Journal of Theoretical and Computational Chemistry
© World Scientific Publishing Company

THE GROWING STRING METHOD FOR FLOWS OF NEWTON TRAJECTORIES BY A SECOND ORDER METHOD

WOLFGANG QUAPP*
Mathematical Institute, University of Leipzig
Johannisgasse 26, Postfach 10 09 20
D-04009 Leipzig, Germany
quapp@rz.uni-leipzig.de

Received 18. 03. 2008
Revised 24. 04. 2008
Accepted Day Month Year

The reaction path is an important concept of theoretical chemistry. We use a definition with a reduced gradient [see W. Quapp et al., Theoret. Chem. Acc. 100, 285 (1998)], also named Newton trajectory (NT). To follow a reaction path, we design a numerical scheme for a method for finding a transition state (TS) between reactant and product on the potential energy surface: the growing string (GS) method. We extend the method [see W. Quapp, J. Chem. Phys. 122, 174106 (2005)] by a second order scheme for the corrector step which includes the use of the Hessian matrix. A dramatic performance enhancement for the exactness to follow the NTs, as well as a dramatic reduction of the number of corrector steps are to report. So we can calculate flows of NTs. The method works in nonredundant internal coordinates. The corresponding metric to work with is curvilinear. The GS calculation is interfaced with the GamessUS package (we have provided this algorithm on a web page). Examples for applications are the HCN isomerization pathway and NTs for the isomerization C7ax ↔ C5 of alanine dipeptide.

Keywords: Potential energy surface; reaction path following; projected gradient; curvilinear metric; internal coordinates.

1. Introduction

The concept of the minimum energy path (MEP) or reaction path (RP) of an adiabatic potential energy surface (PES) is the usual approach to the theoretical kinetics of larger chemical systems.[1,2,3,4,5,6] It is roughly defined as a line in a coordinate space, which connects two minimums by passing the saddle point (SP) or the transition structure (TS) of a PES. The energy of the SP is assumed to be the highest value tracing along the RP. It is the minimal energy a reaction needs to take place. Reaction theories are based on the knowledge of the RP either implicitly

*Telephone: [49] 341-97 32153 Fax: [49] 341-97 32199 web: www.math.uni-leipzig.de/∼quapp

(transition state theory,[1]) or explicitly (variational transition state theory[5]). The latter theory requires local information about the PES along the RP.

The idea of an RP circumvents the dimension problem for large molecules: it is impossible to fully calculate their PES. Any parameterization, s, of the RP $\mathbf{x}(s)=(x^1(s),\ldots,x^n(s))^T$, $n \leq (3N-6)$ is called reaction coordinate. n is the dimension of the problem. It is an uncertainty of the RP definition, how a reaction path ascends to the SP: which is the direction a reaction has to start with? We use here the distinguished or driven coordinate method[7] in the modern form of the NT formulation.[8,9,10,11,12,13] This kind of curves can also be determined uphill. We remark that the search for an appropriate MEP is not necessarily equivalent to the finding of the steepest descent (SD) pathway from the SP,[14,15] the intrinsic reaction coordinate (IRC) which is determined downhill.

The GS method is a combination of predictor steps followed by iterative corrector loops which approximate a node on the searched NT. The exactness of the approximation is set by an ϵ threshold that, of course, also predetermines the order of the number of corrector cycles. For the predictor we use here the direction between two known stationary points (usually minimums). For the definition of different NTs we use different search directions. There are members of a family of NTs, the picture of a flow of curves. The spreading of different NTs allows an insight into the valley formation between the stationary points. A deep valley concentrates the NTs in a narrow tube, where, on the other hand, a valley bifurcation spreads the NTs over a greater region.

It should be noted that the term RP is rather conventional when used for NTs, in the usual case of NTs without a turning or bifurcation point between minimum and TS: every such NT rises monotonically in energy from the minimum to the TS (or vice versa) and fulfills the natural definition of an RP.[16] On the other hand, the emergence of a valley bifurcation is indicated by the bifurcation of the NTs, where the single IRC often may go down through the region without any hint of a problem.

In this paper we execute the corrector step by a long known second order method.[8,9,17] The original proposal for the GS method for NTs,[18,19] on the other hand, was a first order corrector. The reason for the original proposal resulted from the fear of convergence of the second order corrector after a large predictor of the GS step. If one uses a Newton-like method one should always have a good starting value for the corrector iteration. We circumvent the problem of a possible divergence of the first corrector step by a restriction of the step length of each corrector step.

The following sections describe the predictor step of GS, the definition of NTs, and the corrector step including the forced use of a curvilinear metric of z-matrix coordinates. We apply the method to some NT calculations for two sample molecules, and close the paper by a discussion of the results.

The programs used can be obtained on request or retrieved.[20]

2. Predictor Steps along the GS method

We use the linear combination of curvilinear z-matrix coordinates of the current chain point and the final point.[19] We search an RP which should connect the initial minimum, $\mathbf{x}_{ini}$=$\mathbf{x}_0$, with the end $\mathbf{x}_{fin}$=$\mathbf{x}_{m+1}$ by a chain of m nodes $\mathbf{x}_k$. We calculate successive nodes beginning at the initial minimum by a number of copies of the original molecule, somehow interpolated between the current structure, $\mathbf{x}_k$, and the product structure. Since $\mathbf{x}_k$ is an approximate node on the RP obtained by the last corrector loop, we choose a next guess point, $\mathbf{y}_{k+1}$, of the string between the actual node and the final minimum by the predictor step

$$\mathbf{y}_{k+1} = \lambda_k \, \mathbf{x}_k + (1 - \lambda_k) \, \mathbf{x}_{fin} \ , \quad \lambda_k = \frac{m-k}{m+1-k} \ , \quad k = 0, .., m-1 \ . \tag{1}$$

Of course, the formula can only work if $\mathbf{x}_k$ and $\mathbf{x}_{fin}$ have the same sequence of coordinates inclusive the same kind of torsional angles. If the RP was a straight line, the linear combination would create equidistant nodes on the line if we successively set back $\mathbf{x}_{k+1} = \mathbf{y}_{k+1}$ without the corrector and repeat ansatz (1). If the really calculated chain of nodes deviates strongly from the straight line between the start minimums, then the length of the predictor steps becomes larger and larger in the region of the end of the chain. Then the last part of the chain may be useless. However, a solution is simply to interchange start and final minimum and to use the two starting "halves" of the corresponding chains, see the example alanine dipeptide below.

