The Barnes Update Applied in the Gauss-Newton Method: an Improved Algorithm to Locate Bond Breaking Points

Josep Maria Bofill,[*,†,‡] Rosendo Valero,[*,¶,‡] Jordi Ribas-Ariño,[*,¶,‡] and Wolfgang Quapp[*,§]

†*Departament de Química Inorgànica i Orgànica, Secció de Química Orgànica*

and

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

¶*Departament de Ciència de Materials i Química Física, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain,*

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

—- December 17, 2020 —-

E-mail: jmbofill@ub.edu; rosendo.valero@ub.edu; j.ribas@ub.edu; quapp@math.uni-leipzig.de

Abstract

A mechanochemical reaction is a reaction induced by mechanical energy. A general accepted model for this type of reactions consists in a first order perturbation on the associated potential energy surface (PES) of the unperturbed molecular system due to mechanical stress or pulling force. Within this theoretical framework, the so-called *optimal barrier breakdown points or optimal bond breaking points* (BBPs) are critical points of the unperturbed PES where the Hessian matrix has a zero eigenvector that

coincides with the gradient vector. Optimal BBPs are 'catastrophe points' that are particularly important because its associated gradient indicates how to optimally harness tensile forces to induce reactions by transforming a chemical reaction into a barrierless process. Building on a previous method based on a nonlinear least squares minimization to locate BBPs (Bofill et al., *J. Chem. Phys.* **2017**, *147*, 152710-10), we propose a new algorithm to locate BBPs of any molecular system based on the Gauss-Newton method combined with the Barnes update for the nonsymmetric Jacobian matrix, which is shown to be more appropriate than the Broyden update. The efficiency of the new method is demonstrated for a multidimensional model PES and two medium size molecular systems of interest in enzymatic catalysis and mechanochemistry.

1 Introduction

One of the main problems in theoretical chemistry is to study the mechanisms of chemical transformations. An important achievement in the development of models to understand the mechanisms associated with any chemical transformation was the introduction of the following concepts, namely, potential energy surface (PES) and reaction path (RP) as a way to describe the molecular system evolution from reactants to products in geometrical terms and the stationary points of the PES with respect to the coordinates, namely, transition states (TS) or first order saddle points (SP) and minima. In the last years, another type of points has increased its importance to explain some features of the chemical reactivity, namely, the valley-ridged inflection points (VRI) related with the bifurcations of the RP. The VRIs are points of the PES with different mathematical characteristics to those associated with stationary points. Finally, another type of points of the PES has recently turned out to be crucial in the theory that rationalizes the chemical transformations under mechanical forces, the so-called mechanochemistry.[1–12] From a conceptual point of view, the phenomenon of mechanical activation can be understood on the basis of the fact that the PES of a given reactive system changes when this system is subjected to tensile stress. As a result of these

force-induced changes of the PES, the barriers between the minima and the saddle points change.[13] When a mechanical force is applied to a molecular system, the reactivity is either enhanced or suppressed.[14] This is explained using the PES model by the modification of the barrier that separates two minima due to the magnitude and direction of the mechanical force. The easiest case is when the direction of the applied force, $\mathbf{f}$ is constant. The modified PES or force-transformed PES,[15] labeled as, $V_{\mathbf{f}}(\mathbf{x})$ is given by

$$V_{\mathbf{f}}(\mathbf{x}) = V(\mathbf{x}) - \mathbf{f}^T(\mathbf{x} - \mathbf{x}_0) \tag{1}$$

where $V(\mathbf{x})$ is the original potential, T means transposed, $(\mathbf{x}-\mathbf{x}_0)$ is the displacement vector and $\mathbf{x}_0$ is an anchor or reference point. The dimension of the vectors is N. If we expand up to second order the Taylor expansion around $\mathbf{x}_0$ the original potential, that is

$$V(\mathbf{x}) = V(\mathbf{x}_0) + \mathbf{g}(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) + 1/2(\mathbf{x} - \mathbf{x}_0)^T\mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \mathcal{O}(||\mathbf{x} - \mathbf{x}_0||^2)$$

where $\mathbf{g}(\mathbf{x}_0)$ and $\mathbf{H}(\mathbf{x}_0)$ are the gradient and the Hessian of $V(\mathbf{x})$ at $\mathbf{x}_0$, respectively, and $||\mathbf{x} - \mathbf{x}_0|| = [(\mathbf{x} - \mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0)]^{1/2}$. Substituting this expansion in Eq. 1 we have

$$V_{\mathbf{f}}(\mathbf{x}) = V(\mathbf{x}_0) + (\mathbf{g}(\mathbf{x}_0) - \mathbf{f})^T(\mathbf{x} - \mathbf{x}_0) + 1/2(\mathbf{x} - \mathbf{x}_0)^T\mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \mathcal{O}(||\mathbf{x} - \mathbf{x}_0||^2)$$

thus in this model the perturbation due to external constant force only affects to first order in $\mathbf{x}-\mathbf{x}_0$ with respect to the original PES. Regarding this expression we see that the stationary points of the $V_{\mathbf{f}}(\mathbf{x})$ occur when $\mathbf{g}(\mathbf{x}) = \mathbf{f}$. Because the force is constant in direction, we can write $\mathbf{f} = F\,\mathbf{l}$, where $\mathbf{l}$ is a fixed unit vector and F the magnitude. Due to the structure of the force vector, in the stationary points of $V_{\mathbf{f}}(\mathbf{x})$ the relation $||\mathbf{g}(\mathbf{x})|| - F = 0$ holds, and $\mathbf{g}$ is parallel to $\mathbf{l}$. The sequence of points which satisfy the relation that the gradient points to a constant direction, describes a curve, the so-called Newton Trajectory (NT). The tangent

of this curve is governed by the Branin formula[16]

$$\frac{d\mathbf{x}}{dt} = \pm\mathbf{A}(\mathbf{x})\mathbf{g}(\mathbf{x}) = \pm\mathbf{A}(\mathbf{x})\mathbf{l}||\mathbf{g}(\mathbf{x})|| \tag{2}$$

where $\mathbf{A}(\mathbf{x})$ is the adjoint matrix of the Hessian matrix, $\mathbf{H}(\mathbf{x})$ of the original PES, $V(\mathbf{x})$ and t is the parameter that characterizes the curve. The Branin expression, Eq. 2, is a way to construct the Force Displacement Stationary Points (FDSP).[13,14] To see this we compute how the force-transformed PES, Eq. 1, changes through the NT

$$\frac{dV_{\mathbf{f}}(\mathbf{x})}{dt} = (\mathbf{g}(\mathbf{x}) - \mathbf{f})^T(\frac{d\mathbf{x}}{dt}) = \pm(||\mathbf{g}(\mathbf{x})|| - F)(\mathbf{l}^T\mathbf{A}(\mathbf{x})\mathbf{l})||\mathbf{g}(\mathbf{x})|| \tag{3}$$

where in its derivation, Eqs. 1 and 2, the directional derivative has been used. Regarding Eq. 3, it holds $dV_{\mathbf{f}}(\mathbf{x})/dt = 0$ at point $\mathbf{x}$, in the next situations:

(*1*) when the magnitude of the applied force, F, is equal to the gradient norm, $||\mathbf{g}(\mathbf{x})||$, of the original potential, $V(\mathbf{x})$, or

(*2*) when $\mathbf{l}^T\mathbf{A}(\mathbf{x})\mathbf{l} = 0$, or

(*3*) when in the point $\mathbf{x}$ the original PES, $V(\mathbf{x})$, is stationary thus $||\mathbf{g}(\mathbf{x})|| = 0$.

The second situation occurs if the point $\mathbf{x}$ belongs to the valley-ridge-inflection manifold of the original PES.[17,18] When this condition is fulfilled, the magnitude of the applied external force is irrelevant to make $dV_{\mathbf{f}}(\mathbf{x})/dt = 0$ since it depends on the original PES. The NT which goes through a VRI point in its evolution is a singular NT and this type of NT plays an important role in mechanochemistry.[12]

The third condition is an artefact which comes in by the inserted $d\mathbf{x}/dt$ which is zero at the stationary points of the original PES. On the effective PES the stationary points are moved out of their original positions.

According to the previous requirement (*1*) the magnitude of the external force to be applied depends on the behavior of $||\mathbf{g}(\mathbf{x})||$ through the NT curve. For this reason it is important to analyze this norm by considering its variation or slope on the corresponding NT

$$\frac{d||\mathbf{g}(\mathbf{x})||}{dt} = \pm Det(\mathbf{H}(\mathbf{x}))||\mathbf{g}(\mathbf{x})|| \tag{4}$$

where in its derivation the directional derivative concept has been used, Eq. 2, and the relation $\mathbf{A}(\mathbf{x})\mathbf{H}(\mathbf{x}) = \mathbf{I}Det(\mathbf{H}(\mathbf{x}))$ where $\mathbf{I}$ is the unit matrix. The slope is zero in the stationary points and the points where $Det(\mathbf{H}(\mathbf{x})) = 0$ of the original PES. It is important to analyze the variation of $||\mathbf{g}(\mathbf{x})||$ around the manifold of points where $Det(\mathbf{H}(\mathbf{x})) = 0$. Let us assume that $\mathbf{x}_0$ denotes a point of this manifold. If the NT transverses this manifold through this point, then we have

$$||\mathbf{g}(\mathbf{x}(t))|| = ||\mathbf{g}(\mathbf{x}(t_0))|| + 1/2\frac{d^2||\mathbf{g}(\mathbf{x}(t))||}{dt^2}\bigg|_{t=t_0}(t-t_0)^2 + \mathcal{O}((t-t_0)^3) \tag{5}$$

where $\mathbf{x}(t_0) = \mathbf{x}_0$. Note that $d||\mathbf{g}(\mathbf{x}(t))||/dt|_{t=t_0} = 0$ according to Eq. 4 since $Det(\mathbf{H}(\mathbf{x}_0)) = 0$. Using Eq. (B4) of reference 19 and using as the tangent vector the Branin expression, Eq. 2, we have

$$\frac{d^2||\mathbf{g}(\mathbf{x}(t))||}{dt^2}\bigg|_{t=t_0} = \mu_j^2(\mathbf{e}_j^T \left\langle \mathbf{F}(\mathbf{x}_0)\right\rangle_j \mathbf{e}_j)(\mathbf{e}_j^T\mathbf{l})||\mathbf{g}(\mathbf{x}_0)||^2 \tag{6}$$

where $\mathbf{e}_j$ is the normalized eigenvector of the matrices $\mathbf{A}(\mathbf{x}_0)$ and $\mathbf{H}(\mathbf{x}_0)$ with eigenvalues μ_j and 0, respectively. The eigenvalue μ_j is the unique eigenvalue of $\mathbf{A}(\mathbf{x}_0)$ different from zero. The matrix $\left\langle \mathbf{F}(\mathbf{x}_0)\right\rangle_j$ results from the contraction of the third derivative tensor of the potential energy $V(\mathbf{x})$ with respect to the coordinates with the eigenvector $\mathbf{e}_j$. Assuming that we find the first manifold of points that satisfy $Det(\mathbf{H}(\mathbf{x})) = 0$ for all the NTs that emerge from the same minimum and go to the same SP, then it holds the value of $\mathbf{e}_j^T \left\langle \mathbf{F}(\mathbf{x})\right\rangle_j \mathbf{e}_j < 0$, and the NT with $\mathbf{e}_j = \mathbf{l}$ is the one which presents the higher decrease of $||\mathbf{g}(\mathbf{x})||$ after the manifold is transversed. This NT in this point coincides with a Gradient Extremal (GE).[13]

The proof is trivial, since $\mathbf{e}_j = \mathbf{l}$ and $\mathbf{l}$ is the normalized gradient. This implies that the gradient is an eigenvector of the $\mathbf{A}(\mathbf{x})$ matrix and in turn an eigenvector of the corresponding Hessian matrix, being this the requirement that a point belongs to a GE. The fact shows the importance of this manifold in the mechanochemistry theory based on Eq. 1. At the points belonging to the manifold $Det(\mathbf{H}(\mathbf{x})) = 0$ the bond breaking point (BBP) emerges.[7,13] At the latter type of points, the gradient norm along an NT reaches a turning point, and the effective PES, $V_{\mathbf{f}}(\mathbf{x})$, presents a shoulder on the FDSP path. Within the manifold of BBPs of a given PES, there is an *optimal* BBP and also the NT that goes through this point.[13] This BBP defines the lowest force in magnitude and the corresponding pulling direction that should be applied to a molecular system in order to mechanically promote a given chemical transformation by making it barrierless. The *optimal* BBP satisfies the equation[13]

$$\mathbf{H}(\mathbf{x})\mathbf{g}(\mathbf{x}) = \mathbf{0} \tag{7}$$

where $\mathbf{g}(\mathbf{x}) \neq \mathbf{0}$. Eq. 7 tell us that the *optimal* BBP is also a point of a GE where the eigenvalue of the Hessian matrix is zero. It is the point where Eq. 6 takes the value for $\mathbf{e}_j = \mathbf{l}$.

