1 \\ i \neq I}}^I \mathbf{P}_{v_i} \right) \mathbf{H}(\mathbf{q}) \mathbf{v}_I \end{array} \right) \quad (15.a)$$

with the initial conditions

$$\begin{pmatrix} \mathbf{q}(t_0) \\ \mathbf{v}_1(t_0) \\ \vdots \\ \mathbf{v}_I(t_0) \end{pmatrix} = \begin{pmatrix} \mathbf{q}_0 \\ \mathbf{g}(\mathbf{q}_0) \\ \vdots \\ \mathbf{p}_0^I \end{pmatrix} \quad (15.b)$$

where $\mathbf{P}_{v_i} = \mathbf{v}_i \mathbf{v}_i^T / (\mathbf{v}_i^T \mathbf{v}_i)$ for $i = 1, \dots, I$, and $\{\mathbf{p}_0^i\}_{i=2}^I$ is a set of $I-1$ orthogonal vectors which are orthogonal to the vector $\mathbf{g}(\mathbf{q}_0)$. If $I = N$ then the identity matrix \mathbf{I} in Eq. (15.a) should be substituted by the resolution of identity

$$\mathbf{I} = \sum_{i=1}^N \mathbf{P}_{v_i}. \quad (16)$$

The variation energy through this curve is given by

$$\frac{dV}{dt} = -\mathbf{g}^T \frac{d\mathbf{q}}{dt} = -\mathbf{g}^T \left(\mathbf{I} - 2 \sum_{i=1}^I \mathbf{P}_{v_i} \right) \mathbf{g} = 2 \left(\sum_{i=1}^I \frac{(\mathbf{g}^T \mathbf{v}_i)^2}{\mathbf{g}^T \mathbf{g} \mathbf{v}_i^T \mathbf{v}_i} - \frac{1}{2} \right) \mathbf{g}^T \mathbf{g} \quad (17)$$

where the first N components of Eq. (15.a) have been used. Eq. (17) tells us that when the sum of the square of the components of the projected gradient vector in the subspace spanned by the set of \mathbf{v}_i -vectors is higher than $1/2$ then the curve evolves in the direction of increasing the energy. If it is lower than $1/2$ then evolves in the direction of a decrease of energy. When the sum is equal to $1/2$ the curve is at a TP.

CONCLUSION

In this study is reviewed the GAD method for the location of the saddle points which surround a given minimum. It behaves like a growing string method. The integration of the curve can be done by any method suitable to integrate a system of first order ordinary differential equations. The GAD constructs a curve supported by a set of generated vectors. The numerical examples reported also show that GAD has a large convergence region.

ACKNOWLEDGEMENTS

Financial support has been provided by the Spanish Ministerio de Economía y Competitividad (formaly Ministerio de Ciencia e Innovación) through Grant No. CTQ2011-22505 and by the Generalitat de Catalunya through Grant No. 2009SGR-1472, and XRQTC. M.C. gratefully thanks the Ministerio de Economía y Competitividad for a predoctoral fellowship.

FIGURE CAPTIONS

Figure 1.

Behaviour of the GAD curve on the Wolfe-Quapp PES model [33,34]. The curve starts at point (1.2, -1.5) which is near to the minimum located at the point (1.124, -1.486). The curve evolution ends at the TS located at point (0.093, 0.174). The set of \mathbf{v} -vectors generated during the search is indicated by the set of bold arrows.

Figure 2.

Behaviour of the GAD curve on the Wolfe-Quapp PES model [33,34]. The curve starts at the point near the minimum located at (-1.174, 1.477). The curve evolution ends at the TS located at the point (0.093, 0.174). At the point (1.849, 0.635) the curve achieves the highest energy so this point is the TP. The curve leaves the valley where the starting minimum is located at the point (1.842, 0.768). This point is a VRT point. The set of \mathbf{v} -vectors generated during the search are indicated by the set of bold arrows.

Figure 3.