In the state of the method, the use of unscaled nodes, as they come out from predictor and corrector, is a good indicator for any kind of problems along the pathway of interest. Usually the RP is curvilinear, so the predictor nodes are not necessarily fully equidistant under this linear ansatz. We have abstained from constructing a double increasing strain[21,22] meeting at the TS, because we are firstly interested in following the exact NT of interest, and only secondly in getting the TS on the top of this NT. Because we use the Hessian, see below, we can calculate a Newton test step at every node, to guess the distance to the next stationary point. The subsidiary step then indicates the passage of the TS.

3. Definition of NTs

The starting point is a geometrically defined pathway which may serve as a reaction path. Geometrically defined means that only properties of the PES are taken into account like in the definition of the SD, however, no dynamical behavior of the molecule is used. It was proposed to choose a distinguished coordinate along the valley of the minimum, to go a step in this direction, and to perform an energy optimization of the residual coordinates.[7,23,24] Some years ago, this venerable method was transformed into a new mathematical form.[8,9] The concept is that any preselected gradient direction is fixed

$$\mathbf{g}(\mathbf{x})/|\mathbf{g}(\mathbf{x})| = \mathbf{r}, \tag{2}$$

where $\mathbf{r}$ is the unit vector of the search direction, and $\mathbf{g}$ is the gradient of the PES. The search direction may correspond to an assumed start direction of a chemical reaction. For example, it may be the direction between the two minimums of reactant and product. Or it can be the direction along a valley of the minimum, or any other direction. The property (2) is realizable by a projection of the gradient onto the (n−1)-dimensional subspace which is orthogonal to the one-dimensional subspace spanned by the search direction $\mathbf{r}$. A curve belongs to the search direction $\mathbf{r}$, if the gradient of the PES always remains parallel to the direction $\mathbf{r}$ at every point along the curve $\mathbf{x}(s)$ with a suitable parameter s

$$\mathbf{P_r}\,\mathbf{g}(\mathbf{x}(s)) = \mathbf{0} \tag{3}$$

where $\mathbf{P_r}$ projects out of the search direction $\mathbf{r}$. This means $\mathbf{P_r}\,\mathbf{r}$=$\mathbf{0}$. The projector $\mathbf{P_r}$ in Eq. (3) "reduces" the gradient. It is build by $(n-1)$ row vectors being orthogonal to the search direction $\mathbf{r}$. They can be obtained by a modified Gram-Schmidt orthogonality.[25] Thus, $\mathbf{P_r}$ is an $(n-1)\times n$ matrix of rank $(n-1)$. Note that projection operators have a long tradition in theoretical chemistry.[26,27,28] However, the older proposals are not directed to any definition of an RP.

To calculate a point on a curve (3) we need the tangential vector (see also Section 4 below). The derivative of Eq. (3) creates the tangent vector $\mathbf{t} = \mathbf{x}'$ to the curve

$$\mathbf{0} \;=\; \frac{d}{d\,s}[\mathbf{P_r}\;\mathbf{g}(\mathbf{x}(s))] = \mathbf{P_r}\;\frac{d\,\mathbf{g}(\mathbf{x}(s))}{d\,s} = \mathbf{P_r}\;\mathbf{H}(\mathbf{x}(s))\;\mathbf{t}(s)\;. \tag{4}$$

The matrix $\mathbf{H}$ is the Hessian at point $\mathbf{x}(s)$. Equation (4) is a linear equation for the tangent vector $\mathbf{t}$ at the current point. The system is solved by QR decomposition.[17]

For the calculation of the corrector step, we augment the reduced Hessian $\mathbf{P_r}\,\mathbf{H}$, an $(n-1)\times n$ matrix, by the tangent vector to an $(n\times n)$ matrix which is the so-called $\mathbf{K}$ matrix.[17] The $\mathbf{K}$ matrix is the coefficient matrix of a linear system of equations. The right side of the system of equations forms the (negative) reduced gradient now augmented by a zero in the last, n-th entry. Corrector steps are applied if the (Cartesian) norm of the reduced gradient $\mathbf{P_r}\,\mathbf{g}$ of Eq. (3) is greater than a threshold ϵ which is given as a parameter. The corrector step is a kind of a Newton-Raphson step being orthogonal to the tangent.[8] But compare here a further trick proposed in ref.[29] which we will also use: one may distort the pure orthogonality to avoid that corrector loops extinguish the predictor.

To solve the linear $\mathbf{K}$-system, it has to be nonsingular. A way to this property is: the used coordinates have to be nonredundant. It must be $n \leq 3N-6$. Note that n can be smaller than $3N-6$. In this case we speak of a smaller set of collective coordinates.

In contrast to their pure geometric definition, NTs are like molecular dynamics curves: a family of NTs to different directions go over the same TS but may fill the space between the minimum and the TS. However, in contrast to dynamical trajectories, all NTs of the same family concentrate at the same SP. It is simply

a consequence of the definition of the NTs: at the stationary points the gradient vector is the zero vector. All search directions can flow into the zero vector if the gradient disappears. If one imagines the NT to be a curve of the system point driven by a force in a fixed direction, then the NT may indeed fill the gap between the IRC and dynamical trajectories, cf.[14]

4. Metric of z-matrix coordinates and corrector iteration

We work in a curvilinear coordinate space of n non-redundant z-matrix coordinates where, usually, it is $n = 3N - 6$ for non-linear N-atomic molecules, but any non-complete subset with $n < 3N - 6$ may also be used. Measurements of distances or angles have to be done in the corresponding metric. Points or distance vectors are contravariant vectors which have to be measured with the metric matrix $(g_{i\,j})$, thus

$$|\mathbf{x}|^2 = \sum_{i,j=1}^{n} x^i \, g_{i\,j} \, x^j \ . \tag{5}$$

Gradients are covariant vectors where the inverse metric matrix $(g^{i\,j})$ is appropriate.

$$|\mathbf{g}|^2 = \sum_{i,j=1}^{n} g_i \, g^{i\,j} \, g_j \tag{6}$$

is the norm of the gradient vector in Eq.(2). The inverse metric matrix is computed by

$$(g^{ij}) = \left(\sum_{r=1}^{n+6} B(i,r) B(j,r) \right) \tag{7}$$

using the (n,n+6)-B-matrix of the given internal coordinates which is available in the GamessUS[30] if one additionally puts the following orders in the corresponding commando lines:[31]

CONTRL EXETYP=binv COORD=zmt ,
STATPT HESS=calc , and
FORCE METHOD=analytic PRTIFC=.true.

