

# Locating saddle points of any index on potential energy surfaces by the generalized gentlest ascent dynamics

Wolfgang Quapp · Josep Maria Bofill

Received: 27 February 2014 / Accepted: 22 May 2014 / Published online: 13 June 2014  
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** The system of ordinary differential equations for the method of the gentlest ascent dynamics (GAD) is tested to determine the saddle points of the potential energy surface of some molecules. The method has been proposed earlier [E and Zhou in Nonlinearity 24:1831 (2011)]. We additionally use the metric of curvilinear internal coordinates. By a number of examples, we explain the possibilities of a GAD curve; it can find the transition state of interest by a gentlest ascent, directly or indirectly, or not. A GAD curve can be a model of a reaction path, if it does not contain a turning point for the energy. We further discuss generalized GAD formulas for the search of saddle points of a higher index. We calculate diverse examples.

**Keywords** Potential energy surface · Reaction pathway · Saddle point · Rastrigin and Ackley surfaces · LJ<sub>7</sub> cluster · Radicals H<sub>3</sub>CO · NH<sub>5</sub> · CH<sub>5</sub><sup>−</sup>

## 1 Introduction

The potential energy surface (PES) [1–3] and the chemical reaction path [4] are the basic concepts for the theories of chemical dynamics. The PES is a continuous function with

respect to the coordinates of the nuclei in an  $R^N$ ; thus, it is an  $N$ -dimensional hypersurface. It should have continuous derivatives up to a certain order. If Cartesian coordinates are used, then  $N = 3n$ ; however, we may also use  $N = 3n - 6$  nonredundant, internal coordinates for an  $n$ -atomic molecule. Both cases need some special care.

The PES can be seen as formally divided into catchments associated with local minimums [1]. The first-order saddle points (SPs) or transition states (TSs) are located at the deepest points of the boundary of the basins. TS and minimum correspond to stationary points of the PES. Two minimums of the PES can be connected through a TS via a continuous curve in the  $N$ -dimensional coordinate space, describing the coordinates of the nuclei. The curve characterizes a usual reaction path. However, if one looks for dynamical trajectories, then also SPs of a higher index can be touched. One can define many types of curves satisfying the requirement of a general reaction path. In this paper, we test the 'gentlest ascent' pathways proposed by E and Zhou [5] and already used by Samanta and E [6], by Li et al. [7], and the authors of this paper in a preliminary, theoretical letter [8], and by Zeng et al. [9] in a comparison. The name 'gentlest ascent' was coined in a very early paper [10].

Of course, the location of a TS of index 1 is the second oldest problem for the investigation of a PES, besides the search of minimums. Many methods were developed in the last 80 years. For reviews, see refs. [11, 12]. The GAD method is a further ansatz. In this paper, we test its applicability for molecular PES. A comparison with other search methods for the TS search or for the determination of general reaction pathways is a next task. TSs of higher index emerge often [13–18] on the PES of medium and larger molecules. They may be important for dynamical trajectories, as well as for the comparison with conical

W. Quapp (✉)  
Mathematisches Institut, Universität Leipzig, PF 100920,  
04009 Leipzig, Germany  
e-mail: quapp@uni-leipzig.de

J. M. 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  
e-mail: jmbofill@ub.edu

intersections [19]. To follow, GAD trajectories can be used to determine such TSs.

Let  $V(\mathbf{q})$  be an  $N$ -dimensional surface over  $\mathbb{R}^N$ , let  $\mathbf{grad}(\mathbf{q})$  be its gradient vector,  $\nabla_{\mathbf{q}} V(\mathbf{q})$ , and let  $\mathbf{H}(\mathbf{q})$  be the matrix of its second derivatives, the Hessian  $\nabla_{\mathbf{q}} \mathbf{grad}^T$ . The reason for introduction trajectories by a 'gentlest ascent' dynamics (GAD) is the purposeful search for SPs of index 1, in a first step, or for SPs of a higher, fixed index, for example, for 'summits' of index 2. The index-1-GAD curves are trajectories along the following system of  $2N$  coupled differential equations [5, 8]

$$\dot{q}^i = -grad_i + 2 \frac{\sum_j v^j grad_j}{\sum_k v^k v^k} v^i, \quad i = 1, \dots, N \quad (1)$$

$$\dot{v}^i = - \sum_k H_{jk} v^k + \frac{\sum_k \sum_j v^k H_{kj} v^j}{\sum_k v^k v^k} v^i, \quad i = 1, \dots, N \quad (2)$$

with the initial conditions

$$\mathbf{q}(t_0) = \mathbf{q}_0 \quad (3)$$

and

$$\mathbf{v}(t_0) = \mathbf{grad}(t_0). \quad (4)$$

The meaning of the two first equations is the following: If  $\mathbf{v}$  is an eigenvector of  $\mathbf{H}$ , then the right-hand side of Eq. (2) is zero. Thus, no change to  $\mathbf{v}$  emerges. If additionally the eigenvector is the one belonging to the smallest eigenvalue near the SP and the gradient is parallel to this eigenvector, then the ascent to the SP in Eq. (1) is along this direction. This is the search along the SP-col. And at the SP, the gradient disappears, so the dynamical system converges there. In ref. [5] is shown that the index 1 SPs are stable fixed points of the system (1,2). Further aspects of the action of GAD are explained in SubSect. 6.1. Usually, if one starts near a minimum in the corresponding minimum well, then the dynamical system turns itself the  $\mathbf{v}$  vector to the smallest eigenvector, and we find convergence to a corresponding SP. SPs of index 1 are usually the attractors of the dynamical system (1, 2). But we will also report other cases for the radicals  $\text{CH}_3\text{O}$  and  $\text{CH}_5^-$  in Sect. 6. The use of the guide vector  $\mathbf{v}$  in GAD is a generalization of the method of eigenvector following [20, 21].

Recently, an application of the GAD to Physical Chemistry is demonstrated for the point defect activity on a surface [6, 7]. We will apply GAD to single molecules in an empty space. The description of the PES of an  $n$ -atomic molecule is usually done in two different versions:

1. Cartesian coordinates are used. But here emerge the so-called zero eigenvalues of the Hessian. It means that the movement of the molecule as a whole does not change the force constants of the Hessian. If parts of

the corresponding eigenvectors point into the direction of  $\mathbf{v}$  there, then the system of the GAD differential Eqs. (1, 2) can become stiff. The way out is a projector which is given in Sect. 2.

2. One only uses  $3n - 6$  internal coordinates. They have to be nonredundant; however, they are curvilinear. The inclusion of the curvilinear metric of nonredundant internal coordinates is given in Sect. 3.

Formulas for the application of the GAD to SPs of higher index are given in Sect. 4. For every next index, we need  $N$  additional equations of the GAD system of differential equations, and a further guide vector. In Sect. 5, we report the numerical methods used, and in Sect. 6, we present results of calculations. GAD finds the known test SPs of different index, for analytical surfaces and for different molecules: Lennard-Jones ( $\text{LJ}_7$ ) cluster, and the radicals  $\text{H}_3\text{CO}$ ,  $\text{NH}_5$ , and  $\text{CH}_5^-$ . Thus, we demonstrate that the GAD method works; however, we show also some drawbacks. In a last Section, we discuss the results.

## 2 Projection out of the zero-eigenvectors

Let  $\mathbf{z}_i$ ,  $i = 1, \dots, 6$ , be the six so-called zero-eigenvectors of the  $3n \times 3n$  Hessian matrix. The  $\mathbf{z}_i$  are the fix directions under translations and overall rotations of the molecule. We form the dyadic products

$$\mathbf{p}_i = I - \mathbf{z}_i \mathbf{z}_i^T \quad (5)$$

with the  $3n \times 3n$  unit matrix  $I$ . The six single projectors can form the full projection matrix

$$\mathbf{P} = \prod_{i=1}^6 \mathbf{p}_i. \quad (6)$$

The ansatz  $\mathbf{P}\mathbf{v}$  does not contain any parts of the zero-eigenvectors; it projects out these parts out of vector  $\mathbf{v}$ . Thus, if we use the 'projected' Hessian [23] in Eq. (2)

$$\mathbf{H}^p = \mathbf{P}\mathbf{H}\mathbf{P}, \quad (7)$$

then the GAD system avoids roundabout ways into the overall translation or rotation directions. In the application below in Sect. 6.2 to Lennard-Jones clusters, we simply calculate the six zero-eigenvectors of the Hessian at an initial node, and we use these for the projection. It works, although the rotational eigenvectors depend on the current node.