Behaviour of the GAD curve on the Neria-Fischer-Karplus PES model [39] modified according to that proposed by Hirsch and Quapp [38]. The curve starts at point (2.6, -0.2) near the minimum located at (2.71, -0.15). The curve evolution ends at the TS located at point (0, 0). At point (0.871, -1.428) the curve achieves the highest energy value. The point is the TP. The curve leaves the valley where the starting minimum is

located at the point (0.002, -1.28). This point is a VRT point. The set of \mathbf{v} -vectors generated during the search is indicated by the set of bold arrows.

Figure 4.

(a) Behaviour of the GAD trajectory, dotted line, in the Müller-Brown PES model [40]. The curve starts at the point (-0.54, 1.4) near to the minimum located at (-0.56, 1.44). It leads to the TS labeled as SP_1 . The GE is represented by fat, dashed lines, for comparison. (b) ‘Chaotic’ GAD trajectory, continuous lines, starting at point (-0.58, 1.427). The curve goes through the large side valley at the left hand side by repeated arcs. At last it leaves the valley at an accidental point and finds the TS labeled as SP_2 , see text. The GE is represented by fat, dashed lines.

Figure 1

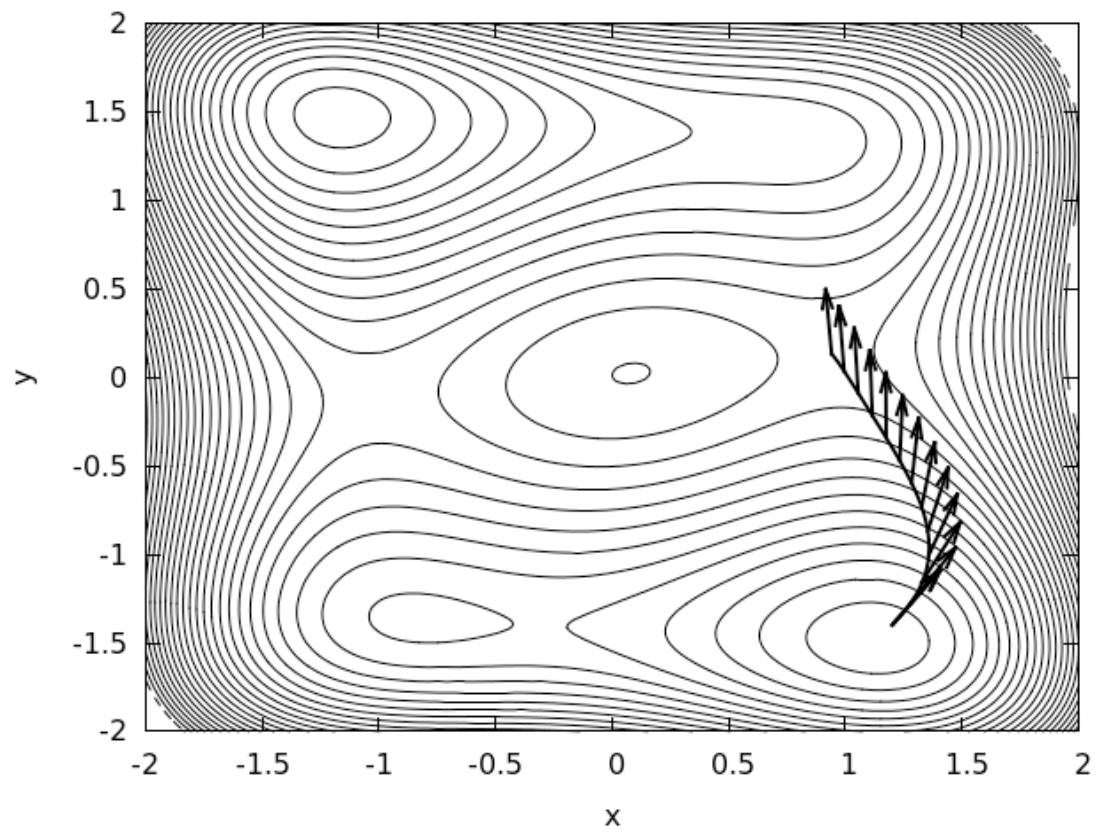


Figure 2.