The B-matrix is the matrix of derivatives of internal coordinates to Cartesian coordinates (without massweighting). From the GamessUS output file, one can grasp the gradient vector and the B- and Hessian matrices of interest by *grep* command; we use it together with an aid program.[20] The values are usually in Bohr and radian, or their combinations, correspondingly. The analytical Hessian is calculated at every step, to clear the possibilities of our method, and to exclude possible problems from this side. Of course, in a routine calculation, it may be updated.

Usually, the search direction $\mathbf{r}$ is built by a difference of two point vectors, thus it has to be measured like a contravariant vector. This also applies to the tangent of the NT. A modified Gram-Schmidt orthogonality[25] for the projection matrix

orthogonal to $\mathbf{r}$ is done in the $(g_{i\,j})$-norm, corresponding to the contravariant form of $\mathbf{r}$. Then the rows of $\mathbf{P_r}$ are also contravariant vectors. This is appropriate to the covariant character of the gradient in Eq.(3). As well as to the Hessian which is two-fold covariant. It is reduced by $\mathbf{P_r}\ \mathbf{H}$ which is the definition for (n-1) lines of the $\mathbf{K}$-matrix. The character of these lines remains covariant. The last line of $\mathbf{K}$ should be the tangent vector from Eq.(4), however, now turned to the covariant form by multiplication of the tangent with $(g_{i\,j})$.

$$t_{cov\,i} = \sum_{j=1}^{n} g_{i\,j}\, t^j \ , \ \mathrm{i} = 1, ..., \mathrm{n} \ . \tag{8}$$

The projection of the gradient becomes the right hand side of the corrector equation: it equals to be $\mathbf{P_r}\ \mathbf{g}$, which is only an $(n-1)$-dimensional vector. It is to be used in a Cartesian metric. It is additionally augmented in the last position (number n) by a zero to build the $\mathbf{K}$-system,[17] a linear equation for the corrector step

$$\begin{pmatrix} \mathbf{P_r}\ \mathbf{H} \\ \mathbf{t}_{cov}^T \end{pmatrix} * CorrStep = \begin{pmatrix} -\mathbf{P_r}\ \mathbf{g} \\ \delta \end{pmatrix} \quad . \tag{9}$$

The length of the corrector step is given automatically with Eq.(9), however, if starting an iteration it is sometimes important to restrict the length to be less than the predictor step. Originally, δ is zero in Eq.(9). If the NT of interest deviates very strongly from the search direction, then it may be that the corrector step extinguishes the predictor. In this case we do not execute the corrector step, but we chose in a second calculation δ to be $0.9\,sin(\alpha)$ times the calculated step length of the corrector, where α is the angle between the predictor and the former tangent of the NT. Then the corrector search goes obliquely to the tangent, and possibly, it does not extinguish the predictor. Note that the scaling of nodes can also be changed: if one concentrates on the calculation of the TS, one may use more nodes in that region,[22] or one may employ a double-ended search method.[32]

The corrector step is a contravariant vector which has to be added to the current point $\mathbf{x}_k$. If a corrector loop is finished, we determine the integrated step length up to the current node, k, with $\mathbf{s}_k = \mathbf{x}_k - \mathbf{x}_{k-1}$ by the new step length piece

$$\mathrm{rpl_k} = (\sum_{i,j} \frac{1}{2}(g_{ij}(k) + g_{ij}(k-1))s_k^i s_k^j)^{1/2} \tag{10}$$

which is added to the reaction path length calculated up so far.

5. Examples

It is natural to ask how the second order corrector including the intrinsic metric compares to the previous method[18,35] in terms of performance. An objective comparison does not seem straightforward since there are different possibilities of internal coordinates. We will try the comparison for the 22-atomic Alanine dipeptide. All calculations have been carried out with the GamessUS[30] suite of programs

for a PC employing the standard 6-31G basis set. It should be noted that the aim of the example alanine dipeptide was to compare the second order GS method with former GS calculations and not to determine quantitative results for RPs or TSs. In this case, the calculations would have to be repeated with higher-level quantum mechanical basis sets. The RHF calculations used here deliver the analytical gradient and Hessian in internal z-matrix coordinates which we directly use in our program.

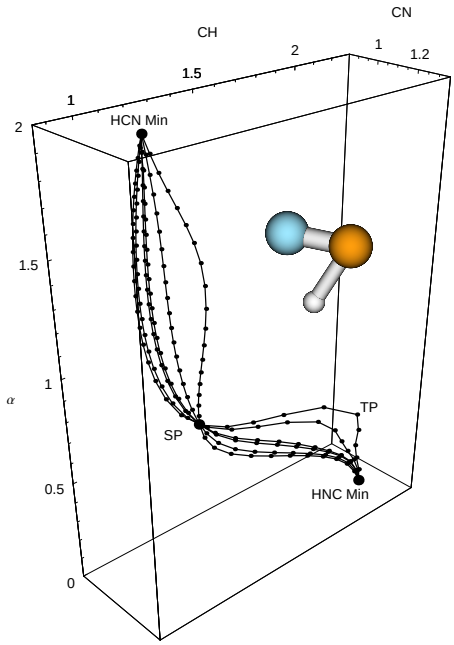

Fig. 1. Toy model HCN. Approximations of NTs (on the PES) of HCN isomerization given in the 3D configuration space of distances CH and CN (in Å), and angle α (scaled by 1/90). The inlay shows the TS. 30 nodes are used. TP is a turning point.