## 3 Metric

We assume  $N = 3n - 6$  nonredundant internal coordinates for an  $n$ -atomic molecule. Now, the corresponding metric relations are used [11, 24], because normally, the PES is

represented in curvilinear coordinates. The calculations are made with the modified ansatz

$$q^i = -g^{ij} \text{grad}_j + 2 \frac{(v^j \text{grad}_j)}{(v^k g_{kj} v^j)} v^i, \quad i = 1, \dots, N \quad (8)$$

$$v^i = -g^{ij} H_{jk} v^k + \frac{(v^k H_{kj} v^j)}{(v^k g_{kj} v^j)} v^i, \quad i = 1, \dots, N. \quad (9)$$

$(g_{kj})$  is the covariant metric matrix, but  $(g^{ij})$  is the contravariant one. In contrast to Eqs. (1, 2), we use the summation convention: Over a double index, an upper and lower one, we have to sum up. Thus, when an index variable appears twice in a single term, it implies the summation of this term over all the possible values of the index. The summation over a set of indexed terms in a formula is used to achieve a notational brevity. Note that the upper indices are not exponents, but are indices depicting a contravariant vector or tensor. We assume that  $\mathbf{q}$  and  $\mathbf{v}$  are contravariant vectors; however,  $\mathbf{grad}$  is a covariant vector, and  $\mathbf{H}$  is (an approximation) of a twofold covariant tensor (at least near the SP of the PES) [11, 25]. Thus, we can form the scalar products  $v^j \text{grad}_j$  in Eq. (1), as well as  $v^k H_{kj} v^j$  in Eq. (2), but for the norm in the denominator in Eqs. (1, 2), we have to include the metric matrix. For the first summand on the right-hand side in Eq. (8), the gradient component, we also have to 'change' the character to adapt it to the left-hand side vector using the inverse metric matrix. This also happens formally for the vector  $H_{kj} v^j$  in Eq. (2), correspondingly in Eq. (9), which is covariant like the gradient.

Equation (9) contains the (contravariant) eigenvector problem for the covariant tensor  $\mathbf{H}$ : To make the characters equal, we have to treat the problem for the eigenvalue  $\lambda$

$$g^{ik} H_{kj} v^j = \lambda v^i \quad \text{or the inverse } H_{rj} v^j = \lambda g_{ri} v^i. \quad (10)$$

In Eq. (2) follows that for an eigenvector  $\mathbf{v}$ , the steplength becomes zero. Contrary, in (1), if  $\mathbf{v}$  becomes parallel to the contravariant gradient vector  $g^{ij} \text{grad}_j = kv^i$  or the covariant component of  $\mathbf{v}$  becomes parallel:  $\text{grad}_j = kg_{ij} v^i$ , (which is the usual behavior on the col at the SP) then the step becomes  $kv^i$  along the corresponding eigenvector with steplength  $k = (\text{norm of the gradient})$ , and it converges along the valley to the SP. From the point of view of Algebra, Eq. (1) is the 'mirror' transformation of the gradient to a plane which is orthogonal to the  $\mathbf{v}$ -direction.

Note that the 'gentlest ascent' dynamics does not always ascend! This happens, e.g., in Fig. 2 in ref. [8]. There  $\mathbf{v}$  is furthermore the col-vector direction to the smallest eigenvalue, and near the eigenvector, but the dynamical trajectory comes from the ridge near the SP downhill. Then, the Eq. (2) brings again zero steps for the  $\mathbf{v}$ -part of the system, but in Eq. (1), the second summand disappears, and the first summand, the minus gradient, now steps downhill the ridge

to the SP of index 1. Here, the examples of Subject. 6.2 (Figs. 3, 4) are also of this type. They show turning points of the energy profile.

#### 4 Search of SPs of a higher index

SPs of index 2 are interesting for Chemistry [13–18]. The field of SPs of index two is a vibrant and active one. Nowadays, Minyaev et al. [16] found that an SP of index 2 can be part of an IRC downhill from an usual SP of index 1. In the Gamess-US program [21, 22], one can use the commands *nneg* = 2 and *method* = *schlegel* to search for an SP of index 2, for example. It uses the quasi-Newton–Raphson optimizer of H.B. Schlegel [21, 22]. E and Zhou [5] made a propose to search SPs of higher index. We use the SP-of-index-2 ansatz from ref. [8], again modified by the nonlinear metric tensors:

$$q^i = -g^{ij} \text{grad}_j + 2 \frac{(v^j \text{grad}_j)}{(v^k g_{kj} v^j)} v^i + 2 \frac{(w^j \text{grad}_j)}{(w^k g_{kj} w^j)} w^i \quad (11)$$

$$v^i = -g^{ij} H_{jk} v^k + \frac{(v^k H_{kj} v^j)}{(v^k g_{kj} v^j)} v^i \quad (12)$$

$$w^i = -g^{ij} H_{jk} w^k + \frac{(w^k H_{kj} w^j)}{(w^k g_{kj} w^j)} w^i, \quad (13)$$

always for  $i = 1, \dots, N = 3n - 6$ . It is a system of  $3N$  differential equations. The initial values are

$$\mathbf{q}(t_o) = \mathbf{q}_o, \quad \mathbf{v}_o = \mathbf{grad}(t_o), \quad \mathbf{w}_o \perp \mathbf{v}_o, \quad (14)$$

where  $\mathbf{w}_o$  is not unique. The reasoning for the introduction of such a system of equations will come true with its success: The two guide directions  $\mathbf{v}$  and  $\mathbf{w}$  are now used analogously to search a summit with its two negative eigendirections of the  $\mathbf{H}$  matrix. This GAD system searches purposefully an SP of index 2. The idea is from the Lanczos diagonalization procedure which is a well-established diagonalization algorithm, see ref. [26] and references therein with respect to the numerical problems. Note, that sometimes this GAD<sub>2</sub> does not converge to an index-2 SP if the vector  $\mathbf{w}$  turns into direction  $\mathbf{v}$ . The two Eqs. (12) and (13) are uncoupled; sometimes they do not lead to correctly orthogonal vectors  $\mathbf{v}$  and  $\mathbf{w}$ . This is avoidable if one uses a Schmidt re-orthogonalization (SO) of  $\mathbf{w}$  against  $\mathbf{v}$  at every step. Results are discussed in Sect. 6. A more sophisticated ansatz is the following:

We define the projection operator to vector  $\mathbf{v}$  by a dyadic product forming an  $(N \times N)$ -matrix

$$\mathbf{P}_v = \frac{1}{(v^k g_{kr} v^r)} (v^i v^j) \quad \text{with entries for} \quad (15)$$

$$i = 1, \dots, N, \quad j = 1, \dots, N.$$

An analogous definition should hold for  $\mathbf{P}_w$ . Be  $G^{-1}$  the inverse metric matrix ( $g^{ij}$ ),  $G$  the usual metric matrix ( $g_{ij}$ ), and  $I$  be the unit matrix, then we put the eqs. (11), (12), and (13) to

$$\mathbf{q}' = - (G^{-1} - 2\mathbf{P}_v - 2\mathbf{P}_w) \text{grad}(\mathbf{q}) \quad (16)$$

$$\mathbf{v}' = - (G^{-1} - \mathbf{P}_v) \mathbf{H}(\mathbf{q}) \mathbf{v} + SO \quad (17)$$

and

$$\mathbf{w}' = - (G^{-1} - \mathbf{P}_w) (I - \mathbf{P}_v G) \mathbf{H}(\mathbf{q}) (I - \mathbf{P}_v G) \mathbf{w} + SO, \quad (18)$$

where again  $SO$  implies a Schmidt orthogonalization in the metric sense to previous guide vectors. Correspondingly, the initial vectors (14) have to be orthogonal with

$$v^i(t_o) g_{ij}(t_o) w^j(t_o) = 0 \quad (19)$$

and the value for the Rayleigh-Ritz relation for  $\mathbf{v}$

$$\lambda_v(\mathbf{q}) = \frac{(v^k H_{kj} v^j)}{(v^k g_{kj} v^j)} \quad (20)$$

has to be in relation to that of  $\mathbf{w}$  by  $ev_1 \leq \lambda_v(\mathbf{q}) \leq ev_2 \leq \lambda_w(\mathbf{q}) \leq ev_3$  for every current node  $\mathbf{q}$  of the  $GAD_2$  trajectory. In diverse tests, see Sect. 6.1, we sometimes found cases where these relations are not always fulfilled; however, along the pathway, they usually stabilize by themselves and converge with the correct relation of  $\mathbf{v}$  and  $\mathbf{w}$ .

If one searches SPs of a still higher index  $s$ , the system has to be extended correspondingly up to the ( $s + \text{first}$ ) line for the guide vectors  $\mathbf{v}_1 = \mathbf{v}$ ,  $\mathbf{v}_2 = \mathbf{w}$ , ...,  $\mathbf{v}_s$

$$\begin{aligned} \mathbf{v}'_s = - (G^{-1} - \mathbf{P}_{v_s}) \left( I - \sum_{r=1}^{s-1} \mathbf{P}_{v_r} G \right) \\ \mathbf{H}(\mathbf{q}) \left( I - \sum_{r=1}^{s-1} \mathbf{P}_{v_r} G \right) \mathbf{v}_s + SO. \end{aligned} \quad (21)$$

A final remark on the above set of equations to the generalized GAD method: If we are interested in the location of a stationary point of index  $N$  where the number is the maximum of the dimension of the space, then the generalized GAD equations are reduced to the steepest ascent equation. The proof is trivial; we take either Eq. (9) or Eq. (13),

$$\begin{aligned} q^i = -g^{ij} \text{grad}_j + 2 \frac{(v_1^j \text{grad}_j)}{(v_1^k g_{kj} v_1^j)} v_1^i + 2 \frac{(v_2^j \text{grad}_j)}{(v_2^k g_{kj} v_2^j)} v_2^i \\ + \dots + 2 \frac{(v_n^j \text{grad}_j)}{(v_n^k g_{kj} v_n^j)} v_n^i \\ = -(g^{ij} - 2g^{ij}) \text{grad}_j = g^{ij} \text{grad}_j. \end{aligned} \quad (22)$$

In Eq. (22), the resolution of identity has been used. The generalized GAD method to locate a stationary point of

index  $N$  is independent of the set of linear independent  $\mathbf{v}$ -vectors,  $\{\mathbf{v}_i\}_{i=1}^N$ .