The purpose of the present study is to present another efficient algorithm to locate the point where Eq. 7 is satisfied. The present study is organized as follows: in Section 2 a new version of the Gauss-Newton method is presented. In Subsection 2.1 the mathematical ground of the algorithm is reported. In Subsection 2.2 it is detailed how to improve the Gauss-Newton method through the restricted step technique. The Subsections 2.3 and 2.4 expose in some detail the basis of the derivation of the Barnes update for a non-symmetric Jacobian as it occurs in the present algorithm. The behavior and performance of the algorithm is shown using some mechanochemical examples in Section 3. Finally, Section 4 concludes the paper.

2 An Algorithm based on the Restricted Gauss-Newton Method

2.1 Basic Formulas and Evaluation of the Displacement Vector

The problem of locating *optimal* BBPs is equivalent to finding the point of coordinates such that the eigenvalue problem, $\mathbf{H}(\mathbf{x})\mathbf{g}(\mathbf{x})\left\|\mathbf{g}(\mathbf{x})\right\|^{-1} = (\mathbf{g}(\mathbf{x})^T\mathbf{H}(\mathbf{x})\mathbf{g}(\mathbf{x})\left\|\mathbf{g}(\mathbf{x})\right\|^{-2})\mathbf{g}(\mathbf{x})\left\|\mathbf{g}(\mathbf{x})\right\|^{-1}$, has in this point the eigenvalue zero, $\mathbf{g}(\mathbf{x})^T\mathbf{H}(\mathbf{x})\mathbf{g}(\mathbf{x}) = 0$. For this purpose if we define the function[19]

$$\sigma(\mathbf{x}) = \frac{\mathbf{g}^T(\mathbf{x})\mathbf{H}^2(\mathbf{x})\mathbf{g}(\mathbf{x})}{\mathbf{g}^T(\mathbf{x})\mathbf{g}(\mathbf{x})} = \mathbf{s}^T(\mathbf{x})\mathbf{s}(\mathbf{x}) \ , \tag{8}$$

then to find the *optimal* BBPs merely consists in locating the point such that the above function takes the minimum value, which is zero, since it is a sum of nonlinear least square functions. In other words, the $\sigma(\mathbf{x})$-function is by definition the scalar product of the vector, $\mathbf{s}(\mathbf{x}) = \mathbf{H}(\mathbf{x})\mathbf{g}(\mathbf{x})\left\|\mathbf{g}(\mathbf{x})\right\|^{-1}$, by itself and reaches its minimum value, the zero value, at the point where $\mathbf{s}(\mathbf{x}) = \mathbf{H}(\mathbf{x})\mathbf{g}(\mathbf{x})\left\|\mathbf{g}(\mathbf{x})\right\|^{-1} = \mathbf{0}$ which is just the condition of an *optimal* BBP given in Eq. 7. We recall that $\mathbf{g}(\mathbf{x})\left\|\mathbf{g}(\mathbf{x})\right\|^{-1} \neq \mathbf{0}$ at the slope of the PES, thus the $\sigma(\mathbf{x})$-function is never zero outside a stationary point except if at this point the corresponding Hessian matrix has at least one zero eigenvalue. Thus the original problem has been transformed into finding the point where $\sigma(\mathbf{x}) = 0$. Since the $\sigma(\mathbf{x})$-function is continuous and differentiable with respect to $\mathbf{x}$ we can expand it around $\mathbf{x}$ up to second order in a Taylor expansion

$$\sigma(\mathbf{x}') = \sigma(\mathbf{x}) + (\mathbf{x}' - \mathbf{x})^T\nabla_{\mathbf{x}}\sigma(\mathbf{x})|_{\mathbf{x}} + \frac{1}{2}(\mathbf{x}' - \mathbf{x})^T\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}^T\sigma(\mathbf{x})|_{\mathbf{x}}(\mathbf{x}' - \mathbf{x}) + \mathcal{O}(||\mathbf{x}' - \mathbf{x}||^3) \ . \tag{9}$$

According to Eq. 8 the derivatives that appear in Eq. 9 with respect to $\mathbf{x}$ have the form

$$\nabla_{\mathbf{x}}\sigma(\mathbf{x}) = 2\ [\nabla_{\mathbf{x}}\mathbf{s}^T(\mathbf{x})\ ]\mathbf{s}(\mathbf{x}) = 2\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$$

and

$$\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}^T\sigma(\mathbf{x}) = 2\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) + 2\sum_{i=1}^{N} s_i(\mathbf{x})\ [\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}^T s_i(\mathbf{x})\ ]$$

where $s_i(\mathbf{x})$ is the i-th component of the $\mathbf{s}(\mathbf{x})$ vector and $\mathbf{J}(\mathbf{x}) = \ [\nabla_{\mathbf{x}}s_1(\mathbf{x})|\cdots|\nabla_{\mathbf{x}}s_N(\mathbf{x})\ ]$. The $\mathbf{J}(\mathbf{x})$ is the $N \times N$ Jacobian matrix. Since $\mathbf{s}(\mathbf{x})$ approaches the zero vector at the desired point where the $\sigma(\mathbf{x})$-function is close to zero, then normally the components $s_i(\mathbf{x})$ for $i = 1, \cdots, N$ are small. This suggests that a good approximation to $\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}^T\sigma(\mathbf{x})$ might be obtained by neglecting the final terms of this matrix to result in $\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}^T\sigma(\mathbf{x}) \approx 2\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$. This is equivalent to making a linear approximation to each component of the $\mathbf{s}(\mathbf{x})$-vector. In this manner using the two elements necessary to compute the gradient, $\nabla_{\mathbf{x}}\sigma(\mathbf{x})$, namely, $\mathbf{s}(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$, are also enough to evaluate the approximate Hessian matrix of this function. The expression for $\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}^T\sigma(\mathbf{x})\ \mathbf{J}(\mathbf{x})$ has been reported in Ref. 19. When the second derivative is approximated, the basic Newton method becomes the Gauss-Newton method or the generalized least squares method.[20] Note that the Gauss-Newton method can fail or can converge slowly. The convergence of the Gauss-Newton algorithm is largely improved if at each iteration the restricted step technique is used.

In the previous algorithm to locate *optimal* BBPs the restricted step was evaluated employing the rational function optimization technique.[21–24] However, the main problem of the rational function optimization technique used to evaluate the restricted step is that the computed displacement vector, $\Delta\mathbf{x} = \mathbf{x}' - \mathbf{x}$, violates the imposed restriction on its length. The length of the displacement vector is sometimes larger or shorter with respect to the required length. For this reason the displacement vector should be scaled. This scaled vector is not the optimal one since the restricted step technique consists in finding the stationary point of minimal energy in a subspace which is generated by the intersection of the PES expanded

up to second order in $\Delta\mathbf{x}$ and a sphere of a given radius centered in the point $\mathbf{x}$. This radius is known as trust radius.[20]

The problem has been treated several times, and now we propose to solve it by diagonalization of a non-symmetric matrix. The restricted step technique in the present case can be formulated as follows: given the $\sigma(\mathbf{x}')$-function centered at the point $\mathbf{x}$ and expanded until second order in $\Delta\mathbf{x}$, find the stationary points located in the intersection of this function with a sphere centered at the point $\mathbf{x}$ of radius r. This problem is solved through the Lagrange multipliers method, that now reads

$$\begin{aligned} &L(\Delta\mathbf{x},\lambda) = \Delta\sigma^{(2)}(\Delta\mathbf{x}) - \lambda(\Delta\mathbf{x}^T\Delta\mathbf{x} - r^2) = \\ &\Delta\mathbf{x}^T\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})\Delta\mathbf{x} + 2\Delta\mathbf{x}^T\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) - \lambda(\Delta\mathbf{x}^T\Delta\mathbf{x} - r^2) \end{aligned} \tag{10}$$

where $\Delta\sigma^{(2)}(\Delta\mathbf{x}) = \Delta\mathbf{x}^T\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})\Delta\mathbf{x} + 2\Delta\mathbf{x}^T\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) \approx \sigma(\mathbf{x}+\Delta\mathbf{x}) - \sigma(\mathbf{x}) = \sigma(\mathbf{x}') - \sigma(\mathbf{x})$. Differentiating $L(\Delta\mathbf{x},\lambda)$ with respect to $\Delta\mathbf{x}$ and λ yields the equations

$$\begin{cases} a) & \mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})\Delta\mathbf{x} + \mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) - \lambda\Delta\mathbf{x} = \mathbf{0} \\ b) & \Delta\mathbf{x}^T\Delta\mathbf{x} - r^2 = 0\ . \end{cases} \tag{11}$$

Note that λ is a function of $\mathbf{x}$, however, this dependence will be omitted to make the notation short. It can be shown that from the set of tuples, $\{(\lambda_k, \Delta\mathbf{x}_k)\}_{k=min}^{max}$, solutions of Eqs. 11, if $\lambda_{max} > \lambda_i > \lambda_j > \lambda_{min}$ then $\Delta\sigma^{(2)}(\Delta\mathbf{x}_{max}) > \Delta\sigma^{(2)}(\Delta\mathbf{x}_i) > \Delta\sigma^{(2)}(\Delta\mathbf{x}_j) > \Delta\sigma^{(2)}(\Delta\mathbf{x}_{min})$. The proof of these inequalities is given in Appendix A. The smallest λ, labeled as λ_{min}, is needed in order to minimize the value of the quadratic approximation of the $\sigma(\mathbf{x})$ function around the point $\mathbf{x}$. From Eqs. 10 and 11 we have that

$$\Delta\sigma^{(2)}(\Delta\mathbf{x}_i) = \lambda_i r^2 + \Delta\mathbf{x}_i^T\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})\ . \tag{12}$$

So in place of the original minimization we can solve the Lagrange Eqs. 11 for λ_{min}. The Lagrange equations 11 can be reduced to a quadratic eigenvalue problem.[25,26] For the deriva-

tion let us assume that λ does not belong to the set of eigenvalues of the matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$, thus from Eq. 11.$a$

$$\Delta\mathbf{x} = -[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^{-1}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) \ . \tag{13}$$

Taking into account the normalization condition for $\Delta\mathbf{x}$, Eq. 11.b we get the secular function

$$f(\lambda) = \mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) - r^2 \tag{14}$$

of which the zeros are to be computed. Now, if we define $\mathbf{t}(\mathbf{x}) = [\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$ so that $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^2\mathbf{t}(\mathbf{x}) = \mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$ then instead of solving the secular function, Eq. 14, we have to solve the system

$$\begin{cases} a) & [\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^2\mathbf{t}(\mathbf{x}) &=& \mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) \\ b) & \mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}(\mathbf{x}) - r^2 &=& 0 \ . \end{cases} \tag{15}$$

Eq. 15.b can be formulated as $r^{-2}\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}(\mathbf{x}) = 1$. Using the factor 1 in Eq. 15.a, we get the quadratic eigenvalue problem,

$$\{[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^2 - r^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\}\mathbf{t}(\mathbf{x}) = \mathbf{0} \tag{16}$$

Note that the restriction that λ does not belong to the set of eigenvalues of matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ is no longer necessary. Of course, we must face the fact that the set of solutions for λ has been extended by these manipulations, because two equations cannot be formulated as a single one without consequences. These consequences are exposed in some detail in Appendix B and are taken into account in the proposed algorithm.