5.1. *A low dimensional case: HCN*[33]

We use the HCN 6-31G potential surface as a first illustrative model. The isomerization pathway of HCN serves as a numeric example. Figure 1 shows calculated NTs in the HCN configuration space in internal coordinates. The distances of CH and CN are given in Å, and α is the angle between the bonds used with a factor of 1/90.

The search direction $\mathbf{r}$ of Eq.(2) can vary for different NTs. We have used for the three coordinates CH, CN, and angle(HCN) fixed values for CN=0.1, angle=1.0, but CH is varied by {-2,-1,-0.1,0.1,1,1.5}. The search is from top to bottom in Figure 1. For the predictor steps we use Eq.(1), and we put $\epsilon = 0.000\,01$ Å for the threshold of

the corrector. (This ϵ is very small.) The GS method leads from the HCN minimum up in energy to the SP, but then downhill in energy to the next minimum, the HNC valley. (The calculation is not terminated at SP to immediately explore the full RP.) The effort is for the sum of corrector steps: {137, 103, 102, 104, 108, and 120}, correspondingly. The reaction path length, Eq.(10), of the shortest NT in the center of the tube is 3.959. It is the NT which belongs to the search direction (0.1,0.1,1). This NT may be the true RP.

Figure 1 shows the expected results: The NTs go along a "tube" and meet at the SP, as well as at the final minimum. The CN distance for all NTs is nearly constant along the pathway. The CH distance has in the extreme cases a bump. The spreading of the tube gives an impression of the depth of the reactive valley: the smaller the tube the deeper is the valley. If the tube spreads a little more then the valley is a little broader. The two outermost NTs with CH parts of -2.0, or +1.5 in $\mathbf{r}$, explore something of the side slope of the PES. On the HCN side of the problem, all NTs increase monotonically in energy, thus all NTs are models of an RP. Only the extreme NT to $\mathbf{r}$=(1.5,0.1,1.0) has a turning point at the slope of the HNC basin, after the TS. The energy profile has a maximum at the TP which is higher then the energy of the TS. Thus, this one NT is not a model of an RP, on its second piece after the TS.

All NTs are nice, smooth curves through the TS. The smoothness of all NTs is analogous to the smoothness of the IRC, the steepest descent curve from TS. (This is in contrast to Fig.1 of ref.[34]) In the case of HCN, the discussed inclusion of the metric, Eqs.(5) to (10), works well.

5.2. *A higher dimensional case: Alanine dipeptide*

Alanine is one of the most common proteinogenic amino acids. Applications of the GS method are done to alanine dipeptide (Acetyl-L-Ala-NHMe),

$$CH_3CO\text{-}NHCHCH_3CO\text{-}NHCH_3.$$

Here we use the usual nickname "dipeptide", though it is a single amino acid with one Cα-atom only in the center, and two caps. But it contains two peptide bonds, C-N, between the two peptide units, the row of atoms O,C,N, and H. Work[19,35] uses NTs to represent RPs for the 66-dimensional configuration space of an alanine dipeptide isomerization, at the lower 3-21G level, and at the quantum chemical level of DFT calculation, B3LYP/6-31G. Maeda and Ohno[36] note that a TS between the two conformers C5 and C7ax of alanine dipeptide is deviated from the midpoint between them. The choice of R = C7ax and P = C5 was made to obtain higher computational efficiency according to the Hammond rule. Applications of the growing string (GS) with a nudged elastic band were also recently done to this system.[21] The method is a MEP optimization. It searches the steepest descent curve (IRC) from the saddle (without guessing the SP in before).

For the larger molecule alanine dipeptide, the use of the full metric matrices

(g_{ij}) and (g^{ij}) does not work well. (We do not understand fully, why.) The diagonal entries in the inverse metric matrix, g^{ii}, have the ordinary values between 0.6 and 2.0. However, the determinant of the matrix is very small, typically $\approx 10^{-17}$, or smaller. It is strange for a metric matrix. Does it come from some elements in any adjoining diagonals? One may speculate that the calculation of the metric matrix (g_{ij}) by matrix inversion from (g^{ij}) becomes erratically, at least in some places. Though the matrix multiplication of $(g^{ij}) \times (g_{ij})$ gives the unit matrix to an exactness of 10^{-6}, or better, one usually gets a (g_{ij}) matrix with some very large entries: Some single entries step over a value of 10^3 which exceeds any expectation of a metric matrix. Also the Hessian has diagonals which, by the calculation in Bohr units, are nearly all less than one. Thus also the determinant of this (60,60)-matrix is very small. It may be a next source of numeric errors. The combination of numeric metric- and Hessian uncertainties may cause that the ratio of the norms of gradient and of reduced gradient runs into false regions. Convergence of corrector loops is missing. The method often diverges. The geometry of the corrector loops often goes wrong.

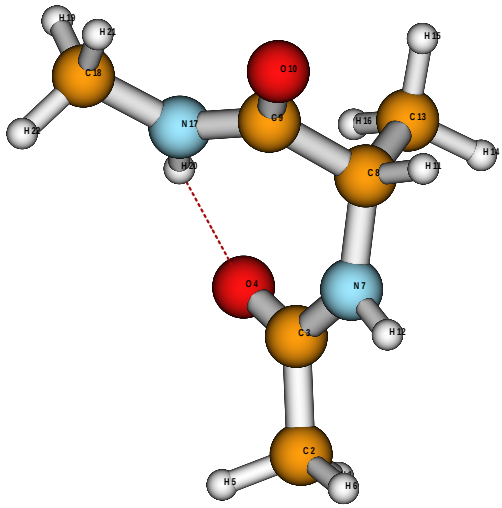

Fig. 2. Numbering of atoms of alanine dipeptide by Chass et al.[37] for use in polyspherical z-matrix coordinates.

One can test some ways out:

(i) We may use the "false" Cartesian metric of a unit matrix for (g_{ij}), as well as (g^{ij}). This brings results here: the method works at all (much to our own surprise). The reason may be that the turning of the alanine torsional angles

between C7ax and C5 is an RP which concerns only one sort of coordinates: even some torsional angles. Nearly all distances and all other bond angles are unchanged. Thus, we do not really need the curvilinearity of the z-matrix coordinates, here in this case.