We summarize, the generalized GAD is an algorithm to locate a stationary point of any index, say  $s$ , on a PES, with  $1 \leq s \leq N$ . Given a set of  $s$  orthogonal vectors,  $\{\mathbf{v}_i\}_{i=1}^s$ , the generalized GAD curve,  $\mathbf{q}(t) = \mathbf{q}(\mathbf{q}_o, \mathbf{v}_1^o, \dots, \mathbf{v}_s^o, t)$ , is the solution of Eqs. (15–22) with  $\mathbf{q}(t = t_o) = \mathbf{q}_o$ ,  $\mathbf{v}_1^o = \text{grad}(t_o)$  and the  $\{\mathbf{v}_i^o\}_{i=2}^s$  are orthogonal as initial conditions, see Eq. (19), being  $\mathbf{v}_i(t = t_o) = \mathbf{v}_i^o$ . Note that the contravariant version of the gradient vector has to be used, and the orthogonality relations are to be taken with the metric. The generalized GAD curve,  $\mathbf{q}(t) = \mathbf{q}(\mathbf{q}_o, \mathbf{v}_1^o, \dots, \mathbf{v}_s^o, t)$ , is guided by the set of vectors,  $\{\mathbf{v}_i(t)\}_{i=1}^s$ , where,  $\mathbf{v}_i(t) = \mathbf{v}_i(\mathbf{q}_o, \mathbf{v}_1^o, \dots, \mathbf{v}_s^o, t)$ . If the  $GAD_s$  method converges, then it stops at an SP of index  $s$ . If  $s = N$  the algorithm is just the steepest ascent method.

## 5 Numerical methods

The GAD equations are of a form that is amenable to a direct numerical solution: The calculation of the GAD trajectories is done with a high-order Runge-Kutta method like in ref. [8]. The performance of such methods is very good, and they converge as fast as possible to a correct solution. The step length along the pathway is automatically selected. It usually costs some additional effort. However, on the other hand, in some regions, the step length becomes large; and one avoids unnecessary steps. The method of choice here is the explicit Runge-Kutta method of order 8(5, 3) [27], subroutine DOP853. Because this paper is a first test of the GAD method for molecules, we are interested in exact pathways. We use here this very exact procedure. That often causes a high number of steps. The same reason holds for not working with updates of the Hessian. In later, more 'practical' calculations, one should use update steps, as well as a coarser Runge-Kutta method.

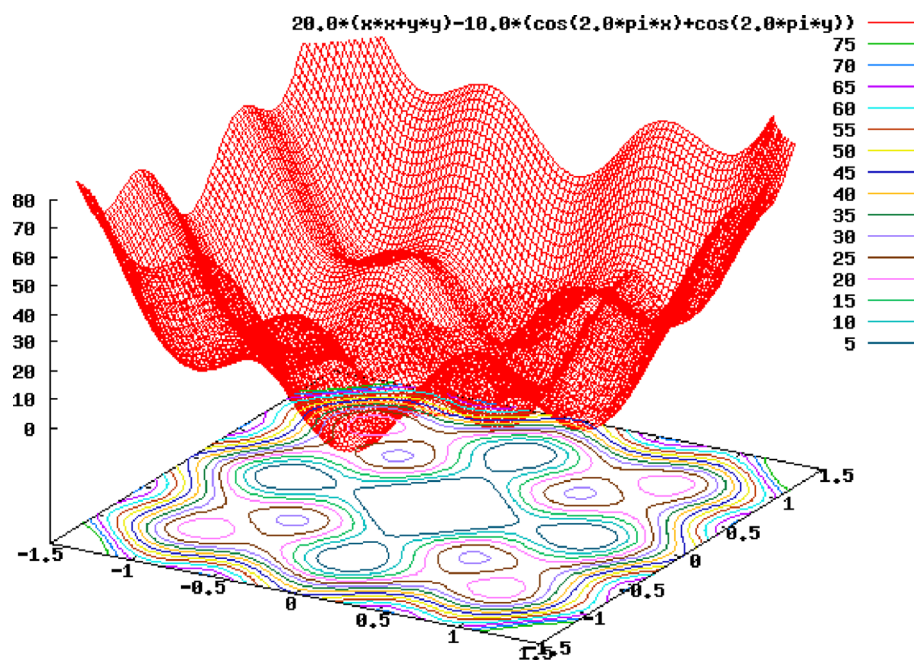
For the PES of interesting molecules, the Gamess-US program is used with functions of the B3YLP 6-311G(d,p) type [21, 22].

## 6 $N$ -dimensional analytical surfaces and chemical systems

### 6.1 $N$ -dimensional surfaces

Before dealing with chemical systems, we test the generalized GAD method as described in Sect. 4 using analytic surfaces as a representation of an  $N$ -dimensional PES. For this aim, the test will be done on the Rastrigin and Ackley functions [28]. Note that for the two examples, the  $\mathbf{G}$  matrix is the unit matrix in the Eqs. (16), (17), and (18).



**Fig. 1** 2-Dimensional Rastrigin surface

### 6.1.1 Rastrigin energy surface

The Rastrigin function is a nonconvex function used as a performance test problem for optimization algorithms. It is an example of a nonlinear function. The two-dimensional case was proposed by Rastrigin [28] and was generalized by Mühlenbein et al. [29]. The search of stationary points of this function is a difficult problem due to the large search space and the large number of stationary points. In Fig. 1, we show the two-dimensional version of the Rastrigin surface. The Rastrigin function is

$$V(\mathbf{q}) = 10N + \mathbf{q}^T \mathbf{q} - 10 \mathbf{1}^T \mathbf{c}(\mathbf{q}) = 10N + \sum_{i=1}^N (q_i^2 - 10 \cos(2\pi q_i)) \quad (23)$$

where  $N$  is the dimension of the system,  $\mathbf{q}^T = (q_1, \dots, q_N)$ ,  $\mathbf{1}^T = (1, \dots, 1_N)$ , and the vectorial function  $\mathbf{c}(\mathbf{q})^T = (\cos(2\pi q_1), \dots, \cos(2\pi q_N))$ . The function is usually used in the hypercubic domain  $q_i \in [-5.0, 5.0]$  for  $i = 1, \dots, N$ . It is easy to calculate gradient vector and Hessian matrix. The Hessian is diagonal at any point  $\mathbf{q}$  of the domain. Each pair of the set of eigenpairs of the Hessian matrix,  $\{h_i, \mathbf{z}_i\}_{i=1}^N$ , has the form

$$h_i = \left( \nabla_{\mathbf{q}} \nabla_{\mathbf{q}}^T V(\mathbf{q}) \right)_{ii} = 2 + 40\pi^2 \cos(2\pi q_i)$$

and  $\mathbf{z}_i^T = (0, \dots, 1_i, \dots, 0)$

for  $i = 1, \dots, N$ . Note that the eigenvectors are the columns of the unit matrix. This result is due to the fact that the Rastrigin function does not have coupling terms between variables, like  $q_i q_j$ .

The global minimum is located in the point  $\mathbf{q}_{gm} = \mathbf{0}$ . In this point,  $V(\mathbf{q}_{gm}) = 0.0$ , and the Hessian matrix is a degenerated matrix of order  $N$ , because  $h_1 = \dots = h_N = 2 + 40\pi^2$ . The properties of the Rastrigin function make it easy to analyze the behavior of the GAD algorithm to locate a stationary point of an index higher than one. If we are interested to locate a stationary point of index, say,  $s$ , where,  $1 \leq s \leq N$ , then the set of the initial guide vectors,  $\{\mathbf{v}_i^0\}_{i=1}^s$ , are selected to be equal to the subset of eigenvectors of the Hessian matrix showing the highest contribution in the absolute value of the initial gradient vector, when the gradient is represented by the set of eigenvectors. In more detail, if at the initial point,  $\mathbf{q}_0$ , the absolute value of the gradient components in the eigenvector space representation satisfies the sequence

$$|\mathbf{z}_1^T \mathbf{grad}(\mathbf{q}_0)| \geq \dots \geq |\mathbf{z}_s^T \mathbf{grad}(\mathbf{q}_0)| \geq \dots \geq |\mathbf{z}_N^T \mathbf{grad}(\mathbf{q}_0)|, \quad (24)$$

then we use the subset of eigenvectors  $\mathbf{z}_1, \dots, \mathbf{z}_s$  for the initial vectors  $\mathbf{v}_1^0, \dots, \mathbf{v}_s^0$ .

With the initial conditions for the set of guide vectors, the GAD equations take their simplest form. The initial guide vectors are eigenvectors of the Hessian, and they are constant in the domain of interest. Consequently, the equality  $d\mathbf{v}_i/dt = \mathbf{0}$  holds for  $i = 1, \dots, s$ , and it is satisfied during the GAD search process. As noted in Sect. 4, the tangent vector,  $d\mathbf{q}/dt$ , is a function of the guide vectors through the Householder transformation, Eqs. (11) or (16), but in the present problem, these guide vectors are  $s$  columns of the unit matrix at every point, and it is

$$\frac{d\mathbf{q}}{dt} = -\left(\mathbf{I} - 2\sum_{i=1}^s \mathbf{v}_i \mathbf{v}_i^T\right) \mathbf{grad}(\mathbf{q}) = \begin{pmatrix} \mathbf{grad}(\mathbf{q})_1 \\ \vdots \\ \mathbf{grad}(\mathbf{q})_s \\ -\mathbf{grad}(\mathbf{q})_{s+1} \\ \vdots \\ -\mathbf{grad}(\mathbf{q})_N \end{pmatrix}. \quad (25)$$

The search direction ascends in the first  $s$  coordinates because the gradient components have the  $+$  sign, and it descends in the remainder  $N - s$  components, where the  $-$  sign multiplies the corresponding components.