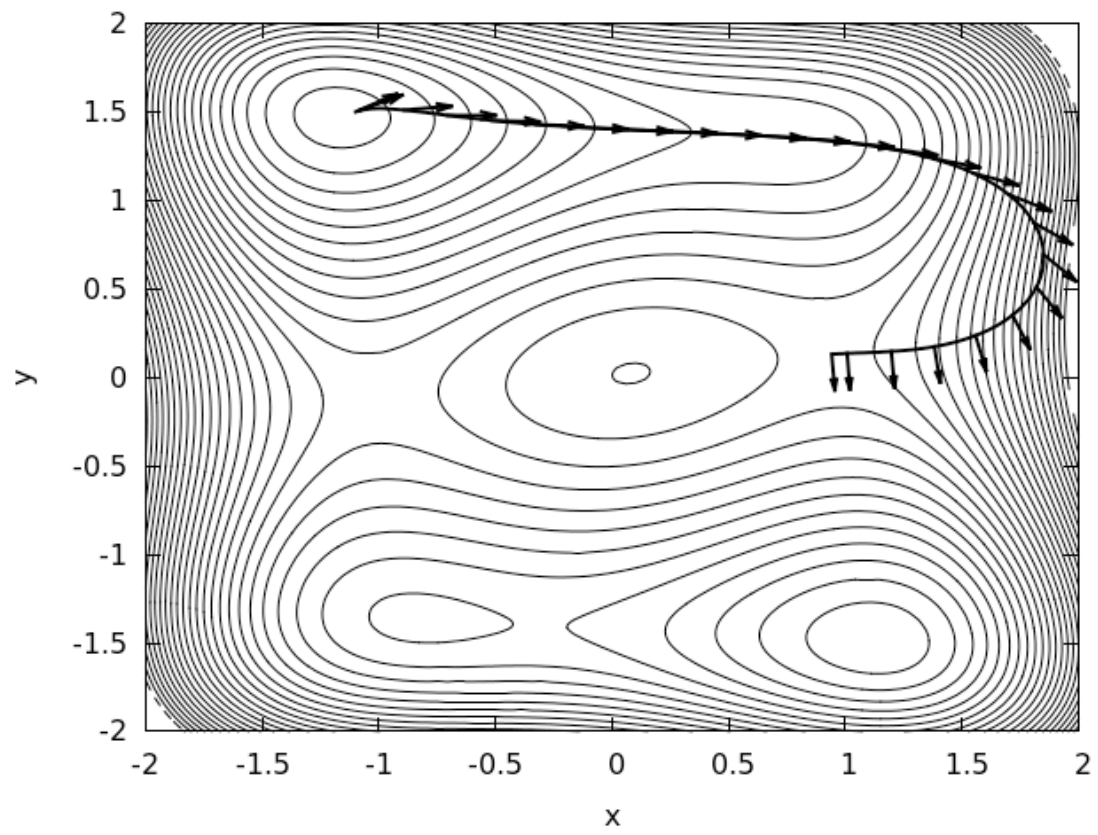


Figure 3.