(ii) We may alternately use a cutted metric: only the diagonal entries of g^{ii}, of the inverse metric matrix are used, and we put zeros in the non-diagonal elements: $g^{ij} = 0$, for $i \neq j$. Thus, we neglect the coupling of mixed derivative terms.[38] The metric matrix simply is the diagonal matrix $g_{ii} = 1/g^{ii}$. The artificial simplification works well, still better then in case (i). Using this "false" metric which is concentrated on the diagonals, results in convergent corrector steps, and indeed acceptable NTs for the isomerization in question. They cross at the TS and can be used to find the TS. However, methods **(i)** and **(ii)** work with an incorrect metric, and the resulting NTs are somewhat "lopsided".

(iii) An observation under a missing convergence of the corrector loop by a full use of metric is an enormous internal rotation of the two outer CH_3-groups of the molecule. The CH_3-groups should stay fixed along the RP in question. Such degrees of freedom do not play any role in the theory of NTs, if they are orthogonal to the search direction. However, here, they are not quiet. The six degrees of freedom of the torsional angles of six H atoms disturb the corrector direction. We may solve the problem by fixing the two CH_3-groups: we fix dih4, the torsional angle between atoms (O4,C3,C2,H1), see Fig.2, and dih22 of (H22,C18,N17,C9), to their start values, and apply the GS method in the remaining 58 dimensions, only. We get the other dihedrals at fixed positions also. Thus, the two CH_3-groups come into the usual role of spectator-groups for the NTs between C5 and C7ax. The small trick makes convergent corrector loops, and nice NTs result. Note: in GamessUS we have to use furthermore the variables number NZVAR=60 in the CONTRL line, but we give in the STATPT line now IFREEZ=6,60 and fix the two dihedrals dih4 and dih22 in the z-matrix.

We had found in the older method[19] for alanine dipeptide that ϵ=0.002 5Å was a possible value for the convergence threshold of the corrector. Here with the second order corrector and the metric in 58 dimensions only, in case **(iii)**, we can use the threshold = 0.000 1Å, and nevertheless, we usually need only 2 or 3 corrector steps in a corrector loop! We present six calculations of NTs with different search directions, two different transit directions, and a number of 30 string nodes for every NT. We use for the predictor the linear combination of curvilinear z-matrix coordinates of the current chain point and the final point[19] with Eq.(1). In a former proposal for the GS method,[35] we needed approximately 40 corrector iterations per node for convergence. Here this is reduced by a factor of 10 to 20, to 2 or 3 corrector steps. Two iterations could be seen to be the "theoretical" minimum of iterations one has to do for an approximation method: one for the first large corrector step to the NT of interest, and a second step, usually very small, to exactly fulfill a very

strong ϵ condition. With such a nice convergence, we are able to calculate much more nodes per NT, as well as much more NTs itself.

The coordinates of alanine dipeptide[37] are used as in ref.[35], see Fig.2. Figure 3 illustrates results of the GS method with 30 nodes. We project all results into the plane of dih9=ϕ, the angle of (O9,C8,N7,C3), and dih17=ψ, the angle of (N17,O9,C8,N7), where an underlying thin grid of level lines is also drawn. The level lines are calculated by setting a raster of tuples, (ϕ,ψ), and optimizing all other degrees of freedom. Thus, the level lines may be in some details not fully consistent with the region where the NTs are going. We first calculate the C5 side of the problem (in contrast to ref.[36]), because in this molecule, we find a qualitative difference between the two minimum regions. The C5 region has direct NTs as model reaction pathways between minimum and TS. A broad valley from C5 leads directly to the TS of interest. There the energy profile of the NTs increases monotonically. This will not hold for the C7ax-side of the problem. For the search direction $\mathbf{r}$ we use for NT no 6 the direction between C5 and C7ax. The soft degrees of freedom, ϕ and ψ, are coupled also to further dihedral turnings. Especially, the five torsional angles which are involved in the isomerization are: dih9=ϕ, dih10, dih11, dih13, and dih17 = ψ. The five coordinates describe the transition from C5 to C7ax, or vice versa. The entries are (in Radian)

$$\begin{pmatrix} dih9 \\ dih10 \\ dih11 \\ dih13 \\ dih17 \end{pmatrix} = \begin{pmatrix} 0.416 \\ -0.474 \\ 0.435 \\ 0.442 \\ -0.463 \end{pmatrix} .$$

For the search directions for NTs no 1 to 5, coordinates are specialized for those five torsional angles only, the angles are varied for $\mathbf{r}$, but all other coordinates are set to zero. The non-zero entries are

$$\begin{pmatrix} 0.416 \\ 0.0 \\ 0.435 \\ 0.442 \\ 0.250 \end{pmatrix} , \begin{pmatrix} 0.416 \\ 0.0 \\ 0.435 \\ 0.442 \\ 0.075 \end{pmatrix} , \begin{pmatrix} 0.416 \\ 0.0 \\ 0.435 \\ 0.442 \\ 0.050 \end{pmatrix} , \begin{pmatrix} 0.416 \\ 0.0 \\ 0.435 \\ 0.442 \\ 0.0 \end{pmatrix} , \begin{pmatrix} 0.416 \\ -0.075 \\ 0.435 \\ 0.442 \\ -0.075 \end{pmatrix} ,$$

correspondingly. The dihs 9, 11, and 13 are on the pattern of no 6, where dih10 and 17 are changed. (Note that after such settings the $\mathbf{r}$ has still to be normalized.) After starting in the C5 region, from left down to above right, the NT curves are numbered by no 1 to no 6. Trajectories no 1 to 5 lead to the TS (all by 15 nodes). NT no 1 has a high turning point (TP). It indicates a ridge of the PES. A TP is an extremizer of the energy profile of the NT. This trajectory is not an RP model. Its energy profile has a maximum higher than that of the TS. One can see it by comparing the level lines which the NT crosses. For this NT, we additionally show the first 20 predictor steps by smaller bullets connected to the nodes of the NT. The predictor steps all point to the C7ax minimum as it is predicted by Eq.(1).