The quadratic eigenvalue problem, Eq. 16, can be reduced to an ordinary eigenvalue problem by properly chosen transformations. With the definition

$\mathbf{p}(\mathbf{x}) = [\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]\mathbf{t}(\mathbf{x})$, the following equations can be established

$$\begin{cases} a) & \mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}(\mathbf{x}) - \mathbf{p}(\mathbf{x}) &= \lambda\mathbf{t}(\mathbf{x}) \\ b) & \mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{p}(\mathbf{x}) - r^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}(\mathbf{x}) &= \lambda\mathbf{p}(\mathbf{x}) \ . \end{cases} \tag{17}$$

In matrix terms this leads to

$$\begin{pmatrix} \mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) & -\mathbf{I} \\ -r^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x}) & \mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) \end{pmatrix} \begin{pmatrix} \mathbf{t}(\mathbf{x}) \\ \mathbf{p}(\mathbf{x}) \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{t}(\mathbf{x}) \\ \mathbf{p}(\mathbf{x}) \end{pmatrix} \ . \tag{18}$$

Thus we have transformed the original quadratic eigenvalue problem into an equivalent linear one that can be solved with traditional techniques. Let us denote the real solutions of Eq. 18 by the triples, $\{(\lambda_i, \mathbf{t}_i^T(\mathbf{x}), \mathbf{p}_i^T(\mathbf{x}))\}_{i=1,..,N}$, where λ_i are given in increasing order. This set of triples has the full set of solutions Eq. 14 and $\lambda_1 = \lambda_{min}$, is the desired solution of Eq. 11. Now, substituting the real triple, $(\lambda_{min}, \mathbf{t}_{min}^T(\mathbf{x}), \mathbf{p}_{min}^T(\mathbf{x}))$, in Eq. 17.$b$ we see that,

$$\Delta\mathbf{x}_{min} = -[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda_{min}\mathbf{I}]^{-1}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) = -\mathbf{p}_{min}(\mathbf{x})[\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}_{min}(\mathbf{x})]^{-1}r^2 \ . \tag{19}$$

If we multiply this equation from the left by $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda_{min}\mathbf{I}]^{-1}$ and using Eq. 17.a we obtain, $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda_{min}\mathbf{I}]^{-1}\Delta\mathbf{x}_{min} = -\mathbf{t}_{min}(\mathbf{x})[\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}_{min}(\mathbf{x})]^{-1}r^2$. Now, multiplying the latest equation from the left by $\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ and taking into account Eq. 13 we have, $\Delta\mathbf{x}_{min}^T\Delta\mathbf{x}_{min} = r^2$, as expected. Note that this result can be applied to any real triple and it is Eq. 19, that transforms the vectors $\mathbf{t}_i^T(\mathbf{x})$ and $\mathbf{p}_i^T(\mathbf{x})$ of the i-th triple to the $\Delta\mathbf{x}_i$ vector that belongs to the $(\lambda_i, \Delta\mathbf{x}_i)$-tuple solution of Eq. 11, that is particularly important for the algorithm. Now, we treat the case that λ_i belongs to an eigenvalue of the the matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$. Let us assume that $\mathbf{t}_i(\mathbf{x})$ is the eigenvector; then, from Eq. 17.a we obtain that $\mathbf{p}_i(\mathbf{x}) = \mathbf{0}$. Now from Eq. 17.$b$ we see that $\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}_i(\mathbf{x}) = 0$ and the triple has the form $(\lambda_i, \mathbf{t}_i^T(\mathbf{x}), \mathbf{0}^T)$. The application of Eq. 19 to this triple gives an undetermined $\Delta\mathbf{x}$ vector. For this reason the triples formed by λ's that belong to the eigenvalue spectra of $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$

matrix are not solutions of Eq. 11, see also Appendix B. Now, we study the origin of the complex solutions. Let λ_i and λ_{i+1} be two consecutive eigenvalues of the matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ such that $\lambda_i^J < \lambda_{i+1}^J$ but they are not solutions of Eq. 18. Now if in the interval $(\lambda_i^J, \lambda_{i+1}^J)$ the secular function, $f(\lambda)$ given in Eq. 14, is greater than zero then there exists a complex tuple and the corresponding conjugate, which are both solutions of Eq. 18. The non-symmetric eigenvalue problem of Eq. 18 always has at least two real tuples, and their λ's correspond to the lowest, λ_{min}, and the highest, λ_{max} values.

The proof of this assertion is the following. As shown above any real tuple solution of Eq. 18 is solution of Eq.14 excepting if, let us say $\lambda_i = \lambda_i^J$, an eigenvalue of $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ matrix, is solution of Eq. 14. Now let us take a r^2 value, then assuming that $\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) \neq \mathbf{0}$ the $f(\lambda)$ increases monotonically from $-r^2$ to ∞ as λ increases from $-\infty$ to λ_1^J, the lowest eigenvalue of $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ matrix. Thus we have to find a unique value of λ in this interval, $\lambda_{min} < \lambda_1^J$, for which $f(\lambda_{min}) = 0$. In a similar manner the $f(\lambda)$ decreases monotonically from ∞ to $-r^2$ as λ increases from λ_N^J, the highest eigenvalue of $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$, to ∞. Thus we have to find a unique value of λ in this interval, $\lambda_{max} > \lambda_N^J$, for which $f(\lambda_{max}) = 0$. This concludes the proof.

Finally, if the solution of Eq. 18 has $2N$ real tuples then from the above results it is easy to proof the corollary about the existence of the next interlace, $\lambda_{min} < \lambda_1^J < \lambda_2 < \lambda_3 < \lambda_2^J < \lambda_4 < \cdots < \lambda_{2N-3} < \lambda_{N-1}^J < \lambda_{2N-2} < \lambda_{2N-1} < \lambda_N^J < \lambda_{max}$, between the $2N$ real tuples and the eigenvalues of the $\mathbf{J}(\mathbf{x})^T\mathbf{J}(\mathbf{x})$, where $min = \lambda_1$ and $max = \lambda_{2N}$.

2.2 Restricted Step Characterization and Trust Radius Update

Many optimization algorithms are based on truncating the Taylor series of the objective function to be minimized, but this truncated function does not have a minimizing point. In other words, $\sigma(\mathbf{x} + \Delta\mathbf{x}) \approx \sigma(\mathbf{x}) + \Delta\sigma^{(2)}(\Delta\mathbf{x})$ and regarding the region about the central point $\mathbf{x}$ in which the quadratic expansion is adequate, the $\sigma(\mathbf{x}) + \Delta\sigma^{(2)}(\Delta\mathbf{x})$ function does not have a minimum point. A way to solve this problem is to assume a region $R_{\mathbf{x}}$ centered

at the point $\mathbf{x}$ in which the function $\sigma(\mathbf{x}) + \Delta\sigma^{(2)}(\Delta\mathbf{x})$ agrees with $\sigma(\mathbf{x}+\Delta\mathbf{x})$ in some sense. Then one chooses $\mathbf{x}' = \mathbf{x} + \Delta\mathbf{x}$ such that $\Delta\mathbf{x}$ minimizes $\Delta\sigma^{(2)}(\Delta\mathbf{x})$ for all $\mathbf{x}' = \mathbf{x} + \Delta\mathbf{x}$ in $R_{\mathbf{x}}$-region. This is the basis of the restricted step method.[20] The trust radius is the parameter that characterizes the $R_{\mathbf{x}}$-region centered at the point $\mathbf{x}$.

The region $R_{\mathbf{x}}$ is normally limited by a sphere, $R_{\mathbf{x}} = \{\mathbf{x}' : (\mathbf{x}' - \mathbf{x})^T(\mathbf{x}' - \mathbf{x}) = \Delta\mathbf{x}^T\Delta\mathbf{x} \leq r^2\}$, and to seek for the solution $\Delta\mathbf{x}$ of the problem

$$\min_{\Delta\mathbf{x}} \Delta\sigma^{(2)}(\Delta\mathbf{x}) \quad \text{subject to} \quad \Delta\mathbf{x}^T\Delta\mathbf{x} \leq r^2 \tag{20}$$

where the equality is nothing more that Eq. 10 solved through Eqs. 18 and 19, and by standard Gauss-Newton for the inequality if the resulting $\Delta\mathbf{x}$ satisfies that $\Delta\mathbf{x}^T\Delta\mathbf{x} < r^2$. Now emerges the question of how the radius r of sphere $R_{\mathbf{x}}$ shall be chosen. The radius r should take into account certain measure of agreement at each step between $\Delta\sigma^{(2)}(\Delta\mathbf{x})$ and $\sigma(\mathbf{x} + \Delta\mathbf{x}) - \sigma(\mathbf{x})$. The ratio

$$ratio = \frac{\sigma(\mathbf{x} + \Delta\mathbf{x}) - \sigma(\mathbf{x})}{\Delta\sigma^{(2)}(\Delta\mathbf{x})} \tag{21}$$

gives a measure of accuracy for which $\Delta\sigma^{(2)}(\Delta\mathbf{x})$ approximates $\sigma(\mathbf{x} + \Delta\mathbf{x}) - \sigma(\mathbf{x})$. Thus when *ratio* is the unity or near to it, it indicates the best accuracy. Following Fletcher[20] an algorithm can be stated that changes r adaptively and attempts to maintain a certain degree of agreement measured by *ratio*, whilst keeping r as large as possible. The algorithm used in the present context at each iteration has the following form:

1) given $\mathbf{x}$ and r, compute $\mathbf{J}(\mathbf{x})$ and $\mathbf{s}(\mathbf{x})$;

2) solve Eq. 20 first by Gauss-Newton, $\Delta\mathbf{x} = -(\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}))^{-1}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$,

and if $\Delta^T\mathbf{x}\Delta\mathbf{x} > r^2$ then solve Eq. 18 and through Eq. 19 obtain the adequate $\Delta\mathbf{x}$;

3) compute $\sigma(\mathbf{x} + \Delta\mathbf{x})$ and hence *ratio* from Eq. 21 using Eq. 12;

4) if *ratio* < 0.25 the new r takes the value $||\Delta\mathbf{x}||/4$,

if *ratio* > 0.75 and $||\Delta\mathbf{x}|| = r$ then the new r takes the value $2r$,

otherwise the value of the new r is the same of the current step;

5) if $r \leq 0$ set $\mathbf{x}' = \mathbf{x}$ else $\mathbf{x}' = \mathbf{x} + \Delta\mathbf{x}$.

One can prove that the constant numbers used, namely, 0.0, 0.25, 0.75, 2.0 are arbitrary and the algorithm is quite insensitive to their change.[20]

2.3 The Barnes Update Formula for the Jacobian Matrix

The main disadvantage of the present algorithm based on the Gauss-Newton method is the calculation of the vector $\mathbf{s}(\mathbf{x})$ and the matrix $\mathbf{J}(\mathbf{x})$. As explained in the subsection 2.1, the vector and the matrix are essential to compute the required gradient and Hessian of the $\sigma(\mathbf{x})$ function at each step or iteration of the algorithm. However, as proved in Ref. 19 its computation involves first, second and third derivatives of the PES with respect to the coordinates. Clearly they are very expensive from a computational point of view, in particular the matrix $\mathbf{J}(\mathbf{x})$ where third derivatives are needed. Thus a way to update the matrix $\mathbf{J}(\mathbf{x})$ at each step is necessary to make the algorithm efficient. The matrix $\mathbf{J}(\mathbf{x})$ is not symmetric thus standard procedures of updating, like the popular Broyden-Fletcher-Goldfarb-Shanno (BFGS)[20] proposed for symmetric cases can not be used in the present case. For this reason in Ref. 19 the Broyden formula[27] was employed to update the $\mathbf{J}(\mathbf{x})$ matrix at each iteration. As will be shown below, the Broyden formula is a simplification of the Barnes formula[28] and the latter has some properties that the former does not have and these differences can be important to improve the algorithm.