Using the numerical Runge-Kutta algorithm [8, 27] described in Sect. 5, we will locate a stationary point of index 5 on the 100-dimensional Rastrigin surface starting from a node near the global minimum,  $\mathbf{q}_{gm}$ . The selected initial point is

$$\begin{aligned} \mathbf{q}_0^T &= (\mathbf{q}_{0-}^T, \mathbf{q}_{0+}^T) = (\mathbf{q}_{0-}^T, \mathbf{0}_+^T) \\ &= (0.001, 0.001, 0.001, 0.001, 0.001, 0, \dots, 0) \end{aligned}$$

being  $\mathbf{0}_+$  a zero vector of dimension 95. The first five components of the gradient vector at the initial node are  $= 0.3968$  in arbitrary units, while the remainder of the components is zero. The eigenvalues of the Hessian matrix there are  $h_- = 396.776$  which is multiply degenerate of order 5 and  $h_+ = 396.784$  which is also degenerate, but of order 95. The eigenvectors of the 5 degenerate eigenvalues correspond to the first five columns of the unit matrix. With the initial conditions, the GAD equations take the form

$$\frac{d\mathbf{q}}{dt} = \begin{pmatrix} \frac{d\mathbf{q}_-}{dt} \\ \frac{d\mathbf{q}_+}{dt} \end{pmatrix} = \begin{pmatrix} \nabla_{\mathbf{q}_-} V(\mathbf{q}) \\ \mathbf{0}_+ \end{pmatrix} \quad (26)$$

$$\frac{d\mathbf{v}_i}{dt} = \mathbf{0} \quad i = 1, \dots, 5. \quad (27)$$

The located point by the Runge-Kutta algorithm is  $q_i = 0.5025$  for  $i = 1, \dots, 5$  where the remainder of the  $\mathbf{q}$  vector elements is zero, as expected. The gradient vector is zero. The eigenvalues of the Hessian at this point are  $h_- = -392.734$  and  $h_+ = 396.784$  which are multiply degenerate in the orders 5 and 95, respectively. The result confirms that the located point is a stationary point of index 5. Note that for this special example, the GAD curve does not have a turning point as defined in ref. [8] because  $dV/dt = (\mathbf{grad}(\mathbf{q}))^T (d\mathbf{q}/dt) = (\nabla_{\mathbf{q}} V(\mathbf{q}))^T (\nabla_{\mathbf{q}_-} V(\mathbf{q})) \neq 0$  between the initial and the converged point, but not including these two points. In this derivation, Eq. (26) has been used. The result indicates that the GAD curve follows the steepest ascent curve and this is due to both the special

structure of the Rastrigin function, Eq. (23), and the selected initial conditions.

### 6.1.2 Ackley surface

The Ackley function was proposed for the same purpose as the Rastrigin function, as a test of new optimizer algorithms [28]. It is defined as

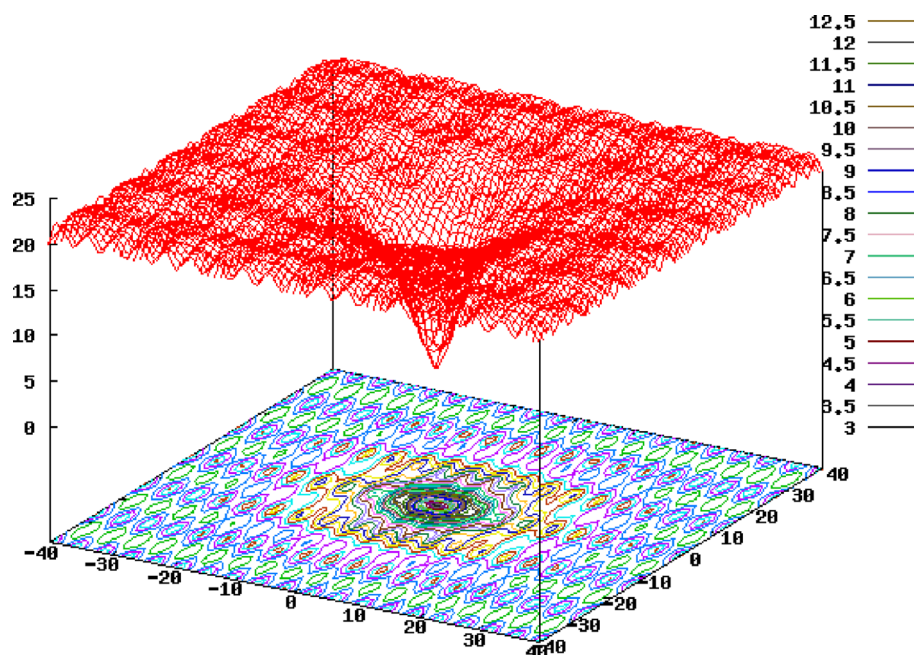
$$V(\mathbf{q}) = -a \exp\left(-b\sqrt{\frac{\mathbf{q}^T \mathbf{q}}{N}}\right) - \exp\left(\frac{\mathbf{1}^T \mathbf{c}(\mathbf{q})}{N}\right) + a + \exp 1 \quad (28)$$

where  $N$  is the dimension of the problem and  $a$ ,  $b$ , and  $c$  are parameters. We use the parameters  $a = 20.0$ ,  $b = 0.2$ , and  $c = 2\pi$ . The meaning of the  $\mathbf{q}$  and  $\mathbf{1}$  vectors is the same as the one given for the Rastrigin function, see Subsection 6.1.1, and it is  $\mathbf{c}(\mathbf{q})^T = (\cos(cq_1), \dots, \cos(cq_N))$ . The function is characterized by a nearly flat outer region, and a large hole at the center. The Ackley function has a useful search region in the hypercubic domain  $q_i \in [-2a, 2a]$  for  $i = 1, \dots, N$ . It is easy to calculate gradient vector and Hessian matrix. Opposite to the Rastrigin surface the Hessian of the Ackley surface has coupling terms between the variables.

The global minimum is located at  $\mathbf{q}_{gm} = \mathbf{0}$ . There holds  $V(\mathbf{q}_{gm}) = 0$ ; however, the gradient vector becomes singular  $\mathbf{grad}^T(\mathbf{q}_{gm}) = (\infty, \dots, \infty)$ , as well as all the elements of the Hessian are indeterminate. This implies that the Ackley surface, Eq. (28), is a continuous function, but not derivable in the point  $\mathbf{q}_{gm} = \mathbf{0}$ . Around the minimum point, the function has the form of a cone with the vertex located at  $\mathbf{0}$ . In Fig. 2, we show the two-dimensional version of the Ackley surface.

Regarding Fig. 2, we may see that the topography of this surface is close to that proposed as the PES model of a protein folding rearrangement, see e.g., [30]. If we take the Ackley surface as a model, the native conformation of the protein is located in the global minimum,  $\mathbf{q}_{gm}$ , and the minimums around the deep valley are related to the conformations associated with the denatured states of the protein. These minimums are connected through a reaction path to the global minimum. They have direct access to the native conformation. The remainder of denatured conformations is related to other minimums. They are called dead end or kinetic traps. To reach the global minimum related to the native conformation, any conformation beginning in the dead-end region of the PES must first pass through the conformations of denatured states that are located around the deep native valley. The features of the PES of a protein folding [30] are well represented by the model of the Ackley surface given in Eq. (28).

**Fig. 2** Two-dimensional Ackley surface



The study of the Ackley surface using the GAD method reveals some limitations in the selection of the initial conditions. Using the same procedure in the selection of the initial conditions as that described in the previous subsection, we attempt to locate a stationary point of index 2 on the four-dimension Ackley surface. We start at the initial node  $q_1^0 = q_2^0 = 0.001, q_3^0 = q_4^0 = 0$ , with a potential energy  $V(\mathbf{q}_0) = 0.0029$  in arbitrary units, which is close to the global minimum. The eigenvalues of the Hessian matrix at this point are  $h_1^0 = 1436.6412, h_2^0 = 1440.8411, h_3^0 = h_4^0 = 1440.8417$  revealing that the eigenpairs 3 and 4 are degenerate. From the Algebra, this implies that the corresponding eigenvectors,  $\mathbf{z}_3^0$  and  $\mathbf{z}_4^0$ , define a subspace of eigenvectors of dimension two. The comparison of projections of the gradient at the initial point is done using inequality (24). We find the space spanned by the eigenvectors of the Hessian matrix where the two highest contributions come from eigenvectors  $\mathbf{z}_1^0$  and  $\mathbf{z}_3^0$ . This version indicates that the initial guess vectors,  $\mathbf{v}_1^0$  and  $\mathbf{v}_2^0$ , are  $\mathbf{z}_1^0$  and  $\mathbf{z}_3^0$ , respectively. However, the degeneracy of  $\mathbf{z}_3^0$  eigenvector precludes to use Eqs. (17) and (18), because they force a partition of the eigenvector space which is not correct.