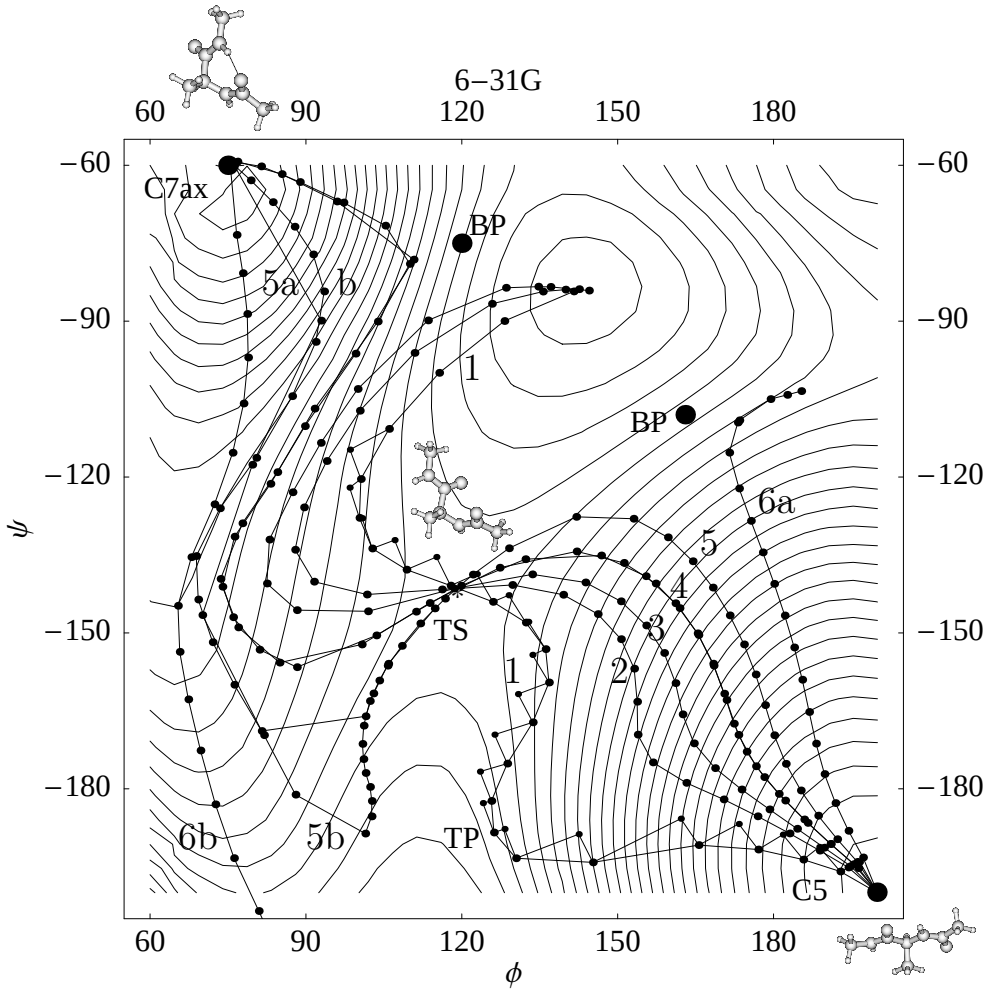

Fig. 3. A flow of Newton trajectories by the GS method for alanine dipeptide between the minimum C5 (start of a numbers) and minimum C7ax (start of b numbers) calculated with ϵ=0.000 1. Different search directions are used to calculate different NTs, see text. The flow is projected from 58 dimensions into the (ϕ , ψ)-plane, 30 nodes are used for the strings. The 20 first predictor steps for the NT no 1 are shown additionally (connected small bullets). They point to C7ax. Five of the NTs meet at the TS (*). However, after the TS the NTs are spread by the C7ax valley form with a shoulder. BPs are bifurcation points. TP is a turning point.

The corrector goes, in some cases, somewhat at an angle, because δ in Eq.(9) is not zero.

The NTs no 2 to 5 are equivalent models for an RP from C5 to TS. They lead continuosly to the TS, and their ascenting energy profile is monotone. In contrast, the NT no 6a bifurcates to the right hand side (and is stopped there: the predictor

points to C7ax, the corrector goes in the contrary direction). A bifurcation point (BP) separates families of NTs. It is the valley-ridge-inflection point of the PES.[10] (Note: the BPs in Fig.3 are only guessed BPs, which are guessed by the behavior of the NTs. The aim of this paper is not the calculation of such points.)

For NTs 1 to 5, we need for the 15 nodes up to the TS the following summary of corrector iterations: 51, 45, 43, 52, and 41, correspondingly. A typical run needs 45 corrector steps, thus in average approximately 3 steps per node.

On the C7ax side of the TS, NTs no 1 to 3 go downhill after the TS, however, after a TP they turn uphill and find a SP of index 2, the hill in Fig.3. It is a correct behavior of NTs: they connect stationary points of an index difference of one. NT no 4 also goes downhill, it also has a TP, but after a second TP it finds the minimum, C7ax. But note that the two TPs of NT no 4 are less in energy than the TS. Thus, this NT can be a semi-well adapted model of an RP.

NT no 5 seems to have four TPs. It leads to a TP near the TS, then slightly uphill, and after a further TP, it goes downhill, but not "directly" to C7ax. It has to circumvent two further TPs. The second TP after the TS is higher than the TS. The NT is no RP model. It seems that no monotonically decreasing NT exists at all, directly between TS and minimum C7ax. (Note: it is possible that between a minimum and its "visible" TS not a direct NT exists, see ref.[39], Fig.8.)

At the C7ax minimum, the dih9=ϕ is the strongest torsion. The main valley of the PES leads along ψ from C7ax to "South". The TS of interest to C5 is just in a side valley. The quasi-shoulder in the main valley seems to make all the complicate NTs. Trajectories to the search direction of no 4 to 6 are also calculated with the start in this minimum, C7ax. The NTs no 4b and 5b meet the others exactly which start in C5. One may see it by twice the number of nodes along the pathways. NT no 6b follows the main valley from C7ax to "South", and does not meet the TS. Maybe it also has a TP and comes back, but this can happen only after the lower edge of the figure where the NT is cut.