In this and the next subsection we use a $\mathbf{d}$-vector rather than $\Delta\mathbf{x}$ to make the notation short. Let us define the pair of differences $\mathbf{y}^{(k)} = \mathbf{s}^{(k+1)} - \mathbf{s}^{(k)}$ and $\mathbf{d}^{(k)} = \Delta\mathbf{x}^k = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$, for $k = 1, \ldots, n$, where $\mathbf{s}^{(k)} = \mathbf{s}(\mathbf{x}^{(k)}) = \mathbf{H}(\mathbf{x}^{(k)})\mathbf{g}(\mathbf{x}^{(k)})||\mathbf{g}(\mathbf{x}^{(k)})||^{-1}$ and n indicates the iteration-n. Then a matrix $\mathbf{J}^{(k+1)}$ can be calculated, where $\mathbf{J}^{(n+1)} = \mathbf{J}(\mathbf{x}^{(n+1)})$, which satisfies

the hereditary condition

$$\mathbf{J}^{(n+1)T}\mathbf{d}^{(k)} = \mathbf{y}^{(k)}, \quad k = 1, \ldots, n. \tag{22}$$

Since $\mathbf{J}^{(k)}$ does not relate $\mathbf{d}^{(k)}$ and $\mathbf{y}^{(k)}$ correctly, thus $\mathbf{J}^{(k+1)}$ is chosen so that it does correctly relate this pair of differences, $\mathbf{J}^{(k+1)T}\mathbf{d}^{(k)} = \mathbf{y}^{(k)}$ being the quasi-Newton condition. One possibility is to have

$$\mathbf{J}^{(k+1)} = \mathbf{J}^{(k)} + \mathbf{C}^{(k)} \ . \tag{23}$$

$\mathbf{C}^{(k)}$ is a matrix, which to avoid carrying over too much information from iteration to iteration, is calculated from the quantities $\mathbf{d}^{(k)}$, $\mathbf{y}^{(k)}$ and $\mathbf{J}^{(k)}$. One can see the updating technique as a perturbation to the current $\mathbf{J}^{(k)}$ matrix by $\mathbf{C}^{(k)}$ to satisfy both the current quasi-Newton condition and the hereditary condition, Eq. 22. Taking into account this premise, to obtain an explicit form for $\mathbf{C}^{(k)}$ we substitute Eq. 23 into the quasi-Newton condition giving

$$\mathbf{C}^{(k)T}\mathbf{d}^{(k)} = \mathbf{J}^{(k+1)T}\mathbf{d}^{(k)} - \mathbf{J}^{(k)T}\mathbf{d}^{(k)} = \mathbf{y}^{(k)} - \mathbf{J}^{(k)T}\mathbf{d}^{(k)} \ . \tag{24}$$

One simple solution for this expression is readily seen to be

$$\mathbf{C}^{(k)T} = \frac{\mathbf{y}^{(k)}\mathbf{v}^{(k)T}}{\mathbf{v}^{(k)T}\mathbf{d}^{(k)}} - \frac{\mathbf{J}^{(k)T}\mathbf{d}^{(k)}\mathbf{w}^{(k)T}}{\mathbf{w}^{(k)T}\mathbf{d}^{(k)}} \tag{25}$$

where $\mathbf{v}^{(k)}$ and $\mathbf{w}^{(k)}$ are two non-zero arbitrary vectors except that neither should be orthogonal to $\mathbf{d}^{(k)}$. The numerators in Eq. 25 are matrices with a rank one. Thus $\mathbf{C}^{(k)}$ has a rank one or two depending upon the choice of $\mathbf{v}^{(k)}$ and $\mathbf{w}^{(k)}$. However, there is evidently still a considerable amount of flexibility in the choice of $\mathbf{C}^{(k)}$. Using the hereditary condition, Eq. 22, we have

$$\mathbf{C}^{(k)T}\mathbf{d}^{(j)} = \mathbf{J}^{(k+1)T}\mathbf{d}^{(j)} - \mathbf{J}^{(k)T}\mathbf{d}^{(j)} = \mathbf{0} \ , \quad j = 1, \ldots, k-1 \ . \tag{26}$$

With $\mathbf{C}^{(k)}$ given by Eq. 25 it therefore follows that

$$\mathbf{v}^{(k)T}\mathbf{d}^{(k)} = \mathbf{w}^{(k)T}\mathbf{d}^{(k)} = 0 \ , \quad j = 1, \ldots, k-1 \ . \tag{27}$$

Thus choose $\mathbf{J}^{(k+1)}$ so as not to change the information in $\mathbf{J}^{(k)}$ relevant to these directions, $j = 1, ..., k-1$. The choice $\mathbf{v}^{(k)} = \mathbf{w}^{(k)}$, a vector which is orthogonal to $\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, \ldots, \mathbf{d}^{(k-1)}$, ensures that the updated matrix satisfies both the quasi-Newton condition and a hereditary property. Thus the resulting update formula is

$$\mathbf{J}^{(k+1)} = \mathbf{J}^{(k)} + \frac{\mathbf{w}^{(k)}(\mathbf{y}^{(k)} - \mathbf{J}^{(k)T}\mathbf{d}^{(k)})^T}{\mathbf{w}^{(k)T}\mathbf{d}^{(k)}} \ . \tag{28}$$

This formula was proposed for the first time by Barnes.[28] Note that the vector $\mathbf{w}^{(k)}$ is the component of $\mathbf{d}^{(k)}$ orthogonal to $\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, \ldots, \mathbf{d}^{(k-1)}$. We recall that N is the dimension of the problem, thus in the $N+1$ iteration the resulting difference pair $\mathbf{y}^{(N+1)}$, $\mathbf{d}^{(N+1)}$ yields too much information to determine $\mathbf{J}^{N+2}$. However, using Eq. 28 and making $\mathbf{w}^{(N+1)}$ orthogonal to $\mathbf{d}^{(2)}, \mathbf{d}^{(3)}, \ldots, \mathbf{d}^{(N)}$ ensures that $\mathbf{J}^{(N+2)T}\mathbf{d}^{(j)} = \mathbf{y}^{(j)}$ for $j = 2, \ldots, N+1$. The difference pair $\mathbf{y}^{(1)}$, $\mathbf{d}^{(1)}$ is dropped out and replaced by $\mathbf{y}^{(N+1)}$, $\mathbf{d}^{(N+1)}$. In this manner for any $k > N$ an approximate matrix $\mathbf{J}^{(k+1)}$ can be calculated which satisfies $\mathbf{J}^{(k+1)T}\mathbf{d}^{(j)} = \mathbf{y}^{(j)}$ for the most recent N difference pairs.

The Broyden update[28] is derived from the Barnes formula, Eq. 28, by taking $\mathbf{w}^{(k)} = \mathbf{d}^{(k)}$. This election clearly destroys the hereditary property, Eq. 22, and from this termination in step $N + 1$ for the case that the vector $\mathbf{s}(\mathbf{x})$ is linear. Also some invariant properties preserved in Barnes update formula are lost in the Broyden formula. This makes the use of Barnes formula more attractive for the present type of problem. The price to be paid is the orthonormalization of the space generated by the displacement vectors at each step to compute the $\mathbf{w}^{(k)}$ vector.

2.4 Orthonormalization of the Displacement Vectors

Eq. 28 represents a modification of the matrix $\mathbf{J}^{(k)}$ by a rank one update. A useful representation that helps to understand this update is based on the definition of the matrices $\mathbf{Y}^{(k)} = [\mathbf{y}^{(1)}|\ldots|\mathbf{y}^{(k)}]$ and $\mathbf{D}^{(k)} = [\mathbf{d}^{(1)}|\ldots|\mathbf{d}^{(k)}]$ whose columns contain the most recent N (or k if $k \leq N$) difference pairs $\mathbf{y}^{(k)} = \mathbf{s}^{(k+1)} - \mathbf{s}^{(k)}$ and $\mathbf{d}^{(k)} = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$. If we assume that these two matrices have full rank (min (k, N)) then the representation $\mathbf{J}^{(k+1)T} = \mathbf{Y}^{(k)}\mathbf{D}^{+(k)}$ satisfies Eq. 22 and is equivalent to the use of Eq. 28 when the vector $\mathbf{w}^{(k)}$ is the component of $\mathbf{d}^{(k)}$ orthogonal to the set $\{\mathbf{d}^{(j)}\}_{j=1}^{k-1}$ where $k \leq N$. Thus we only need a way to compute at each step the corresponding $\mathbf{w}^{(k)}$ vector orthogonal to the current iterative subspace of displacement vectors needed in the Barnes update formula, Eq. 28. For this reason we describe a stable algorithm for this purpose. The matrix $\mathbf{D}^{+(k)}$ denotes the full rank generalized inverse of $\mathbf{D}^{(k)}$ having the form $\mathbf{D}^{+(k)} = (\mathbf{D}^{(k)T}\mathbf{D}^{(k)})^{-1}\mathbf{D}^{(k)T}$. The dimension of the matrix $\mathbf{D}^{+(k)}$ is $k \times N$. Thus using Eq. 28 one needs only a way to update the $\mathbf{D}^{+(k-1)}$ matrix to $\mathbf{D}^{+(k)}$, and in this computation the $\mathbf{w}^{(k)}$ vector is obtained. This is done using a procedure proposed by Fletcher[29] that is now briefly described for the purpose of being applied in the present context. Another procedure is due to Ben-Israel.[30] It is also an iterative procedure but we will not discuss it here.

We define a projection matrix $\mathbf{P}^{(k-1)}$ as $\mathbf{D}^{(k-1)}\mathbf{D}^{+(k-1)}$ where $\mathbf{D}^{+(k-1)}\mathbf{D}^{(k-1)} = \mathbf{I}_{(k-1)\times(k-1)}$ being $\mathbf{I}_{(k-1)\times(k-1)}$ the unit matrix of dimension $(k-1)\times(k-1)$ since $\mathbf{D}^{+(k-1)}$ is the "generalized inverse" of $\mathbf{D}^{(k-1)}$ as noted previously. Now, $\mathbf{D}^{(k-1)}$ is a matrix of dimension $N \times (k-1)$ whereas $\mathbf{D}^{+(k-1)}$ has dimension $(k-1) \times N$. Normally the orthogonalization of a vector, say $\mathbf{v}$, with respect to the set of vectors, $\{\mathbf{d}^{(j)}\}_{j=1}^{k-1}$, can be achieved by computing $\mathbf{Q}^{(k-1)}\mathbf{v} = \mathbf{v} - \mathbf{P}^{(k-1)}\mathbf{v} = \mathbf{v} - \mathbf{D}^{(k-1)}\mathbf{D}^{+(k-1)}\mathbf{v}$. Following Fletcher,[29] we give an expression showing how to update the "generalized inverse" $\mathbf{D}^{+(k-1)}$ when a column is added to $\mathbf{D}^{(k-1)}$ matrix. In other words, we are going from $\mathbf{D}^{+(k-1)}$ to $\mathbf{D}^{+(k)}$ when $\mathbf{D}^{(k-1)}$ extends to $\mathbf{D}^{(k)}$. We denote by $\mathbf{e}_i$ the i-th column vector of a unit matrix. Thus, the j-th element is δ_{ji} (Kro-

necker's delta). The relation $\mathbf{D}^{+(k-1)}\mathbf{D}^{(k-1)} = \mathbf{I}_{(k-1)\times(k-1)}$ and from this $\mathbf{D}^{+(k-1)}\mathbf{d}^{(i)} = \mathbf{e}_i$ for $i = 1, \ldots, k-1$, will be used. Let $\mathbf{D}^{(k-1)}$ and $\mathbf{D}^{+(k-1)}$ be the current matrices and $\mathbf{d}^{(k)}$ the new vector that extends $\mathbf{D}^{(k-1)}$ to $\mathbf{D}^{(k)} = [\mathbf{D}^{(k-1)}|\mathbf{d}^{(k)}]$. Then, $\mathbf{D}^{+(k)}$ is uniquely determined by the next conditions:

1) $\mathbf{D}^{+(k)}\mathbf{d}^{(i)} = \mathbf{D}^{+(k-1)}\mathbf{d}^{(i)} = \mathbf{e}_i$ for $i = 1, \ldots, k-1$;

2) $\mathbf{D}^{+(k)}\mathbf{d}^{(k)} = \mathbf{e}_k$;

3) $\mathbf{D}^{+(k)}\mathbf{z} = \mathbf{0}$ for any vector $\mathbf{z}$ orthogonal to $\{\mathbf{d}^{(i)}\}_{i=1}^{k}$.