Note that the generalized GAD equations to find a stationary point of index say,  $s$ , imply a subspace spanned by the selected guide vectors,  $\{\mathbf{v}_i\}_{i=1}^s$ , and the corresponding orthogonal subspace. If at the initial point the selected guide vectors are some eigenvectors of the Hessian matrix, the selection should take into account the possible degeneracy of the eigenpairs. The partition of the guide vectors should be in accord to the degeneracy of the Hessian matrix

at the converged stationary point. If at the converged stationary point a guide vector, which is already an eigenvector of the Hessian matrix is degenerate with another eigenvector not included in the subspace of the guide vectors, then the GAD procedure has been converged to a stationary point of index  $s + 1$ . The reason is that at the initial node, the Rayleigh-Ritz quotient is stationary with respect to the guide vectors and during the GAD search procedure, this stationarity is preserved as much as possible. The present example shows the last situation. The coordinates of the converged stationary point are,  $q_1^0 = q_2^0 = q_3^0 = 0.6635$  and  $q_4^0 = -1.1635$ . The eigenvalues of the Hessian matrix are,  $h_1 = -8.8893, h_2 = h_3 = -2.9014, h_4 = 1.8146$ . One easily concludes that the converged point is a stationary point of index 3 rather than index 2 as it was required in the construction of the GAD search equations.

Finally, we present a case where we locate a stationary point of a desired index starting from a point such that the Hessian matrix is highly degenerate. The surface is the 100-dimension Ackley function. We are now interested in the location of a stationary point of index 5 starting from the node with coordinates  $q_1^0 = 10.0, q_2^0 = \dots = q_{100}^0 = 0$ . At this point,  $V(\mathbf{q}_0) = 3.6254$  in arbitrary units. This point is far from the global minimum. The eigenvalues of the Hessian matrix are  $h_1^0 = -2.1756, h_2^0 = \dots = h_{100}^0 = 1.1059$ , indicating that there is a multiple eigenvalue of order 99. Proceeding as before, the generalized GAD algorithm localized at a stationary point of index 5 with coordinates,  $q_1 = 10.5323, q_2 = 0.5323, q_3 = q_4 = q_5 = -0.4677, q_6 = \dots = q_{100} = -0.0323$ .

The eigenvalues of the Hessian matrix are  $h_1 = -4.1457, h_2 = -0.9047, h_3 = h_4 = h_5 = -0.9031, h_6 = \dots = h_{100} = 0.9263$ , showing that the stationary point indeed has index 5.

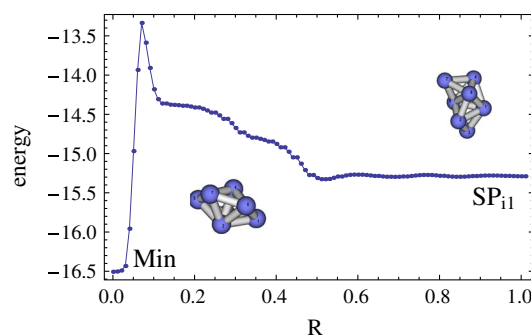
## 6.2 Lennard-Jones cluster

Before coming to molecules, we choose a cluster of  $n = 7$  abstract LJ kernels ( $\sigma = 1, \epsilon = 1$ ) for the GAD experiment. It has an analytical PES. Energy, Gradient, and Hessian can be explicitly calculated, without approximation. Geometries of LJ<sub>7</sub> for the start and for comparisons are taken from the Doye homepage [31], for the formulas see an Appendix. We use 21 Cartesian coordinates in the full  $3n$ -dimensional space. Thus, we have to use the projected Hessian matrix (7) for the 42-dimensional GAD system (2). A test starts near the minimum, MIN<sub>1</sub>, (in the Doye numbering) with small initial deviations for particle 4 ( $\delta x = 0.002, \delta y = -0.001, \delta z = 0.0003$ ). GAD finds the index-1 SP (number 3 in the Doye numbering), see Fig. 3. Up to the cumulative steplength  $R = 1$  one needs 80 steps and obtains an accuracy  $|grad|_{max} = 0.16$  of the largest component of the gradient. To get any accuracy, one has to carry on the calculation: For example, for  $|grad|_{max} = 1.6 \times 10^{-5}$ , GAD needs 517 steps, and additionally 14 steps for step length corrections. Less than the tenth number of the steps leads to the neighborhood of the SP, the remainder is used for the refinement of the SP. So, if one stops GAD 'near' the SP of interest, and if one jumps there to a Newton–Raphson step, one can avoid the long convergence process of the DOP853 routine.

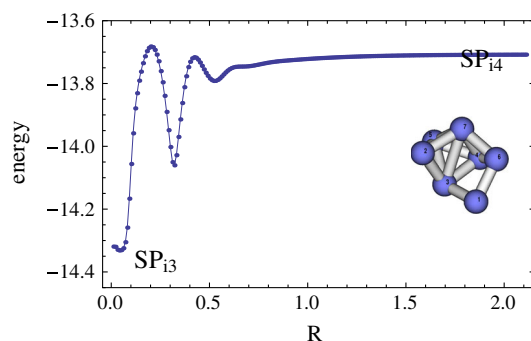
Note that bullets on the profile of Fig. 3 are drawn at fixed  $R$  steps. They do not reflect the steps of the DOP853 routine. Of course, there are also many start values near the minimum which lead to 'dissociative' geometries of the clusters: LJ<sub>6</sub>+LJ<sub>1</sub> or LJ<sub>5</sub>+LJ<sub>2</sub>, for many combinations of the seven particles. Such structures are not really interesting here. We search for an 'isomerization' like in Fig. 3. A 'good' choice of the initial deviation from the minimum is a critical aspect of the GAD method.

A next test for the GAD method is the search of SPs of a higher index, along the Eqs. (11) to (13). The metric here is the trivial unit matrix, and the additional vector  $\mathbf{w}$  is orthogonalized to  $\mathbf{v}$  at every step of the procedure (Or of vectors  $\mathbf{v}_i, 2 \leq i \leq s$ , if  $s$  is the index of the searched SP. We then use  $(s - 1)$  additional equations like Eq. (13)). If we look for an index-2 SP, we need one more vector  $\mathbf{w}$ , and the index-2-GAD system has  $3 \times 3n = 63$  dimensions.

For small LJ clusters, the description of the numbers and characters of diverse SPs are given by Doye and Wales [20]. So, we can stop here our test after demonstrating that an index- $s$ -GAD procedure can find an SP of index  $s$ . We



**Fig. 3** A GAD trajectory leads from a LJ<sub>7</sub> minimum, Min, to a LJ<sub>7</sub> SP of index 1 (right). Shown is the energy profile over the reaction coordinate,  $R$



**Fig. 4** Energy profile over an index-4-GAD trajectory beginning at a LJ<sub>7</sub>-SP of index 3, to a LJ<sub>7</sub>-SP of index 4.  $R$  is an abstract 'reaction coordinate', the inlet is the shape of the SP<sub>*i4*</sub>

start near a corresponding SP of a deeper index, here the SP<sub>*i1*</sub>, after randomly perturbing the coordinates of the initial node. The initial choice of the second guide vector  $\mathbf{w}$  can be an eigenvector being orthogonal to the direction of the saddle valley. We automatically perform so many iterated searches up to a success: Here, we needed 11 iterations for an SP<sub>*i2*</sub> with  $E = -14.723$  units. At  $R = 2.85$  units, we get a convergence of  $|grad|_{max} = 1.4 \times 10^{-4}$ . The trajectory has 971 nodes, and additionally 47 steps are done for step length calculations. Here, nearly the full trajectory, up to  $R = 2.25$  units, is used to meander over the PES. There is  $|grad|_{max} = 0.1$ .

A further index-2-GAD test is done for the deflation ansatz of E and Zhou [5], Eqs. (16, 17) there. Because the Hessian is symmetric, we can use here  $3 \times 3n = 63$  dimensions only with the two guide vectors  $\mathbf{v}, \mathbf{w}$ . We get the same SP of index 2, in a third iteration, at  $R = 1.47$  units with the accuracy result  $|grad|_{max} = 0.1$ , and we need 121 steps. The ansatz works without an additional Schmidt orthogonalization of the test vectors  $\mathbf{v}, \mathbf{w}$ .

The next task is the search of an SP<sub>*i3*</sub> starting (randomly) near the found SP<sub>*i2*</sub>. The further guide vectors  $\mathbf{v}_i$ ,



( $i = 2, 3$ ), can again be eigenvectors being orthogonally to the first direction of the saddle valley. We chose the eigenvector  $\mathbf{e}_2$  to the negative eigenvalue 2, and the eigenvector  $\mathbf{e}_9$  to the first-positive eigenvalue 9. Again index-3-GAD was successful: A trajectory with 328 nodes found an index-3 SP at  $E = -14.348$  units at an accuracy of  $|\text{grad}|_{\text{max}} = 0.1$ , at  $R = 0.55$  units. If one calculates further up to  $R = 4.45$  units, one gets an accuracy of  $|\text{grad}|_{\text{max}} = 4.7 \cdot 10^{-5}$  after 2068 nodes, where only three step length searches are done. The energy profile is here monotonically increasing. The second iteration was the good one, whereas the first again departed to a dissociative SP.

And so on: An index-4-GAD can find an SP of index 4, see Fig. 4, at  $E = -13.708$  units. Now the search space of the Runge-Kutta method has the dimension  $5 \cdot 3n = 105$ . The energy profile over the trajectory is shown. It enrolls the complicate search with some turning points of energy. Again the region of the  $\text{SP}_{i4}$  is already touched at  $R \approx 0.6$  units, but the remainder of the calculations, up to  $R = 6$  units, is done for the exact convergence to the SP with a threshold of  $|\text{grad}|_{\text{max}} < 2 \cdot 10^{-5}$ . The full trajectory makes 3,560 steps; note that the bullets in Fig. 4 are drawn at fixed  $R$ -steps. They do not reflect the much more steps of the GAD trajectory.