At all, the picture of the NTs no 1 to 5 develops propitiously from C5 up to the TS. They form a flow of curves. However, after the TS, the flow separates into 2 "half-good" curves (no 4 and 5), and three curves (no 1-no 3) which go the other way. The reason seems to be a structure like a quasi-shoulder, in the C7ax valley, and the valley-ridge-inflection point, the upper BP in the Figure. The valley of the TS downhill is only a side valley which anywhere flows into the main valley. The calculated NTs indicate it.

The pathway of every NT-approximation is not rescaled; the shown curves with their nodes are the direct outcome of the corrector loops. Because of our very small ϵ, we can assume that the calculated nodes are exact points of the theoretical NTs. Because the curves of the NTs are not a staight line between C5 and C7ax, we sometimes "lose" some distance at the end, and the predictor steps with Eq.(1) become larger and larger: compare the NTs no 4a,b and 5a,b. We accept this here

to be within the nature of the construction. We simply calculate the corresponding NTs at the C7ax side by a new series of NTs starting at C7ax.

The roundabout way curve no 5 in Fig.3 is no RP model, at the C7ax side of the TS, because it has a TP higher than TS. But it indicates a possible adventure of the GS method. The jump of NT no 5a indicates that the GS method is adapted to our aim, to find a possible pathway between two stationary points of character minimum, also by NTs. The GS machinery helps to leave the blind valley (in some cases). In the inverse direction it does not work so well: the NT 5b does not well cut the corner, and it needs 28 nodes to reach the TS. (It is cut there.) A predictor for the NT 5a along the tangent of that NT[8,9] would stay anywhere in the blind valley and possibly not reach the next basin. In any case we again have a flow of curves which indicates the character of the PES. The TPs of diverse NTs indicate the difficult character of the C7ax valley, and its quasi-shoulder. The separation of the NTs at the BPs creates a clear guess for the valley-ridge-inflection points of the PES.

In Table 1 we compare the effort of the calculations in this paper with previous works. It shows the dramatic improvement of the second order method.

Table 1. Comparison of effort for alanine dipeptide.

Method	Threshold of corrector	# Nodes	# Corr.steps per node	# Chains
1.order[a]	0.008 0	8	47	1
1.order[a]	0.000 5	23	151	1
1.order[b]	0.003 5	5	43	1
1.order[b]	0.003 5	7	40	1
1.order[b]	0.003 5	9	31	1
2.order[c]	0.000 1	30	2-3	9

Note: there is a qualitative difference of the NT approximations in ref. 35, Fig.1, and here, also.
[a] ref. 19, basis set 3-21G only: it is another PES
[b] ref. 35
[c] this paper

6. Discussion

It is known that in spite of its great success IRC (the steepest descent from saddle point) is actually not completely universal.[40,41,42,43,44] NT is a pre-dynamical allegory for a chemical reaction path[45] like the IRC. So it is also an RP definition. An NT may be understood to be a line where the uphill driving force, or the downhill damping force, always have the same direction, the force which has to overpower the gradient (or to hold it back).

New and notable is: we have extended the former algorithms of a growing string (GS) method for NTs to a second order corrector. It works in nonredundant inter-

nal coordinates. The Newton trajectories can follow the RP definition, as exactly as one needs it, to find a corresponding SP by going uphill, or a minimum by going downhill from SP. NT curves are definably independent of the coordinate system.[33] Usually, if the metric of the coordinate system is correct, this needs a very small number of corrector steps per loop, 2 to 3. A much lower number than in the older GS programs for NTs. The NT based TS determination (in this paper a side effect only) has been very successful in elucidating the valley character at a level of detail that otherwise would be inaccessible.

Two examples illustrate the usefullness of NTs. The standard example of the HCN isomerization shows the existence of the isomerization valley with its flow picture. The flow of NTs does not spread too wide after the start at a stationary point, and at the next stationary point, it again flows together. Along the valley floor way, there the curvature of level-hypersurfaces is maximal. A great range of gradient directions occurs in a small region of the PES. Thus, the NTs to corresponding gradient directions concentrate in this small region. They form a "tube". Usually, the flow of NTs is a fine background for the IRC. It surrounds the IRC. The flow is a skeleton for the MEP where the NT with the lowest integrated path length, Eq.(10), may be the MEP.

This can also be observed in the alanine dipeptide example in the C5 basin: some different NTs concentrate in the valley between minimum and TS. We have a flow which forms a tube around the steepest descent from TS. However, here we find a different situation at the other side of the TS. The C7ax basin has a main valley in "North-South" direction, but the SP of interest is situated above a valley which flows into the main valley from a side direction. Additionally, the main valley has a quasi-shoulder. Here the NTs spread over a wide region of the PES. But of course, some meet again at the SP region. The fan of NTs may indicate that possible reactions from C7ax may often go along other ways, but do not meet that SP.

The turning point (TP) case[7,23,46,47] divides NTs into those which can serve as loose RPs, and others: if the NT does not contain a TP at the pathway from minimum up to the SP, as in the C5 basin, it may be used as an RP model. However, if a TP emerges, it can have an energy higher than the energy of the respective SP and, hence, this path does not meet the meaning of an RP. This is the case at the C7ax side left from the TS, in Fig.3, NT no 5, as well as NT no 1 in the C5 basin.

A corrector with gradient descent, or conjugate gradient, needs very larger numbers of corrector iterations.[18,19] However, one advantage is the possibility to work with redundant coordinates, like Cartesians. Here under the Eq.(9) we have to work with non-redundant coordinates. In the contrary case the equation system (9) does not work.

Next to the problem of defining a suitable curve for the reaction path of chemistry is the possibility of RP branching. The corresponding points are bifurcation

points. C7ax and the TS are not joined by a narrow valley, however, there are ramified reaction valleys. The branching of Newton trajectories is connected to the valley-ridge inflection points of the PES. The method of following a reduced gradient has succeeded in computing symmetric VRI points.[10] (Un-symmetric VRIs can be found by following the valley extremals.[48]) The GS method can also give a hint to BPs by observing a corresponding flow of NTs, see Fig. 3. If predictor steps are "large" enough, the GS can jump over a BP, or can jump over a kink, and follow the next branch of the NT. It is a global advantage over a predictor along the tangent of the NT.