An expression that satisfies these three conditions is

$$\mathbf{D}^{+(k)} = \begin{pmatrix} \mathbf{D}^{+(k-1)} \\ \mathbf{0}^T \end{pmatrix} + \begin{pmatrix} -\mathbf{D}^{+(k-1)}\mathbf{d}^{(k)} \\ 1 \end{pmatrix} \frac{\mathbf{w}^{(k)T}}{\mathbf{w}^{(k)T}\mathbf{d}^{(k)}} \tag{29}$$

where $\mathbf{w}^{(k)}$ is the component of $\mathbf{d}^{(k)}$ orthogonal to the columns of $\mathbf{D}^{(k-1)}$ matrix, $\{\mathbf{d}^{(i)}\}_{i=1}^{k-1}$, which is written as $\mathbf{w}^{(k)} = \mathbf{d}^{(k)} - \mathbf{D}^{(k-1)}\mathbf{D}^{+(k-1)}\mathbf{d}^{(k)} = \mathbf{Q}^{(k-1)}\mathbf{d}^{(k)}$ and $\mathbf{0}$ is the zero vector of dimension N. This $\mathbf{w}^{(k)}$ vector is the one that appears in the Barnes update formula, Eq. 28. In fact the $\mathbf{w}^{(k)T}/(\mathbf{w}^{(k)T}\mathbf{d}^{(k)})$ vector of Eq. 28 is the k-th row of the $\mathbf{D}^{+(k)}$ matrix, the added row that has transformed the $\mathbf{D}^{+(k-1)}$ matrix to the $\mathbf{D}^{+(k)}$ matrix. This row is also labeled as $\mathbf{d}^{+(k)T}$. The $||\mathbf{w}^{(k)}||^2$ is a measure of the linearly dependent character of the set of vectors $\{\mathbf{d}^{(i)}\}_{i=1}^{k}$. If it is close to zero, it implies that this set of vectors is nearly linearly dependent.

Now we consider the case that $\mathbf{D}^{+(k)}$, $\mathbf{D}^{(k)}$, and $\mathbf{d}^{(k)}$ are given, and that $\mathbf{D}^{+(k-1)}$ and $\mathbf{D}^{(k-1)}$ are required. This consideration is important when $k > N$, since as explained previously the oldest vector of the set $\{\mathbf{d}^{(k)}\}_{k=1}^{N}$ should be substituted by the newest one. We see that $\mathbf{D}^{(k-1)} = \mathbf{D}^{(k)} - [\mathbf{O}^{(k-1)}|\mathbf{d}^{(k)}]$ where $\mathbf{O}^{(k-1)}$ is the zero matrix of dimension $N \times (k-1)$. Using Eq. 29 the k-th row of $\mathbf{D}^{+(k)}$ matrix, that is $\mathbf{d}^{+(k)T}$, is just $\mathbf{w}^{(k)T}/\mathbf{w}^{(k)T}\mathbf{d}^{(k)}$ as noted

above thus the last equation can be rearranged in the form

$$\begin{pmatrix} \mathbf{D}^{+(k-1)} \\ \mathbf{0}^T \end{pmatrix} = \mathbf{D}^{+(k)} - \mathbf{u}^{(k)}\mathbf{d}^{+(k)T} \ . \tag{30}$$

Now the rows of $\mathbf{D}^{+(k-1)}$ are linear combinations of the columns of $\mathbf{D}^{(k-1)}$, and $\mathbf{d}^{+(k)}$ is orthogonal to all of these. Hence multiplying this equation by $\mathbf{d}^{+(k)}$ on the right gives

$$\mathbf{u}^{(k)} = \frac{\mathbf{D}^{+(k)}\mathbf{d}^{+(k)}}{\mathbf{d}^{+(k)T}\mathbf{d}^{+(k)}} \tag{31}$$

so that $\mathbf{D}^{+(k-1)}$ is now determined completely and the formula for contracting the basis is established. These formulae enable the basis set $\{\mathbf{d}^{(j)}\}_{j=1}^{k}$ to be varied at will.

In summary, using Barnes update formula, Eq. 28, if the iteration, say k, is lower than N then we have to use Eq. 29 to compute the current $\mathbf{w}^{(k)}$ vector and the extended $\mathbf{D}^{+(k+1)}$ matrix. Also, the $\mathbf{D}^{(k)}$ matrix should be extended by a column to obtain $\mathbf{D}^{(k+1)}$. If $k > N$ then first we have to use Eq. 30 to eliminate the corresponding row of the current $\mathbf{D}^{+}$ matrix and the column of the $\mathbf{D}$ matrix, both of dimension $N \times N$, and second, we have to use Eq. 29 to compute the new row of the $\mathbf{D}^{+}$ matrix, the actual $\mathbf{w}^{(k)}$ vector and to replace the column of the $\mathbf{D}$ matrix. Finally, the algorithm has an important property detailed in Appendix C.

3 Examples and Performance of the Algorithm

In this section, we will present the results of the algorithm applied to a toy model (Frenkel-Kontorova) and to two different and relevant medium size molecular systems. These two molecular systems are the ring opening of a Benzocyclobutene derivative and the Chorismate Mutase rearrangement.

3.1 The Frenkel-Kontorova model

We use a Frenkel-Kontorova model as the first example.[31–35] It is a simple, yet non-trivial, model. One often divides in solid-state physics a set of particles into a one-dimensional subsystem of interacting elements, and the remaining part as a substrate. The latter acts by a potential on the extracted subsystem. In chemistry, the Frenkel-Kontorova chain is to a certain degree similar to a linear chain-like alkane elongation.[36]

N equal atoms or ions may be described by the vector of coordinates $\mathbf{x} = (x_1, .., ..., x_N)$. The positions x_i are on an axis. The coordinates of all particles satisfy $x_i < x_{i+1}$. They are sorted in a fixed order. We search for a movement of the full chain along the x-axis. The potential energy surface (PES) of the model is

$$V(\mathbf{x}) = v \sum_{i=1}^{N} [1 - cos(\frac{2\pi x_i}{a_s})] + \sum_{i=1}^{N-1} \frac{k}{2} \left[x_{i+1} - x_i - a_o\right]^2 . \tag{32}$$

We calculate the case $N = 10$. For simplicity we use here the ratio $a_o/a_s = 1$ thus we exclude a misfit of the chain and the substrate.[33] The parameter a_o is the average spacing of the chain, but a_s is the periodicity of the substrate potential. We usually put the factor at the sinusoidal potential, v=1. Then the spring constant, k, is also the ratio of the strength. of the sinusoidal potential to that of the spring potential. Here we also simplify $k = 1$ throughout. All quantities referred to are dimensionless.

Because we have no misfit, the ground state of the chain is the sitting of every atom in one well of the substrate. The lowest eigenvalue of the Hessian matrix of the PES is then 1, and the corresponding eigenvector of the first normal mode is purely translational.[31] This movement leaves the chain unchanged, it only vibrates 'collectively' in the bowls of the substrate, every atom in its own bowl.

However, on the PES such a 'collective' unified movement of the chain is not the direction to a low energy path through the PES mountains. In contrast, it leads to a SP of index four, where all atoms are raised to the tops of the substrate, compare Figs.14 and 15 of

ref. 35. This SP_4 has a very high energy. Consequently, a Newton trajectory (NT) with search direction $(1,1,...,1,1)^T$ does not lead to the known lowest energy path over consecutive SPs of index one, these SPs being kinks or anti-kinks of the chain.[33–35] Note that kinks are stretched structures, but anti-kinks are compressed structures of the chain.

Paying attention to this, we start our test with an NT along a pulling direction that does not involve all atoms: $(1,1,1,1,0,0,..,0)^T$ and obtain a corresponding BBP_{1111} on the NT profile. This guess is at (1.795, 7.942, 13.973, 19.879, 25.53, 31.568, 37.757, 44.005, 50.274, 56.553). We use it for the initial structure for the Barnes algorithm. It converges after 21 iterations to the optimal BBP, at energy 3.05,

BBP_{opt}=(2.04, 7.606, 13.26, 19.182, 25.285, 31.484, 37.729, 43.995, 50.271, 56.552).

The structures of the chain along the 21 steps of the optimization are represented in an SI-file by an animation: *F*Km10DanimateBarnesBBP.gif. One can additionally use the gradient vector with norm 1.88 at the optimal BBP for the optimal force normalized direction

$\mathbf{f}_{opt}$=(0.86, 0.47, 0.2, 0.08, 0.03, 0.01, 0.004, 0.002, 0.001, 0.0003) (*)

for the search direction of an 'optimal' NT. The energy profile of this NT is shown in Fig. 1. It performs a nice ascent to the first floor of the anti-kink SPs over the (red) optimal BBP. However, this NT does not demonstrate the correct minimum energy path. At the end, it deviates uphill and finds an SP of index 2 on the second floor of combined kink and anti-kink structures. For such a combined structure to be possible the chain has to have a length of at least 10 atoms.

In comparison to the Barnes update, a calculation with the long known Broyden update needs 30 iterations, to converge to the same optimal BBP.

The optimal BBP structure of the chain is depicted in Fig. 2. The SPs of index 1 are anti-kinks sitting on the tops of the substrate. The first one is on top 3, the next SPs of index one are the SPs of the tops 4 to $7 = N - 3$. After some turning points (TP), the NT here deviates from the MEP and increases uphill to the second floor. (Every NT connects SPs of an index difference of one,[37] so one cannot know the further direction of an NT from

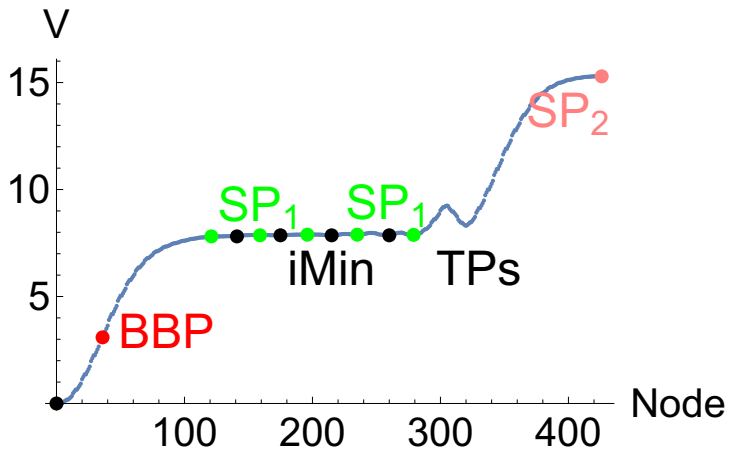


Figure 1: Energy profile over the NT to the optimal direction. Chain with 10 particles and misfit parameter 1. Optimal BBP red, Minima black, SP_1 green, and SP_2 pink.

a given SP_1: it can go downhill to the next minimum or uphill to an SP_2.) The gallery of the SP_1 and intermediate minima in between is already shown in Fig.16 of ref. 35.