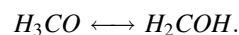
A conclusion of this test is the generalized Gad can find SPs of any index of the PES. It is a determined and purposeful method. To search for an SP of index  $s$ , one should start near an SP of the lower index ( $s - 1$ ). However, any other initial nodes are possible.

### 6.3 $\text{H}_3\text{CO}$ , methoxy radical

The radical  $\text{H}_3\text{CO}$  is an important, prototypical part of combustion chemistry [32]. A deep exploration of its PES was done by Knyazev 2002 [33], see also references therein. The radical consists of a methyl group bound to oxygen [34]. Its ground state has  $C_s$  symmetry; thus, the umbrella symmetry  $C_{3v}$  of  $\text{CH}_3$  is slightly changed. It is an electron-donating group and can cause an organic compound to become less acidic. Although this group exerts a negative inductive effect, its positive mesomeric effect is, however, greater and thus dominates, and so this group is an electron-donating group. The radical has two radical electronic states. In the  $\sigma$ -bond, the electron is localized in

an orbital collinear with respect to the CO-bond, and in the  $\pi$ -electronic state, the electron is localized in the orbital (SOMO) orthogonal to the CO-bond.

Note that we do not treat the isomerization to the hydroxymethyl radical [35, 36] here in this work



The used method is UHF, and the basic set is B3LYP 6-311G(p,d). The minimum of the methoxy radical is given (in ANG,DEG internal coordinates) in Table 1

The  $\text{H}_2\text{CO}+\text{H}$  van der Waals minimum is given in Table 2. Note that the B3LYP basis set does not well describe such a van der Waals minimum; however, this is meaningless for the test of the GAD method.

The SP number 1 ( $\text{SP}_1$ ) of index 1 is given in Table 3. It is the threshold for an escaping H-atom. The SP number 2 ( $\text{SP}_2$ ) of index 1 is given in Table 4. The eigenvector of this SP to the col direction is the lowering of the bond of the second escaping H-atom of the full radical. Thus, the  $\text{SP}_2$  forms an unsymmetric structure for the dissociation of an  $\text{H}_2$  part.

Three index-1-GAD trajectories are shown in Fig. 5 where we used formulas (8) and (9) for the GAD method. The trajectories show the three H-atoms for three different GAD runs for the radical  $\text{H}_3\text{CO}$ . One trajectory finds  $\text{SP}_1$ , one finds  $\text{SP}_2$ , and one goes to the valley of the van der Waals minimum at  $\text{H}_2\text{CO}+\text{H}$  without crossing an SP. The right triple of curves (bullets) depicts the distances of the three H-atoms for a trajectory from the minimum to the  $\text{SP}_1$  of index 1. In the representation of Fig. 5, the bullets now depict the real nodes of the calculated trajectory. One can clearly see the advance of the used Runge-Kutta method DOP853: The steplength is variable. At curved pieces, the steps are denser. On a straight pathway are large steps, but near the SP the steps become very short, and the SP is exactly located. The method stops there after convergence to a given threshold.

The region of a valley-ridge inflection (VRI) is indicated here as well. It is a bifurcation (bif) of a valley between the two SPs which is not directly attracted by GAD trajectories. It is detected by Newton trajectories, see ref. [37]. A next search trajectory (thin lines), however, overcomes the barrier between minimum and dissociation to  $\text{H}+\text{H}_2\text{CO}$ , but does not meet an SP. The left depicted minimum in the Figure is  $\text{H}_2\text{CO}$ , but the corresponding third, the single-dissociating H escapes over the frame. Thus, this GAD trajectory misses its aim to converge at an SP. It connects two minimum regions without meeting an SP. The calculation is stopped after the trajectory leaves the frame of the Figure. The case opens a gap of GAD: If it converges to a node, then it is a saddle point, but it converges not in every case.

The middle trajectory (dashes + bullets) meets the  $\text{SP}_2$  of index 1. It is the pathway to the (unsymmetric)

**Table 1**  $\text{H}_3\text{CO}$ : z-matrix for the minimum

C						
O	1	1.380				
H	1	1.089	2	106.07		
H	1	1.087	2	111.69	3	-117.81
H	1	1.087	2	111.63	3	117.72

**Table 2** H<sub>3</sub>CO: coordinates of a van der Waals minimum, see text

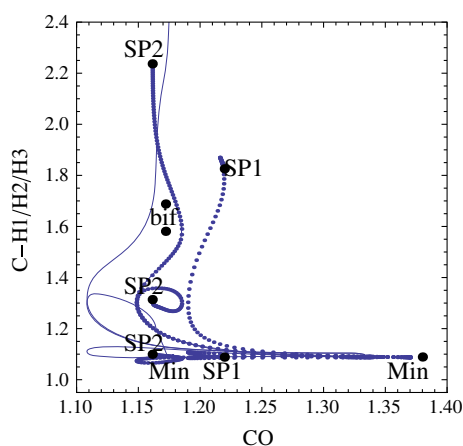
1.178		
1.095	122.09	
1.095	122.10	−180.0
4.020	54.69	180.0

**Table 3** H<sub>3</sub>CO: coordinates of the SP number 1

1.220		
1.825	100.18	
1.088	121.07	−95.85
1.088	121.07	95.85

**Table 4** H<sub>3</sub>CO: coordinates of the SP number 2 (SP2)

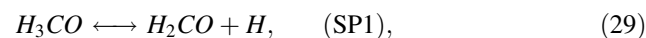
1.161		
1.328	121.31	
1.1	125.86	−180.0
2.319	120.67	0.0



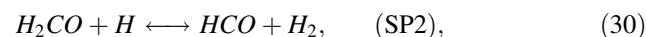
**Fig. 5** Three GAD trajectories for the three H-atoms, correspondingly. An index −1 SP search for radical H<sub>3</sub>CO has been created. The right triple of curves (*bullets only*) depicts the distances of the three H-atoms to the C-atom for a trajectory from the minimum (being at bottom right) to the SP1 of index 1. It is the pathway to the dissociation to H<sub>2</sub>CO+H. A next trajectory (*thin lines*), the left triple of curves, does not meet any SP though it passes nearby to SP2. The left 'minimum' in the Figure is H<sub>2</sub>CO, and the corresponding single-dissociating H escapes over the upper frame. The middle trajectory (*dashes + bullets*) meets the SP2 of index 1. It is the pathway to the dissociation to HCO+H<sub>2</sub>. (Compare ref. [33]; Fig. 4 there)

dissociation to HCO+H<sub>2</sub>. The trajectory shows the advance of the Runge-Kutta method even more. Near the SP, the steps become very short, and the trajectory can 'spirally' include the SP which again is exactly located (In the case of a fixed step length of the Runge-Kutta method; however, the GAD trajectory would circumvent the SP2 in

an infinite, false 'limit cycle'). Note that also the SP2 is a result of a direct GAD search for SPs of index 1, compare ref. [33] where this was not possible. The GAD trajectory goes across the 'ascending' ridge from SP1 over which the direct dissociation should go to HCO+H<sub>2</sub>. We find the SP2 for this reaction by a direct trajectory which 'misses' the SP1 (Compare ref. [38] for another 2-dimensional picture of such an SP constellation). There are now the SPs of index 1 for the reactions



and with ref. [33], the rearrangement



which is also found coming from the minimum H<sub>3</sub>CO. However, a direct symmetric dissociation



should go over a putative SP of index 2, corresponding to Fig. 5 of ref. [39]. It forms a 'summit' in the corresponding two-dimensional section. It is a task for GAD to search this higher index SP.

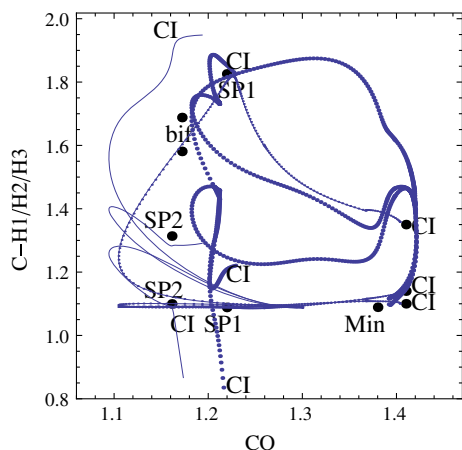
Diverse GAD<sub>2</sub> search tests, however, go to regions with an extremely negative eigenvalue of the Hessian; however, the norm of the gradient does not converge to zero: This is a hint for a conical intersection (CI) point nearby [19]. Usually, a CI point has two negative eigenvalues, like an SP of index 2. Thus, the search along a 'full two-dimensional direction' may find CI points, as well [19]. The difference to an SP is that at the CI, the gradient is not zero and the GAD system will not converge. But a strong indicator for the corresponding region is, behind the large negative first eigenvalue, that the Runge-Kutta procedure extremely shortens the steplength. The reason is the large eigenvalue, indicating a strong curvature of the PES. This causes a short steplength. For example, we find a CI point at a node with coordinates given in Table 5. The Hessian at the CI node has two negative eigenvalues:  $\lambda_1 = -5.42$  and  $\lambda_2 = -0.748$ . The first one is larger than 'normal' by a factor of 20.