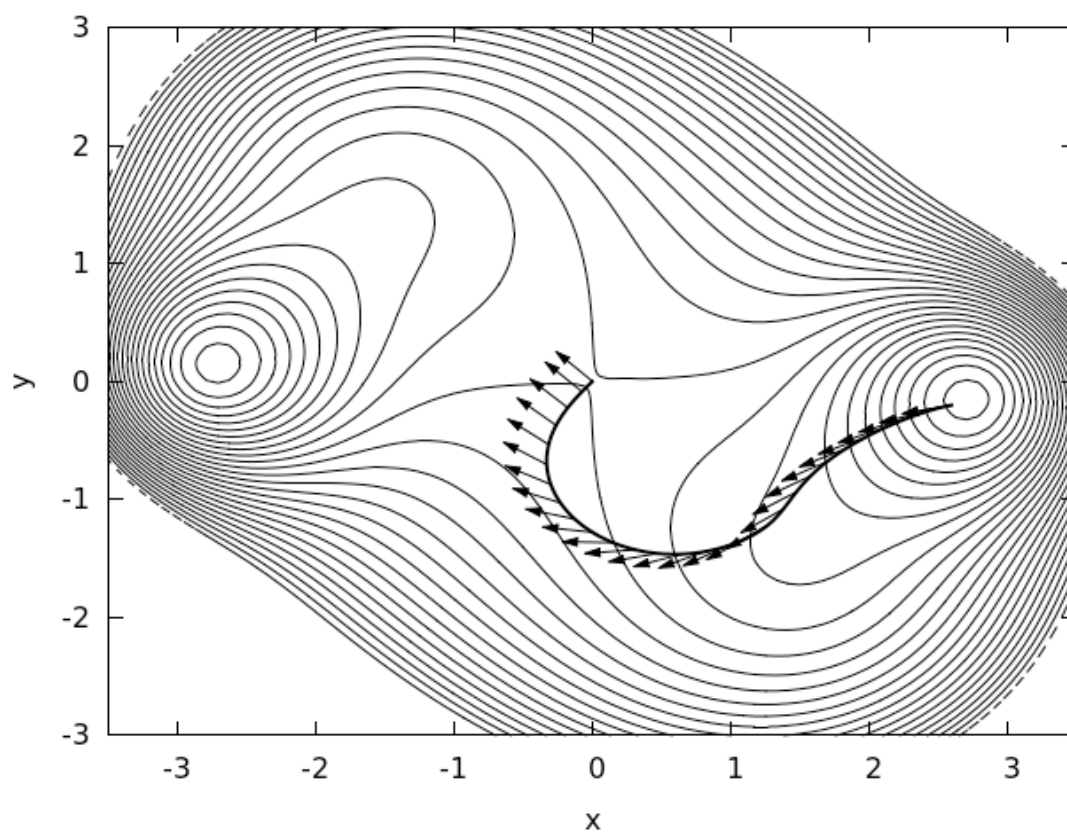


Figure 4a.