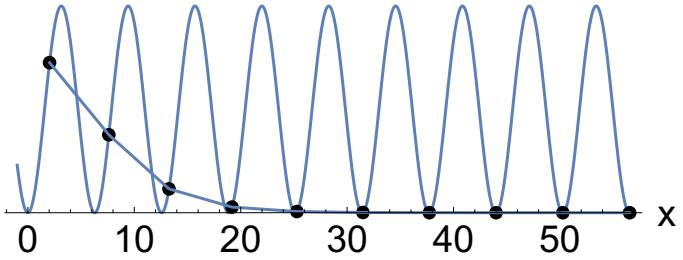


Figure 2: Optimal BBP of the chain with 10 particles and no misfit as result of the Barnes program. (The cos-fringes are the substrate potential. For better guiding of the imagination, we have shifted corresponding moved atoms of the chain on places of the substrate potential; the chain itself is a straight line. The spring potential between the atoms is not depicted here.)

In Table 1 we still compare the energy of the optimal BBP with the first BBPs of other NTs.

3.2 The Benzocyclobutene ring opening

The first real chemical example is devoted to the conrotatory electrocyclic ring opening of cis-1,2-dimethyl-benzocyclobutene to yield a E, Z-diene. This transformation was studied by us in a previous paper on the so-called gentlest-ascent-dynamics-conjugate-directions

Table 1: Comparison of BBPs

Direction	Energy
$(1,1,1,1,1,1,1,1,1,1)^T$	10.077
$(1,1,1,1,0,0,0,0,0,0)^T$	4.239
$\mathbf{f}_{opt}$, see (*)	3.047

(GAD-CD) transition state search algorithm[38] and will be named *bcb* in the following. The computational details for the *bcb* reaction are analogous to those presented below for the chorismate to prephenate reaction. The BBP of the *bcb* reaction will be located using the algorithm herein presented and also the algorithm introduced in Ref. 19, which was based on a Broyden update. This will allow us to compare the performance of the two algorithms. The stationary points (namely, reactant, product, and TS) on the bcb reaction were first determined, and their relative energies were computed, taking reactants as the energy zero. The structures of the stationary points are depicted in Fig. 3. Also, three geometrical pa-

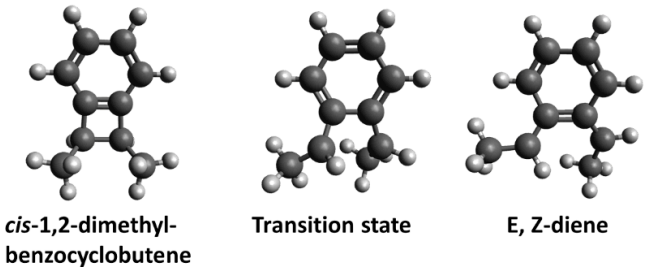


Figure 3: Structures of minima for reactants and products and of the transition state for the cis-1,2-dimethyl-benzocyclobutene to E, Z-diene reaction.

rameters which most characterize the bcb transformation are evaluated: the C−C distance for the carbon atoms of the cyclobutene part of the molecule that experience the ring opening process, and the two dihedral angles corresponding to these carbon atoms, the methyl substituents, and two of the carbon atoms on the benzene ring. For a precise definition of these dihedrals, we refer the reader to Fig. 6 of Ref. 38. The results are presented in Table

2, which also shows the energies and geometric parameters for the located BBP.

Table 2: Energies (in au), energies relative to cis-1,2-dimethyl- benzocyclobutene minimum (in kcal·mol^{-1}), and geometrical parameters (in Å and degrees) of stationary points and BBP for the *bcb* reaction at the M06 − 2X/def2-SVP level.

Structure	Energy	ΔE	C−C distance	Dihedral 1	Dihedral 2
bcb	−387.791756	0.0	1.586	59.5	−59.7
Transition state	−387.715414	47.9	2.333	39.1	−116.6
BBP	−387.725221	41.8	2.601	42.2	−129.7
E, Z-diene	−387.763148	18.0	3.050	4.9	−170.9

The BBP was determined setting the MNVAR variable to different values to ascertain the efficiency of the algorithm in these different situations, but the rest of computational parameters are identical. The MNVAR = 0 calculation corresponds to the update method of Ref. 19, and values of MNVAR of 5, 10, 20, and 30 were chosen for the Barnes update algorithm. As can be seen in Table 3, all calculations converge in a relatively small number of iterations and yield the BBP and the force vector at the BBP. In this particular case the optimal value for MNVAR is 10. Remarkably, for this value MNVAR the new algorithm, herein presented, converges much faster than the algorithm employed in Ref. 19. Thus, it can be concluded that the Barnes update is significantly more efficient than the previously reported algorithm. It is observed that the BBP of the *bcb* transformation is located in between the transition state and the product, E, Z-diene, but much closer both energetically and geometrically to the transition state.

Table 3: Values of MNVAR, number of iterations needed for convergence, energies (in au) and force vector norm (in au·bohr^{-1}) at the BBP for the BBP searches carried out for the *bcb* reaction.

MNVAR[a]	No. of iterations	Energy	Force vector norm
0	187	−387.723980	0.01491
5	134	−387.722273	0.01285
10	69	−387.725221	0.01471
20	82	−387.726186	0.01448
30	168	−387.724221	0.01417

[a] MNVAR = 0 corresponds to the algorithm of Ref. 19.

The molecular structure of the BBP and the force vector superimposed on this structure are

depicted in Fig. 4. The force vector is equal to the gradient at the BBP and represents the optimal pulling direction to carry out the conversion from *cis*-1,2-dimethyl-benzocyclobutene to E, Z-diene in the gas phase.

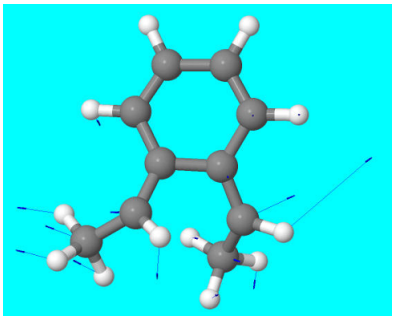


Figure 4: The molecular structure of the converged BBP for the *bcb* reaction and the force vector superimposed on it are shown.

3.3 The Chorismate-Mutase rearrangement

The second real chemical example that we have studied is the intramolecular Claisen rearrangement of chorismate to prephenate. This reaction is catalyzed by the chorismate mutase enzyme and is it part of the synthetic pathway to phenylalanine and tyrosine in bacteria, fungi, and higher plants.[39] Significant controversy about the enzymatic mechanism of this reaction, one of the few examples of enzyme-catalyzed pericyclic reaction, has arisen. Thus, it was discussed which of a transition state stabilization by electrostatic interactions in the enzyme active center, or a near attack conformation effect, was the main factor accounting for the enzyme catalytic enhancement of the reaction rate constant.[40] More recently, entropic effects have been proposed to be also part of the picture.[41]

In the present work, we do not aim at a thorough study of the chorismate to prephenate reaction, but rather we provide a first study of the reaction from the point of view of mechanochemistry. In particular, we have computed the BBP for prephenate production by interfacing the BBP-finding algorithm with the TURBOMOLE[42,43] electronic structure package. The M06-2X[44] density functional and Ahlrich's def2-SVP[45] basis set were chosen. An SCF error tolerance of 10^{-8}, a large $m5$ grid, and weight derivatives were introduced to improve the accuracy of the gradients and Hessians along the BBP search. Convergence of the BBP search procedure is achieved when either the maximum component of the σ-function or the square root of the quotient of the σ- function norm by the number of degrees of freedom is less than 1.0×10^{-3}.

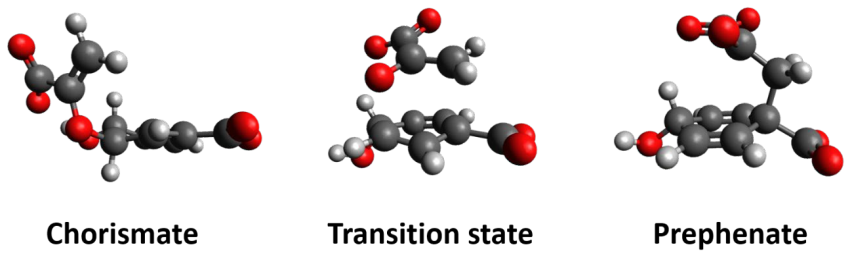


Figure 5: Structures of minima for reactants and products and of the transition state for the chorismate to prephenate rearrangement.

First, the stationary points on the gas-phase PES were determined separately. The structures of the stationary points are presented in Fig. 5. Note that the reaction involves the rupture of the initial C−O bond of chorismate and the formation of a C−C bond in prephenate. Thus, we can single out these bond distances as defining the reaction path. The energies of chorismate, the TS, and prephenate, along with their C−O and C−C bond distances, and the energies relative to chorismate, are detailed in Table 4. Chorismate has different conformers in the gas phase, and the conformer determined here is called 'OH-out' in the literature.[46] In this conformer, the −OH group is pointing towards one of the oxygens of the $-CO_2$ group, thereby forming an hydrogen bond. Another conformer, called 'OH-in', has the −OH group

pointing in the opposite direction with respect to the $-CO_2$ group, and originates a different reaction path. Note the evolution of the rearrangement from a short C−O distance and a long C−C distance in chorismate, to intermediate distances in the TS, and a long C−O distance and short C−C distance in prephenate. The relative energies reported in Table 4 can be compared to those reported in the literature at the B3LYP/6 − 31+G(d, p) level.[47] Thus, the $\Delta E^{\neq}$ energy is 50.8 kcal·mol^{-1} at the M06 − 2X/def2-SVP level, and it is 37.5 kcal·mol^{-1} at the B3LYP/6 − 31+G(d, p) level; and the respective ΔE^{R} energies are 5.2 and −10.5 kcal·mol^{-1}. Despite these differences, note that the aim of the present study is to determine the mechanochemical behaviour of the system rather than trying to obtain accurate results for the reaction.

Table 4: Energies (in au), energies relative to chorismate minimum (in kcal·mol^{-1}), and geometrical parameters (in Å) of stationary points and BBP for the chorismate to prephenate rearrangement at the M06 − 2X/def2-SVP level.

Structure	Energy	ΔE	C−O distance	C−C distance
Chorismate	−836.260031	0.0	1.369	5.282
Transition state	−836.179151	50.8	1.904	2.283
BBP	−836.201071	37.0	2.412	1.896
Prephenate	−836.251765	5.2	3.401	1.564

To proceed with the determination of the BBP, an intrinsic reaction coordinate (IRC) path was calculated starting with the TS geometry and ending in the prephenate product. Then, the point on the IRC with the largest gradient norm was chosen as the starting point for the BBP search. Note that this strategy has been shown to be appropriate for the 1, 2-sigmatropic H-shift rearrangement of cyclopentadiene in Ref. 19. The Cartesian coordinates of the initial point and of the BBP for the conversion to prephenate are printed in Supporting Information. The energy at the initial point is −836.210927 au, and the σ-function value is relatively small, i.e., 0.008130. The algorithm presented in this work was applied with different values of MNVAR, where MNVAR corresponds to the number of previous displacement vectors during their orthogonalization procedure (see Section 2.4). In particular, MNVAR values of 5, 10, 20, and 30 were defined. Note that the number of degrees of freedom is 72

for the chorismate-prephenate 24-atom system.