Three index-2-GAD trajectories leading to CI points are shown in Fig. 6. None of the trajectories cross the region of the VRI bifurcation because only one H-atom moves in this region; however, the VRI concerns all three H-atoms. Usually, the CI points form a high-dimensional seam [19] on the intersecting PESs, so the CI points which we met may be only representants of this seam, or even of different seams.

A next test with the formulas of E and Zhou [5], Eqs. (15), (17) there, was done for a search of SPs of index 2: Here, the vectors **v** and **w** stay orthogonal; however, we again do not find any stationary points. In contrast, again CI points are localized.

Note: Using the GAD ansatz, Eqs. (11)–(13), we found that sometimes the vectors  $\mathbf{v}$  and  $\mathbf{w}$  do not stay strictly orthogonal, but the search can be successful, nevertheless, without a strict orthogonalization of both vectors. The Schmidt orthogonalization (SO) makes the GAD method safer; however, it may also disturb the steplength search of procedure DOP853. That can cause a bigger effort, besides the effort of the orthogonalization itself.

A last observation in this subsection: If the GAD trajectory crosses a region of a VRI bifurcation, then at least



**Fig. 6** Three GAD trajectories for the three H-atoms, correspondingly. An index 2 search for SPs of radical  $\text{H}_3\text{CO}$  has been created. Every triple of curves depicts the distances of the three H-atoms to the C-atom. The first trajectory (thin bullets and dashes) from the minimum touches the SP1 of index 1. Then, it goes further uphill to a CI point on the right-hand side at  $r_{\text{CO}} \approx 1.42$ . The second trajectory (thin lines only) starts at  $r_{\text{CO}} \approx 1.3$ , it finds quite another CI point at  $r_{\text{CO}} \approx 1.16$ . Two H-atoms are near the SP2 positions; however, the third H-atom is not. So, the CI is far from the SP2 (after crossing the CI, the trajectory continues on an upper surface. There it is stopped). The third trajectory (fat bullets only) starts at  $r_{\text{CO}} \approx 1.4$ . It ends at a CI point at  $r_{\text{CO}} \approx 1.22$  where one H-atom is near the position of SP1, but the other two H-atoms diverge from the SP1 position

**Table 5**  $\text{H}_3\text{CO}$ : coordinates of a CI

1.418		
1.113	102.53	
1.259	85.90	−214.09
1.165	150.65	102.65

**Table 6**  $\text{NH}_5$ : z-matrix for the minimum

N						
H	1	1.0				
H	1	1.0	2	106.99		
H	1	2.868	2	111.25	3	121.99
H	1	1.0	2	106.99	4	123.63
H	1	3.604	3	111.43	5	123.85

one curvature of the PES is nearly zero: To 'feel' the true pathway, in this case, the DOP853 procedure also uses very tiny steplengths. Thus, one needs a large effort to cross such a region. It is quite the same like in the contrary case of large curvatures at CI nodes. But here it seems to be a gap of the DOP853 procedure.

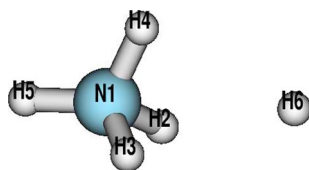
Conclusion: Apparently, an index-2-SP of the PES does not exist in the interesting region, compare also ref. [39]. On the contrary, many Conical Intersections (CI) emerge there. They can also be indicated by the GAD method. In contrast to the LJ clusters in Subsect. 6.2 where one single, unique PES exists, the relations in real molecules are quite more complicated.

#### 6.4 $\text{NH}_5$

Is the  $\text{NH}_5$  structure possible?  $\text{NH}_4^+$  is formed by the spare pair of electrons on the nitrogen atom bonding to an  $\text{H}^+$  ion. This is a co-ordinate bond where both electrons come from the nitrogen. There are now no electrons left on the nitrogen to form either another regular covalent or a co-ordinate bond. To add another hydrogen, one had to form a co-ordinate bond from the hydrogen, which needed two electrons there, a hydride ( $\text{H}^-$ ) ion. Even when one can get these to interact, this forms an ionic bond instead of a pure covalent bond. It is an 'ammonium hydride' salt, rather than a true  $\text{NH}_5$  compound. Unfortunately, this salt is stable only at a high pressure. When one decreases the pressure, it simply decomposes back into regular ammonia and hydrogen again.

Here, we treat the  $\text{NH}_3 \cdots \text{H}_2$  complex under a pyramidal inversion of the ammonia molecule, see Ref. [15]. The used method to calculate the PES is RHF, and the basic set is 6-311G(3d,2p). The van der Waals minimum  $\text{NH}_3 \cdots \text{H}_2$  is given by its z-matrix (in ANG, DEG internal coordinates) in Table 6. A calculation along the GAD<sub>2</sub> ansatz Eqs. (11–13) starts anywhere nearby to successfully search the SP of index 2 [15], see Fig. 7. Note: We circumvent possible SPs of index 1 by passing directly from the minimum region to the SP<sub>2</sub>.

At the SP of index 2, we find the first eigenvalues of H (in Bohr, Radiant system) −0.036, −0.030, and 0.036. The run needs the effort: Function calls 6007, steps 451, accepted steps 348, and steplength searches 103.



**Fig. 7**  $\text{NH}_5$ : SP of index 2

**Table 7**  $\text{CH}_5^-$ : z-matrix for the symmetric SP of index 1, see Fig. 8

C						
H	1	1.733				
H	1	1.061	2	90.0		
H	1	1.061	2	90.0	3	120.0
H	1	1.061	2	90.0	4	120.0
H	1	1.733	2	180.0	5	119.8

## 6.5 $\text{CH}_5^-$ radical

There is a recent paper on this radical, see ref. [16]. The used method is RHF, and the basic set is 6-311G(3d,2p). In Table 7, we report an SP (in ANG,DEG internal coordinates) with  $D_{3h}$ -symmetry [40].

In a first result, we find this high lying  $\text{SP}_1$  of index 1 if we start midway on the ridge after a valley-ridge inflection (VRI) uphill to the  $\text{SP}_1$ , see ref. [16]. We start here above the VRI region to avoid endless tiny steps, see the former Subsect. 6.3 for an explanation.

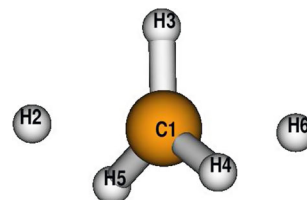
A second result for  $\text{CH}_5^-$  is a so-called roundaboutway: We start anywhere in the plane of a kind of van der Waals minimum of the molecule. However, the trajectory comes back to a minimum structure in the 'plane' of a dissociating H from the molecule  $\text{CH}_4$ . There is a turning point (TP) at the emerging mountains of the PES, and then the trajectory returns, like in the strange case of the MB test [8]. At the TP, we find the first three eigenvalues of H (in ohr,Radiant system)  $-0.009$ ,  $0.001$ , and  $0.018$ . Thus, the GAD trajectory 'finds' a region with one negative eigenvalue, but turns back. GAD goes over the TP further and further, but then back to the minimum. GAD is a dynamical system method. It is programmed that the system trajectory locally holds a given direction. GAD may not necessarily converge globally [5].

A third result concerns the search of an SP of index 2. We again start anywhere in the plane of a kind of a van der Waals minimum of the molecule.  $\text{GAD}_2$  works successfully, it ends at the  $\text{SP}_{12}$  given in Table 8.

At the SP, the first three eigenvalues of H are  $-0.004$ ,  $-0.001$ , and  $0.001$ . The small first eigenvalues show that the PES of this radical is extremely flat, see [16]. The run needs the following effort: Function calls 8478, steps 671, accepted steps 365, but many steplength searches: 306.

**Table 8**  $\text{CH}_5^-$ : Internal coordinates of the SP of index 2 (structure 3 in ref. [16])

1.085		
1.079	110.13	
1.079	110.16	117.47
1.085	109.21	121.30
3.871	54.09	122.15



**Fig. 8**  $\text{CH}_5^-$ : symmetric SP of index 1

## 7 Discussion and conclusion

We have shown the applicability of the GAD method to small molecules. Two different kinds of coordinate systems are used: First Cartesian coordinates including the projection out of the zero-eigenvectors of the Hessian in the GAD ansatz, or, second, internal coordinates where the curvilinear metric has to be included in the GAD formulas.

The GAD method is a good 'growing string' method for the exploration of a PES. We found all SPs of index 2 of interest, or the method indicates a CI case. The sticking point is the handicap of the initial node, and the initial guide vectors. Thus, one has to guess an abstract region of attractivity of the corresponding SP (which is, of course, usually unknown). To looking to an SP of interest, we have to aim in the correct direction. In ref. [7], the authors worked with an 'available good initial guess': A coarse minimum energy path is determined before by a double-ended string method. If no SP of index  $s$  is found, then we cannot say that there is no SP of the searched index. Because GAD is 'only' a growing string method, we have no indication of how many SPs of a given index it overlooks. In the case of Newton trajectories [11], we know that they connect SPs of an index difference of one in the regular case. This does not fit here. We report a  $\text{GAD}_2$  trajectory from the minimum region to an SP of index 2.

In our tests, not in every case, the searched kind of SP is really found; the GAD trajectory can have turning point(s) and can go 'wrong' on a way back. However, if one knows what to look for then the method can be useful. If the true region of the SP is found, then an exact localization of the SP is usually not very effective. One may jump automatically to another method, in such a case. The convergence of the dynamical system of GAD is slow,



compare also ref. [9]. However, if GAD converges, then it converges to every given threshold.