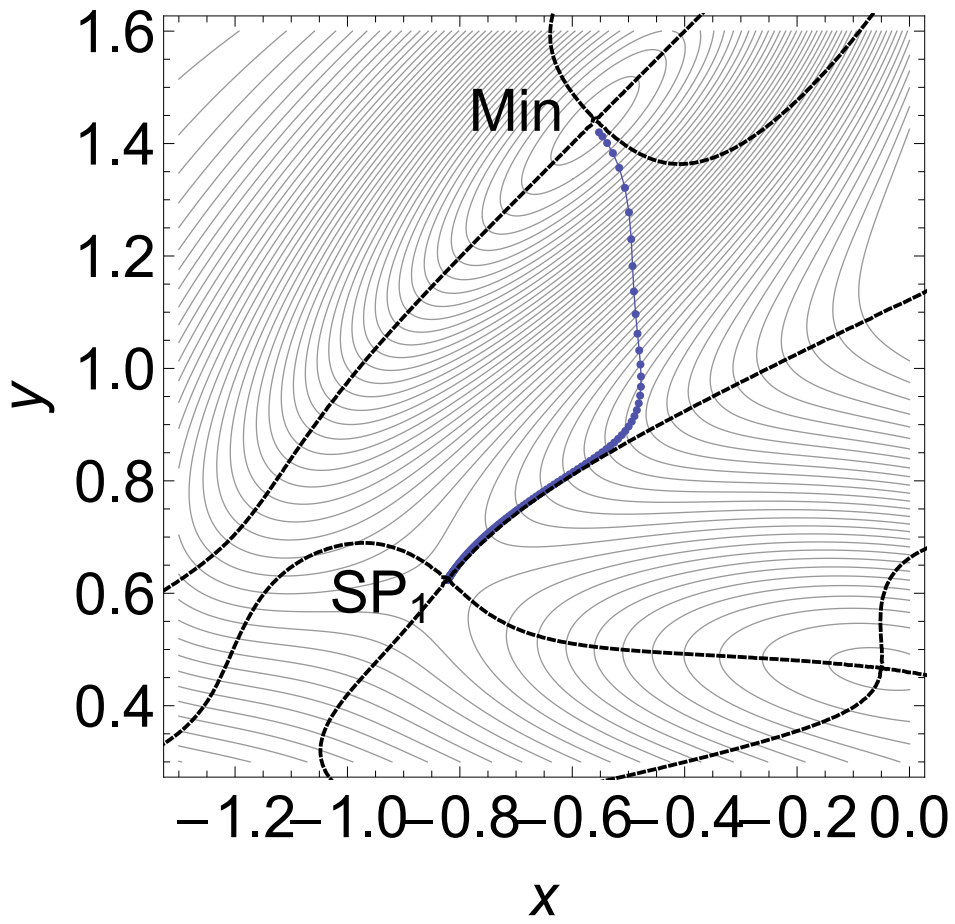
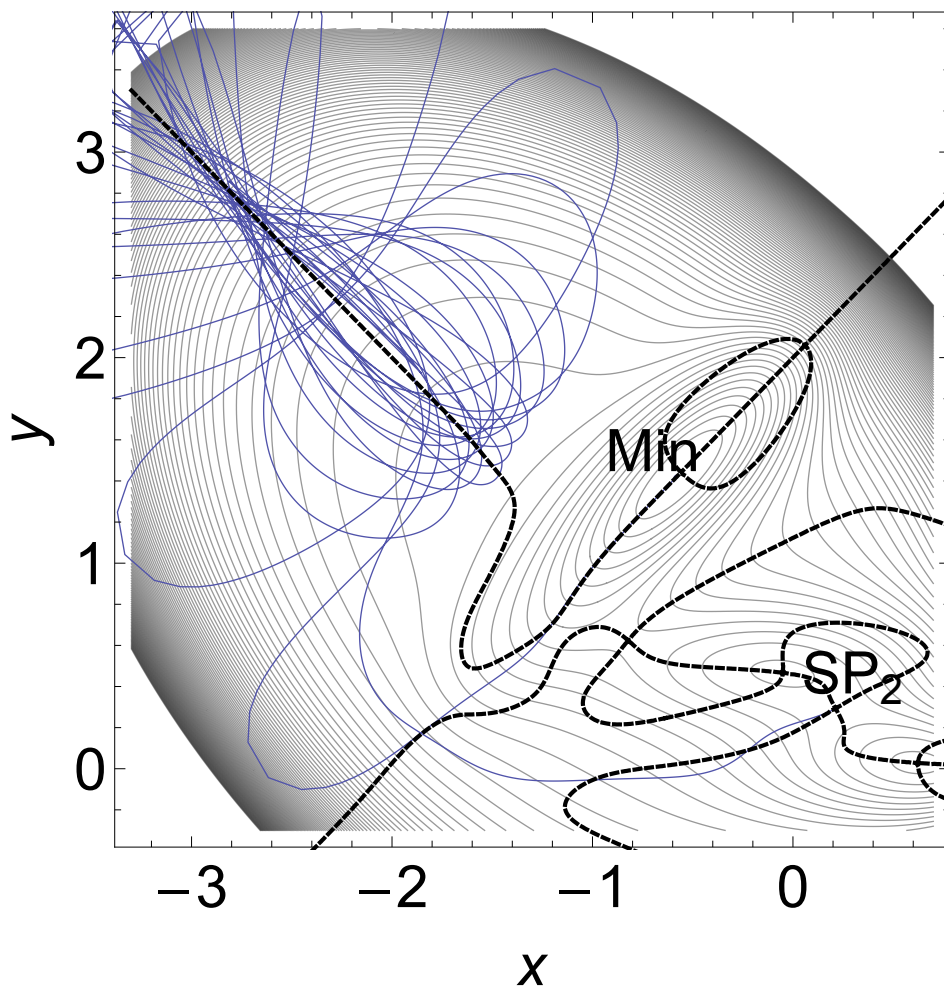


Figure 4b.