The number of iterations required to converge to the BBP, the energies and the force vector norm at the BBP are presented in Table 5. Note that for all calculations, it is always fulfilled that the square root of the quotient of the sigma-function and the number of degrees of freedom at the BBP is less than 1.0×10^{-3}, but the norm of the σ- function is always slightly larger than 1.0×10^{-3} (around 3.0×10^{-3}). More precisely, the value of the σ-function is about 7.0×10^{-5} at the BBP. As can be seen, the number of iterations to converge to the BBP vary significantly with MNVAR, and the optimal value is MNVAR = 30. This value is slightly less than half the number of degrees of freedom of the system. A BBP search was also carried out with the algorithm presented in Ref. 19, but it failed to converge within the 1000 maximum iteration limit. This farther demonstrates the significantly better performance of the new algorithm compared with respect to the original one. The energy, the C−O and C−C bond distances, and the relative energy to chorismate for the BBP are shown in Table 4. Comparing with the stationary points, one can observe that the BBP is located in between the TS and prephenate, but much closer both energetically and geometrically to the TS. The molecular structure of the BBP and the force vector superimposed on this structure are depicted in Fig. 6. The force vector is equal to the gradient at the BBP and represents the optimal pulling direction to carry out the conversion from prephenate to chorismate in the gas phase.

Table 5: Values of MNVAR, number of iterations to convergence, energies (in au) and force vector norm (in au·bohr^{-1}) at the BBP for the BBP searches.

MNVAR[a]	No. of iterations	Energy	Force vector norm
0	-[b]	-	-
5	996	−836.202004	0.05126
10	383	−836.201602	0.04949
20	199	−836.201072	0.04937
30	185	−836.202530	0.05016

[a] MNVAR = 0 corresponds to the algorithm of Ref. 19.
[b] No convergence was achieved after 1000 iterations.

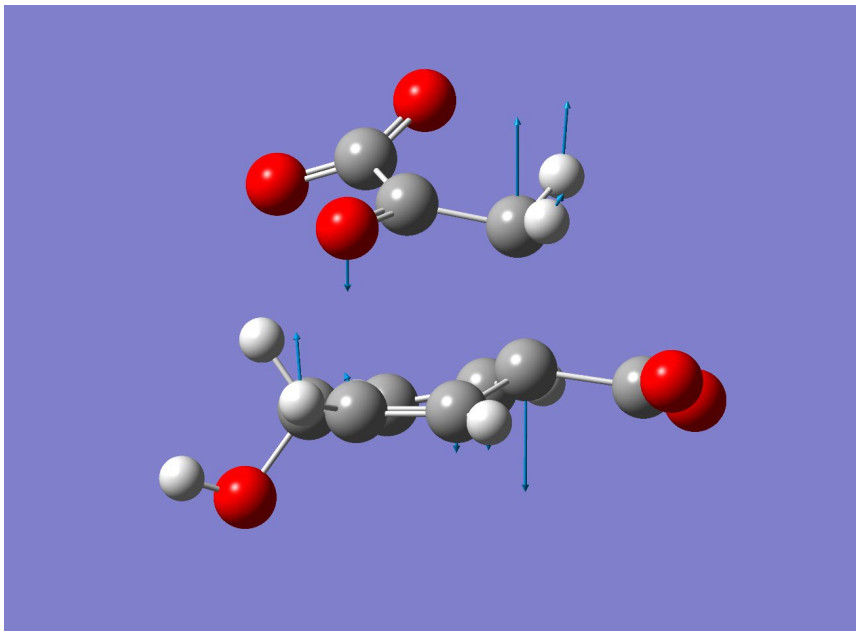


Figure 6: The molecular structure of the converged BBP and the force vector superimposed on it are shown.

4 Conclusion

We have presented a Gaussian-Newton method combined with a Barnes update of a Jacobian matrix that enables the location of points on the PES where the Hessian matrix has a zero eigenvalue and the corresponding zero eigenvector coincides with the gradient. The algorithm has been demonstrated to work efficiently for the Frenkel-Kontorova model PES and the ring opening of a benzocyclobutene derivative, and the Chorismate Mutase rearrangement. The critical points located by the algorithm herein presented are particularly important in the context of mechanochemistry because they coincide with the optimal BBPs of the system. The gradient in these points indicates how a pulling force should be applied to a molecular system to render a given chemical transformation barrierless using the smallest possible force. Therefore, our algorithm will be useful for identifying efficient ways of employing mechanical stress to enforce chemical reactions.

5 Appendix A: Consequences of the increasing Order of the Lagrangian Multipliers

For any tuple, let say i, Eq. 11.a can be multiplied from the left by $\Delta\mathbf{x}_i^T$. Taking the expression of Eq. 12 we have $\Delta\sigma^{(2)}(\Delta\mathbf{x}_i) = \lambda_i r^2 + \Delta\mathbf{x}_i^T\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$. Now we want to prove that if $\lambda_j > \lambda_i$ then $\Delta\sigma^{(2)}(\Delta\mathbf{x}_j) > \Delta\sigma^{(2)}(\Delta\mathbf{x}_i)$. For this purpose we first multiply from the left Eq. 11.a for the i-tuple by $\Delta\mathbf{x}_j^T$ and analogously for the j-tuple by $\Delta\mathbf{x}_i^T$, respectively. The difference between the two resulting equations gives

$$(\Delta\mathbf{x}_j - \Delta\mathbf{x}_i)^T\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) = -(\lambda_j - \lambda_i)\Delta\mathbf{x}_j^T\Delta\mathbf{x}_i$$

since $\Delta\mathbf{x}_j^T(\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}))\Delta\mathbf{x}_i = \Delta\mathbf{x}_i^T(\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}))\Delta\mathbf{x}_j$. Now taking this result we have

$$\Delta\sigma^{(2)}(\Delta\mathbf{x}_j) - \Delta\sigma^{(2)}(\Delta\mathbf{x}_i) = (\lambda_j - \lambda_i)(r^2 - \Delta\mathbf{x}_j^T\Delta\mathbf{x}_i) \ .$$

Any $\Delta\mathbf{x}_k$ vector for $k = 1, \ldots, N$ can be written as $\Delta\mathbf{x}_k = r\mathbf{u}_k$ where $\mathbf{u}_k^T\mathbf{u}_k = 1$, thus

$$\Delta\sigma^{(2)}(\Delta\mathbf{x}_j) - \Delta\sigma^{(2)}(\Delta\mathbf{x}_i) = (\lambda_j - \lambda_i)r^2(1 - \cos\theta_{j,i}) \tag{33}$$

$\theta_{j,i}$ being the angle formed between the unitary vectors $\mathbf{u}_i$ and $\mathbf{u}_j$. We conclude that $\Delta\sigma^{(2)}(\Delta\mathbf{x}_j) > \Delta\sigma^{(2)}(\Delta\mathbf{x}_i)$ if $\lambda_j > \lambda_i$ since $(1 - \cos\theta_{j,i}) > 0$. Note that the zero case corresponds to the situation that i-tuple and j-tuple are the same. A consequence of this relation is the following: the difference between $\Delta\sigma^{(2)}(\Delta\mathbf{x}_{max}) - \Delta\sigma^{(2)}(\Delta\mathbf{x}_{min})$ can be expressed as

$$\Delta\sigma^{(2)}(\Delta\mathbf{x}_{max}) - \Delta\sigma^{(2)}(\Delta\mathbf{x}_{min}) =$$

$$(\Delta\sigma^{(2)}(\Delta\mathbf{x}_{max}) - \Delta\sigma^{(2)}(\Delta\mathbf{x}_{max-1})) + (\Delta\sigma^{(2)}(\Delta\mathbf{x}_{max-1}) - \Delta\sigma^{(2)}(\Delta\mathbf{x}_{max-2})) + \ldots$$

$$+(\Delta\sigma^{(2)}(\Delta\mathbf{x}_{min-2}) - \Delta\sigma^{(2)}(\Delta\mathbf{x}_{min-1})) + (\Delta\sigma^{(2)}(\Delta\mathbf{x}_{min-1}) - \Delta\sigma^{(2)}(\Delta\mathbf{x}_{min})) \ .$$

Substituting in the latest equality Eq. 33, and after some rearrangements we obtain an interesting expression between the angles $\theta_{j,i}$

$$\cos\theta_{max,min} = \frac{1}{\lambda_{max} - \lambda_{min}} \sum_{k=min}^{max-1} (\lambda_{k+1} - \lambda_k) \cos\theta_{k+1,k} \ . \tag{34}$$

Finally, we prove that in the set of tuple-solutions of Eqs. 11 there is no degeneracy. For this purpose we assume that $\lambda_i = \lambda_j$. Then from the relation $(\Delta\mathbf{x}_j - \Delta\mathbf{x}_i)^T \mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) = -(\lambda_j - \lambda_i)\Delta\mathbf{x}_j^T \Delta\mathbf{x}_i = 0$ we conclude that it is satisfied if and only if $\Delta\mathbf{x}_j = \Delta\mathbf{x}_i$ since $\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) \neq \mathbf{0}$. Thus here the tuples i and j have to be the same tuple.

6 Appendix B: Relation between the solutions of Eqs. 11 and Eq. 16

In this appendix we show that the set of solutions for λ has been extended in Eq. 16 with respect to Eqs. 11.[25,26] As one will see this fact does not have consequences in the proposed algorithm and they already are taken into account and discussed in Section 2.1.

1. Let us start by assuming that λ and $\Delta\mathbf{x}$ are the solution of Eqs. 11, then Eq. 16 has a solution for this λ. We have to consider two cases, the first one is this that λ is an eigenvalue of $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ and the second case is that λ is not an eigenvalue of this matrix.

   (a) In the first case, let $\mathbf{t}(\mathbf{x})$ the eigenvector of the matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ with eigenvalue λ, then with Eq. 11.a we have $-\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) = [\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]\Delta\mathbf{x}$ implying that $-\mathbf{t}^T(\mathbf{x})\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) = \mathbf{t}^T(\mathbf{x})[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]\Delta\mathbf{x} = \mathbf{0}$. From this result follows that the first term of the left hand side part of Eq. 16, $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^2\mathbf{t}^T(\mathbf{x}) = \mathbf{0}$, and the second term, $-r^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}(\mathbf{x}) = \mathbf{0}$. Thus the λ solution of Eq. 11 being aneigenvalue of the matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$, with its corresponding eigenvector,

$\mathbf{t}(\mathbf{x})$, satisfies Eq. 16. Notice that in this case the vectors $\mathbf{t}(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$ are orthogonal.

(b) We consider now the second case where λ is not an eigenvalue of the matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$. From the definition of the vector $\mathbf{t}(\mathbf{x})$ given after Eq. 14 we can write, $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^2\mathbf{t}(\mathbf{x}) = -[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]\Delta\mathbf{x} = \mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$, where Eq. 13 as a solution of Eq. 11.a has been used. Thus the first term of the left hand side of Eq. 16 is equal to $\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$. Now the second term of the left hand side of Eq. 16 has the value, $-r^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}(\mathbf{x}) = -r^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) = -r^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})\Delta\mathbf{x}^T\Delta\mathbf{x} = -\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$, where the definition of the vector $\mathbf{t}(\mathbf{x})$, Eq. 13 and that $\Delta\mathbf{x}$ satisfies Eq. 11.b has been used. Thus Eq. 16 is again satisfied. Consequently from the pair $\Delta\mathbf{x}$ and the solution λ of Eq. 11, we always obtain an eigenvalue λ and the corresponding vector, $\mathbf{t}(\mathbf{x})$, being solutions of Eq. 16.

2. Now we treat the reverse situation, λ and $\mathbf{t}(\mathbf{x})$ are solutions of Eq. 16. We try to find a solution of Eq. 11. We have to consider to cases when λ is not an eigenvalue of the matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ and when it is an eigenvalue.

(a) In the first case since λ is not an eigenvalue of $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ always Eq. 13 exists obtaining $\Delta\mathbf{x}$ and thus Eq. 11.a is satisfied. Now multiplying Eq. 16 from the left by $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^{-2}$, we obtain, $\mathbf{t}(\mathbf{x}) = r^{-2}[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}(\mathbf{x})$. Since the eigenvector $\mathbf{t}(\mathbf{x}) \neq \mathbf{0}$ then the last equality implies that $\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}(\mathbf{x}) \neq 0$ and we can write $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^{-2}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x}) = \mathbf{t}(\mathbf{x})(\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{t}(\mathbf{x}))^{-1}r^2$. Multiplying this equality from the left by $\mathbf{s}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ and using Eq. 13 we obtain that the norm of $\Delta\mathbf{x}$ is r^2, thus Eq. 11.b is satisfied too. This result is also given in Section 2.1.