**Acknowledgments** Financial support from the Spanish Ministerio de Ciencia e Innovación, DGI project CTQ2011-22505 and, in part from the Generalitat de Catalunya projects 2009SGR-1472 is fully acknowledged. We thank the referees for suggestions and comments.

## Appendix: Hessian matrix of the Lennard-Jones cluster

The potential energy surface (PES) of an  $n$ -atomic Lennard-Jones (LJ) cluster is as follows:

$$V(x_1, y_1, z_1, \dots, x_n, y_n, z_n) = 4\epsilon \sum_{i=1}^n \sum_{j=i+1}^n \left[ \sigma^{12} \left( \frac{1}{r_{ij}} \right)^{12} - \sigma^6 \left( \frac{1}{r_{ij}} \right)^6 \right] \quad (32)$$

where the double sum can be shortened to  $\sum_{i < j}$ , and

$$r_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 \quad (33)$$

is the square of the distance between atoms  $i$  and  $j$ .  $\epsilon, \sigma$  are parameters of the current cluster.

$V(\mathbf{q})$  is a  $3n$ -dimensional surface over  $\mathbf{R}^{3n}$ . Let  $\mathbf{grad}(\mathbf{q})$  be its gradient vector,  $\nabla_{\mathbf{q}} V(\mathbf{q})$ , and let  $\mathbf{H}(\mathbf{q})$  be the matrix of its second derivatives, the Hessian  $\nabla_{\mathbf{q}} \mathbf{grad}^T$ . We could not find the calculation of the Hessian in the older literature; thus, we give it here in the appendix. We get for the distance between two atoms

$$\frac{\partial r_{ij}}{\partial x_i} = \frac{x_i - x_j}{r_{ij}}, \quad \frac{\partial r_{ij}}{\partial x_j} = \frac{x_j - x_i}{r_{ij}}. \quad (34)$$

A single component of the gradient becomes

$$\frac{\partial V}{\partial x_i} = 4\epsilon \sum_{i \neq j}^n (x_i - x_j) \left[ -12\sigma^{12} \left( \frac{1}{r_{ij}} \right)^{14} + 6\sigma^6 \left( \frac{1}{r_{ij}} \right)^8 \right]. \quad (35)$$

If we put the abbreviation

$$T1(ij) =: \left[ -12\sigma^{12} \left( \frac{1}{r_{ij}} \right)^{14} + 6\sigma^6 \left( \frac{1}{r_{ij}} \right)^8 \right] \quad (36)$$

and if we set for simplification of the summation  $T1(ii) = 0$ , then we can write

$$\frac{\partial V}{\partial x_i} = 4\epsilon \sum_{j=1}^n (x_i - x_j) T1(ij), \quad (37)$$

and analogously we have for  $i = 1, \dots, n$

$$\begin{aligned} \frac{\partial V}{\partial y_i} &= 4\epsilon \sum_{j=1}^n (y_i - y_j) T1(ij) \quad \text{and} \\ \frac{\partial V}{\partial z_i} &= 4\epsilon \sum_{j=1}^n (z_i - z_j) T1(ij). \end{aligned} \quad (38)$$

We further define the abbreviation

$$T2(ij) =: \left[ 12 * 14\sigma^{12} \left( \frac{1}{r_{ij}} \right)^{16} - 6 * 8\sigma^6 \left( \frac{1}{r_{ij}} \right)^{10} \right]. \quad (39)$$

Again, we set  $T2(ii) = 0$ . Then, single components of the Hessian are for the same atom

$$H(x_i, x_i) = \frac{\partial^2 V}{\partial x_i^2} = 4\epsilon \sum_{j=1}^n \left( T1(ij) + (x_i - x_j)^2 T2(ij) \right), \quad (40)$$

and

$$H(x_i, y_i) = \frac{\partial^2 V}{\partial x_i \partial y_i} = 4\epsilon \sum_{j=1}^n (x_i - x_j) * (y_i - y_j) T2(ij), \quad (41)$$

and analogously one gets  $H(x_i, z_i)$  and  $H(y_i, z_i)$  and their symmetric counterparts. If different atoms with numbers  $i$  and  $j$  are involved, then we get

$$H(x_i, x_j) = \frac{\partial^2 V}{\partial x_i \partial x_j} = -4\epsilon \left( T1(ij) + (x_i - x_j)^2 T2(ij) \right), \quad (42)$$

and

$$H(x_i, y_j) = \frac{\partial^2 V}{\partial x_i \partial y_j} = -4\epsilon (x_i - x_j) * (y_i - y_j) T2(ij), \quad (43)$$

and analogously one gets  $H(x_i, z_j)$  and  $H(y_i, z_j)$  and their symmetric counterparts.

## References

- Mezey PG (1987) Potential energy hypersurfaces. Elsevier, Amsterdam
- Heidrich D, Kliesch W, Quapp W (1991) Properties of chemical interesting potential energy surfaces. Springer Lecture Notes in Chemistry 56, Berlin
- Wales DJ (2003) Energy Landsc. Cambridge Univ Press, Cambridge
- Heidrich D (1995) In: Heidrich D (ed) The reaction path in chemistry: current approaches and perspectives. Kluwer, Dordrecht
- Weinan E, Zhou X (2011) Nonlinearity 24:1831
- Samanta A, Weinan E (2012) J Chem Phys 136:124104
- Li T, Zhang P, Zhang W (2013) Multiscale Model Simul 11:385
- Bofill JM, Quapp W, Caballero M (2013) Chem Phys Lett 583:203
- Zeng Y, Xiao P, Henkelman G (2014) J Chem Phys 140:044115
- Crippen GM, Scheraga HA (1971) Arch Biochem BioPhys 144:462
- Quapp W (2003) J Theor Comput Chem 2:385

12. Hratchian HP, Schlegel HB (2005) Theory and Applications of Computational Chemistry: the first forty years, Dykstra CE, Frenking G, Kim KS, Scuseria GE (eds) p195 Elsevier, Amsterdam.
13. Heidrich D, Quapp W (1986) Theor Chim Acta 70:89
14. Khait YG, Panin AI, Averyanov AS (1994) Int J Quant Chem 54:329
15. Minyaev RM, Getmanskii IV, Quapp W (2004) Russ J Phys Chem 78:1494
16. Minyaev RM, Quapp W, Schmidt B, Getmanski IV, Koval VV (2013) Chem Phys 425:170
17. Maugière FA, Collins P, Ezra GS, Wiggins S (2013) J Chem Phys 138:134118
18. Nagahata Y, Teramoto H, Li C-B, Kawai Sh, Komatsuzaki T (2013) Phys Rev E Stat Nonlin Soft Matter Phys 88:042923
19. Quapp W, Bofill JM, Caballero M (2012) Chem Phys Lett 541:122
20. Doye JPK, Wales DJ (2008) J Chem Phys 116:3777
21. Schmidt MW, Baldrige KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347
22. Gordon MS, Schmidt MW (2005) Theory and applications of computational chemistry: the first forty years. In: Dykstra CE, Frenking G, Kim KS, Scuseria GE (eds) Advances in electronic structure theory: GAMESS a decade later. Elsevier, Amsterdam, p 1167
23. Miller WH, Handy NC, Adams JE (1980) J Chem Phys 72:99
24. Quapp W, Heidrich D (1984) Theor Chim Acta 66:245
25. Tachibana A (1981) Theor Chim Acta 58:301
26. Parlett BN (1980) The symmetric eigenvalue problem. Prentice-Hall, Englewood Cliffs NJ
27. Hairer E, Nørsett SP, Wanner G (1993) Solving ordinary differential equations I nonstiff problems. 2nd Ed. Springer, Series Computat Math 8, see also [www-unige.ch/~hairer/software.html](http://www-unige.ch/~hairer/software.html)
28. Hedar A-R (2013) Global optimization test problems. Technical report, see also [www-optima.amp.i.kyotou.ac.jp/member/student/hedar/Hedar\\_files/TestGOhtm](http://www-optima.amp.i.kyotou.ac.jp/member/student/hedar/Hedar_files/TestGOhtm)
29. Mühlenbein H, Schomisch D, Born J (1991) Parallel Comp 17:619
30. Ghosh K, Ozkan SB, Dill KA (2007) J Am Chem Soc 129:11920
31. See [www-doye.ch.cam.ac.uk/networks/LJn.html](http://www-doye.ch.cam.ac.uk/networks/LJn.html)
32. Balucani N, Leonori F, Bergeat A, Petrucci R, Casavecchia P (2011) Phys Chem Chem Phys 13:8322
33. Knyazev VD (2002) J Phys Chem A 106:8741
34. Carter JT, Cook DB (1991) J Mol Struct Theochem 251:111
35. Kamarchik E, Rodrigo C, Bowman JM, Reisler H, Krylov AI (2012) J Chem Phys 136:084304
36. Ryazanov M, Rodrigo C, Reisler H (2012) J Chem Phys 136:084305
37. Schmidt B, Quapp W (2012) Theor Chem Acc 132:1305
38. Bofill JM, Quapp W, Caballero M (2012) J Chem Theor Comput 8:927
39. Marcy TP, Diaz RR, Heard D, Leone SR, Harding LB, Klippenstein SJ (2001) J Phys Chem A 105:8361
40. Ishida K, Morokuma K, Komornicki A (1977) J Chem Phys 66:2153