REFERENCES

-
- [1] P. G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987).
- [2] D. Heidrich (Ed), *The Reaction Path in Chemistry: Current Approaches and Perspectives* (Kluwer, Dordrecht, 1995).
- [3] P. G. Mezey, *Theor. Chim. Acta* 58, 309 (1981).
- [4] H. B. Schlegel, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 1, 790 (2011).
- [5] K. Levenberg, *Quart. Appl. Math.* 2, 164 (1944).
- [6] D. W. Marquardt, *SIAM J.* 11, 431 (1963).
- [7] R. Fletcher, *Practical Methods of Optimization* (Wiley, New York, 1987).
- [8] G. M. Crippen and H. A. Scheraga, *Arch. Biochem. Biophys.* 144, 462 (1971).
- [9] S. Bell and J. S. Crighton, *J. Chem. Phys.* 80, 2464 (1984).
- [10] A. Banerjee, N. Adams, J. Simons, and R. Shepard, *J. Phys. Chem.* 89, 52 (1985).
- [11] J. Baker, *J. Comput. Chem.* 7, 385 (1986).
- [12] C. M. Smith, *Theor. Chim. Acta* 74, 85 (1988).
- [13] C. M. Smith, *Int. J. Quantum Chem.* 37, 773 (1990).
- [14] T. Helgaker, *Chem. Phys. Lett.* 182, 503 (1991).
- [15] P. Culot, G. Dive, V. H. Nguyen, and J. M. Ghuysen, *Theor. Chim. Acta* 82, 189 (1992).
- [16] J. M. Bofill, *J. Comput. Chem.* 15, 1 (1994).
- [17] P. Y. Ayala and H. B. Schlegel, *J. Chem. Phys.* 107, 375 (1997).
- [18] E. Besalú and J. M. Bofill, *Theor. Chem. Acc.* 100, 265 (1998).
- [19] L. J. Munro and D. J. Wales, *Phys. Rev. B* 59, 3969 (1999).
- [20] W. E and X. Zhou, *Nonlinearity* 24, 1831 (2011).
- [21] J. Q. Sun and K. Ruedenberg, *J. Chem. Phys.* 101, 2157 (1994).
- [22] A. S. Householder, *J. ACM* 5, 339 (1958).
- [23] T. Helgaker, *Lectures in Quantum Chemistry*, in *Lecture Notes Chemistry* 58, edited by B. Roos, (Springer, Berlin, 1992), p. 295.
- [24] W. E, W. Ren, and E. Vanden-Eijden, *J. Chem. Phys.* 126, 164103 (2007).

-
- [25] R. Courant, D. Hilbert, in *Methods of Mathematical Physics, Part 1*, Second English edn., ed. by R. Courant (Wiley, New York, 1953).
- [26] D. K. Hoffman, R. S. Nord, and K. Ruedenberg, *Theor. Chim. Acta* 69, 265 (1986).
- [27] W. Quapp, *Theor. Chim. Acta* 75, 447 (1989).
- [28] H. B. Schlegel, *Theor. Chim. Acta* 83, 15 (1992).
- [29] J.-Q. Sun and K. Ruedenberg, *J. Chem. Phys.* 98, 9707 (1993).
- [30] M. Hirsch and W. Quapp, *Chem. Phys. Lett.* 395, 150 (2004).
- [31] J. M. Bofill, W. Quapp, and M. Caballero, *J. Chem. Theory Comput.* 8, 927 (2012).
- [32] E. Hairer, S.P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I Nonstiff Problems*, Springer Series in Computational Mathematics, Second Edition, (Springer-Verlag, Heidelberg, 1993).
- [33] S. Wolfe, H. B. Schlegel, I. G. Csizmadia, and F. Bernardi, *J. Am. Chem. Soc.* 97, 2020 (1975).
- [34] W. Quapp, *J. Chem. Phys.* 122, 174106 (2005).
- [35] W. Quapp, M. Hirsch, and D. Heidrich, *Theor. Chem. Acc.* 112, 40 (2004).
- [36] W. Quapp, *J. Mol. Struct. (THEOCHEM)* 695–696, 95 (2004).
- [37] J. M. Bofill and W. Quapp, *J. Math. Chem.* 51, 1099 (2013).
- [38] M. Hirsch and W. Quapp, *J. Math. Chem.* 36, 307 (2004).
- [39] E. Neria, S. Fischer, and M. Karplus, *J. Chem. Phys.* 105, 1902 (1996).
- [40] K. Müller and L. D. Brown, *Theor. Chim. Acta* 53, 75 (1979).
- [41] W. Quapp, *J. Chem. Soc. Faraday Trans.* 90, 1607 (1994).
- [42] D. Heidrich and W. Quapp, *Theor. Chim. Acta* 70, 89 (1986).
- [43] R. M. Minyaev, I. V. Getmanskii, and W. Quapp, *Russ. J. Phys. Chem.* 78, 1494 (2004).