Acknowledgments

I thank Dr. A.A. Granovsky for some hints to use the GamessUS program, and Profs. Dr. E. Kraka and Dr. D. Cremer for the discussion of some examples of lower dimension, cf. ref.[45], which are not reported in the paper, but which made it possible to better understand the numerical problems of the NTs of alanine dipeptide. The author thanks a referee for a suggestion.

References

1. K. Laidler, *Theory of Reaction Rates* (McGraw-Hill, New York, 1969).
2. W. Miller, N.C. Handy, J.E. Adams, J. Chem. Phys. **72**, 99 (1980).
3. P.G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987).
4. D. Heidrich, W. Kliesch, W. Quapp, *Properties of Chemically Interesting Potential Energy Surfaces* (Springer, Berlin, 1991).
5. D. Heidrich, *The Reaction Path in Chemistry: Current Approaches and Perspectives* (Kluwer, Dordrecht, 1995).
6. W. Quapp, D. Heidrich, Theor. Chim. Acta **66**, 245 (1984).
7. I.H. Williams, G.M. Maggiora, J. Mol. Struct. (Theochem) **89**, 365 (1982).
8. W. Quapp, M. Hirsch, O. Imig, D. Heidrich, J. Comp. Chem. **19**, 1087 (1998).
9. W. Quapp, M. Hirsch, D. Heidrich, Theor. Chem. Acc. **100**, 285 (1998).
10. M. Hirsch, W. Quapp, D. Heidrich, Phys. Chem. Chem. Phys. **1**, 5291 (1999).
11. J.M. Anglada, E. Besalu, J.M. Bofill, R. Crehuet, J. Comp. Chem. **22**, 387 (2001).
12. J.M. Bofill, J.M. Anglada, Theor. Chem. Acc. **105**, 463 (2001).
13. R. Crehuet, J.M. Bofill, J.M. Anglada, Theor. Chem. Acc. **107**, 130 (2002).
14. S.C. Ammal, H. Yamataka, Eur. J. Org. Chem. 4327 (2006).
15. J.G. López, G. Vayner, U. Lourderaj, S.V. Addapalli, S. Kato, W.A. deJong, T.L. Windus, W.L. Hase, J. Am. Chem. Soc. **129**, 9976 (2007).
16. M. Hirsch, W. Quapp, J. Molec. Struct., THEOCHEM, **683**,1 (2004);
    M. Hirsch, W. Quapp, J. Math. Chem. **36**, 307 (2004).
17. E. L. Allgower, K. Georg, *Numerical Continuation Methods - An Introduction* (Springer, Berlin, 1990).
18. W. Quapp, J. Comp. Chem. **25**, 1277 (2004).
19. W. Quapp, J. Chem. Phys. **122**, 174106 (2005).
20. W. Quapp, *E-mail:* quapp@rz.uni-leipzig.de
    Web: www.math.uni-leipzig.de/∼quapp/gs2teord
21. B. Peters, A. Heyden, A.T. Bell, A. Chakraborty, J. Chem. Phys. **120**, 7877 (2004).

22. S.K. Burger, W. Yang, J. Chem. Phys. **127**, 164107 (2007).
23. K. Müller, L.D. Brown, Theor. Chim. Acta **53**, 75 (1979).
24. J. Cioslowski, A.P. Scott, L. Radon, Mol. Phys. **91**, 413 (1997).
25. A. Kiełbasiński, H. Schwetlick, *Numerische Lineare Algebra* (Deutscher Verl. Wiss., Berlin, 1988).
26. P. Pulay, "Direct use of the gradient for investigating molecular energy surfaces," in: Application of Electronic Structure Theory, ed. H.F. Schaefer (Plenum, New York, 1977, p.153).
27. W. Miller, N.C. Handy, J.E. Adams, J. Chem. Phys. **72**, 99 (1980).
28. D.-h. Lu, M. Zhao, D.G. Truhlar, J. Comp. Chem. **12**, 376 (1991).
29. M. Hirsch, W. Quapp, J. Comp. Chem. **23**, 887 (2002).
30. M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S.J. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, J. Comp. Chem. **14**, 1347 (1993).
31. A.A. Granovsky, PC GAMESS version 7.1, http://classic.chem.msu.su/gran/gamess/index.html
32. E.F. Koslover, D.J. Wales, J. Chem. Phys. **127**, 134102 (2007).
33. W. Quapp, J. Math. Chem. **36**, 365 (2004).
34. J. Gonzales, X Gimenez, J.M. Bofill, J. Comp. Chem. **28**, 2102 (2007).
35. W. Quapp, J. Comp. Chem. **28**, 1834 (2007).
36. S. Maeda, K. Ohno, Chem. Phys. Lett. **404**, 95 (2005).
37. G.A. Chass, M.Sahai, J.M.S. Law, S. Lovas, Ö. Farkas, A. Perczel, L.-L. Rivail, I.G. Csizmadia, Int. J. Quant. Chem. **90**, 933 (2002).
38. J. Stare, J. Chem. Inf. Model. **47**, 840 (2007).
39. W. Quapp, J. Theor. Comp. Chem. **2**, 385 (2003).
40. T.B. Woolf, Chem. Phys. Lett. **289**, 433 (1998).
41. V.D. Knyazev, J. Phys. Chem. **106**, 8741 (2002).
42. R. Meyer, T.-K. Ha, Mol. Phys **101**, 3263 (2003);
43. M.Hirsch, W. Quapp, Chem. Phys. Lett. **395**, 150 (2004).
44. C.S. Tautermann, A.F. Voegele, K.R. Liedle, J. Chem. Phys **120**, 631 (2004);
45. W. Quapp, E. Kraka, D. Cremer, J. Phys. Chem. **111**, 11287 (2007);
H. Joo, E. Kraka, W. Quapp, D. Cremer, Mol. Phys. **105**, 2697 (2007).
46. W. Quapp, J. Comp. Chem. **22**, 537 (2001).
47. D. Lauvergnat, A. Nauts, Y. Justum, X. Chapuisat, J. Chem. Phys. **114**, 6592 (2001).
48. W. Quapp, M. Hirsch, D. Heidrich, Theor. Chem. Acc. **112**, 40 (2004);
W. Quapp, J. Mol. Struct. **695-696**, 95 (2004).