(b) In the second case we assume that λ is an eigenvalue of $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$. We use a

spectral decomposition. We define the vector, $\Delta\bar{\mathbf{x}} = [\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^+\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$, where the pseudoinverse $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]^+$ represents the inverse of the matrix $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]$ without the eigenvector of the matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ with eigenvalue λ. To solve in this case Eqs. 11, we have to consider the next distinctions:

(i) First, $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]\Delta\bar{\mathbf{x}} \neq \mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$, thus Eqs. 11 do not have a solution for this λ. The reason is that the vector $\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$ has a non-null component in the eigenvector associated with eigenvalue λ.

(ii) Second, $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda\mathbf{I}]\Delta\bar{\mathbf{x}} = \mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$. In this case three subcases appear. In the first subcase it is $\Delta\bar{\mathbf{x}}^T\Delta\bar{\mathbf{x}} > r^2$, thus Eqs. 11 do not have a solution for this λ because the normalization condition r^2 is not satisfied. The second subcase occurs when, $\Delta\bar{\mathbf{x}}^T\Delta\bar{\mathbf{x}} = r^2$, thus $\Delta\bar{\mathbf{x}} = \Delta\mathbf{x}$, being the unique solution of Eqs. 11 for this λ.

(iii) Third, $\Delta\bar{\mathbf{x}}^T\Delta\bar{\mathbf{x}} < r^2$, and Eqs. 11 have several solutions for this λ. Let us assume that the eigenvalue λ is l-degenerate being $\{\mathbf{l}^{(i)}\}_{i=1}^{l}$ the set of this l-degenerate orthonormal vectors. Notice that $\Delta\bar{\mathbf{x}}$ is orthogonal to this set of l-degenerate orthonormal vectors. With this notation we can write $\Delta\mathbf{x} = \Delta\bar{\mathbf{x}} + c_1\mathbf{l}^{(1)} + \cdots + c_l\mathbf{l}^{(l)}$, with $r^2 = \Delta\mathbf{x}^T\Delta\mathbf{x} = \Delta\bar{\mathbf{x}}^T\Delta\bar{\mathbf{x}} + c_1^2 + \cdots + c_l^2$ being solutions of Eqs. 11. This set of solutions constitutes a manifold of dimension $l - 1$.

In summary, the solvability of Eqs. 11 for the smallest λ_{min} must satisfy that $\lambda_{min} \leq \lambda_1^J$, being λ_1^J the smallest eigenvalue of the matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$. We assume that λ_{min} is the smallest eigenvalue of Eq. 16, with the above results we have the next three cases.

1) If it is $\lambda_{min} < \lambda_1^J$ then this implies that λ_{min} lies outside the spectrum of $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ and $\Delta\mathbf{x} = [\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda_{min}\mathbf{I}]^{-1}\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$ fulfills Eqs. 11 being the unique solution.

2) If it is $\lambda_{min} = \lambda_1^J$ then the vector $\Delta\bar{\mathbf{x}} = [\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda_{min}\mathbf{I}]^+\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$ must satisfy

the equation $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda_{min}\mathbf{I}]\Delta\bar{\mathbf{x}} = \mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$. If in this case $\Delta\bar{\mathbf{x}}^T\Delta\bar{\mathbf{x}} = r^2$, then $\Delta\bar{\mathbf{x}} = \Delta\mathbf{x}$ is the unique solution of Eqs. 11.

3) If $\lambda_{min} = \lambda_1^J$ and $\Delta\bar{\mathbf{x}}$ satisfies $[\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) - \lambda_{min}\mathbf{I}]\Delta\bar{\mathbf{x}} = \mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$ but $\Delta\bar{\mathbf{x}}^T\Delta\bar{\mathbf{x}} < r^2$ then we must find a vector $\bar{\mathbf{l}}$ being a linear combination of the set of eigenvectors, $\{\mathbf{l}^{(i)}\}_{i=1}^{l}$, corresponding to the l-degenerate eigenvalue λ_1^J of the matrix $\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x})$ such that $r^2 = \Delta\bar{\mathbf{x}}^T\Delta\bar{\mathbf{x}} + \bar{\mathbf{l}}^T\bar{\mathbf{l}}$. Then $\Delta\mathbf{x} = \Delta\bar{\mathbf{x}} + \bar{\mathbf{l}} = \Delta\bar{\mathbf{x}} + c_1\mathbf{l}^{(1)} + \cdots + c_l\mathbf{l}^{(l)}$ represents one of the many solutions of Eqs. 11.

7 Appendix C: The Variance-Covariance Matrix obtained from the Algorithm

An important property of this algorithm is that it provides an approximation to the least squares variance–covariance matrix. According to Eq. 8 the derivatives that appear in Eq. 9 with respect to $\mathbf{x}$ have the form $\nabla_{\mathbf{x}}\sigma(\mathbf{x}) = 2\ [\nabla_{\mathbf{x}}\mathbf{s}^T(\mathbf{x})]\mathbf{s}(\mathbf{x}) = 2\mathbf{J}(\mathbf{x})\mathbf{s}(\mathbf{x})$ and $\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}^T\sigma(\mathbf{x}) = 2\mathbf{J}(\mathbf{x})\mathbf{J}^T(\mathbf{x}) + 2\sum_{i=1}^{N} s_i(\mathbf{x})\ [\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}^T s_i(\mathbf{x})]$ where $s_i(\mathbf{x})$ is the ith component of the $\mathbf{s}(\mathbf{x})$ vector and $\mathbf{J}(\mathbf{x}) = \ [\nabla_{\mathbf{x}}s_1(\mathbf{x})|\cdots|\nabla_{\mathbf{x}}s_N(\mathbf{x})]$. It is obvious that if $\sigma(\mathbf{x}_0) = 0$ then $\mathbf{s}(\mathbf{x}_0) = \mathbf{0}$ thus the derivatives that appear in Eq. 9 take the values $\nabla_{\mathbf{x}}\sigma(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_0} = \mathbf{0}$ and $\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}^T\sigma(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_0} = 2\mathbf{J}(\mathbf{x}_0)\mathbf{J}^T(\mathbf{x}_0)$. From these results Eq. 9 reduces to $\sigma(\mathbf{x}) = (\mathbf{x}-\mathbf{x}_0)^T\ [\mathbf{J}(\mathbf{x}_0)\mathbf{J}^T(\mathbf{x}_0)\ ](\mathbf{x}-\mathbf{x}_0)$. Taking the definition of the function $\sigma(\mathbf{x})$, the latter equality can be written as

$$\sigma(\mathbf{x}) = Tr\ [(\sigma(\mathbf{x})/N)\mathbf{I}\ ] = Tr\ [\mathbf{s}(\mathbf{x})\mathbf{s}^T(\mathbf{x})\ ] = Tr\ [(\mathbf{x}-\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)^T\ \mathbf{J}(\mathbf{x}_0)\mathbf{J}^T(\mathbf{x}_0)\ ]$$

where Tr means trace, $\mathbf{I}$ is the identity matrix of dimension N x N and the cyclic permutation of trace has been used. Note that $Det[\mathbf{J}(\mathbf{x}_0)\mathbf{J}^T(\mathbf{x}_0)\ ] > 0$ since the function $\sigma(\mathbf{x})$ has a minimum in $\mathbf{x}_0$. If we take off the Tr operator we can write

$$(\sigma(\mathbf{x})/N)\mathbf{I} = (\mathbf{x}-\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)^T\ \mathbf{J}(\mathbf{x}_0)\mathbf{J}^T(\mathbf{x}_0)\ .$$

Multiplying this equality from the right by $[\mathbf{J}(\mathbf{x}_0)\mathbf{J}^T(\mathbf{x}_0)\,]^{-1}$ we obtain the variance-covariance matrix $\mathbf{V}$

$$\mathbf{V} = N(\mathbf{x}-\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)^T = [\mathbf{J}(\mathbf{x}_0)\mathbf{J}^T(\mathbf{x}_0)\,]^{-1}\sigma(\mathbf{x})\ . \tag{35}$$

Thus the $\mathbf{V}$-matrix is a multiple of the inverse matrix of $\mathbf{J}(\mathbf{x}_0)\mathbf{J}^T(\mathbf{x}_0)$. The diagonal elements of $\mathbf{V}$ give the variance of the elements of $\{x_i - x_i^0\}_{i=1}^N$ in the maximum probability solution, $\mathbf{x}_0$, whereas the off-diagonal elements give the covariance between $x_i - x_i^0$ and $x_j - x_j^0$. The $\{x_i\}_{i=1}^N$ and $\{x_i^0\}_{i=1}^N$ are the elements of the vectors $\mathbf{x}$ and $\mathbf{x}_0$, respectively. Thus $\mathbf{V}$ is directly related with the inverse of the matrix $\mathbf{J}(\mathbf{x}_0)\mathbf{J}^T(\mathbf{x}_0)$ according to Eq. 35. Using the notations, definitions and results of Section 2.4 we have that $\mathbf{J}\mathbf{J}^T = \mathbf{D}^{+T}\mathbf{Y}^T\mathbf{Y}\mathbf{D}^+$ since $\mathbf{J}^T = \mathbf{Y}\mathbf{D}^+$. We recall that the matrices, $\mathbf{Y} = [\mathbf{y}^{(1)}|\ldots|\mathbf{y}^{(N)}]$, $\mathbf{D} = [\mathbf{d}^{(1)}|\ldots|\mathbf{d}^{(N)}]$ and $\mathbf{D}^+ = (\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T$ are built if the Barnes update formula, Eq. 28, is used. Note that now the $\mathbf{D}$ matrix has the full dimension $N \times N$ where N is the dimension of the full space. Thus it holds $\mathbf{D}^+ = \mathbf{D}^{-1}$. The equality $(\mathbf{J}\mathbf{J}^T)^{-1}(\mathbf{J}\mathbf{J}^T) = \mathbf{I} = \mathbf{D}\mathbf{D}^+$ is fulfilled, where the resolution of identity has been used. Substituting in the last equality the expression for $\mathbf{J}\mathbf{J}^T$ we have, $(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{D}^{+T}\mathbf{Y}^T\mathbf{Y}\mathbf{D}^+ = \mathbf{D}\mathbf{D}^+$. Multiplying this expression from the right, first by $\mathbf{D}$, second by $(\mathbf{Y}^T\mathbf{Y})^{-1}$, third by $\mathbf{D}^T$ and taking into account that $\mathbf{D}^{+T}\mathbf{D}^T = \mathbf{D}(\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T = \mathbf{I}$ we obtain $(\mathbf{J}\mathbf{J}^T)^{-1} = \mathbf{D}(\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{D}^T \approx \mathbf{V}$, which is the desired result.

Acknowledgement

The authors thank the financial support from the Spanish Ministerio de Economía y Competitividad, Projects No. PID2019-109518GB-I00, CTQ2017-87773-P/AEI/ FEDER, Spanish Structures of Excellence María de Maeztu program through grant MDM-2017-0767 and Generalitat de Catalunya, Project No. 2017 SGR 348.

Supporting Information Available

The following file is available free of charge.

- Filename: Supporting-Info.pdf

  Cartesian coordinates of the initial point and the BBP for the conversion to prephenate.

An animation file with filename: FKm10DanimateBarnesBBP.gif is attended.

References

(1) Eying, H.; Walter, J.; Rimball, G. E. *Quantum Chemistry*; John Wiley and Sons, Inc., New York, 1944.

(2) Bell, G. I. Models for the specific adhesion of cells to cells. *Science* **1978**, *200*, 618 – 627.

(3) Bustamante, C.; Chemla, Y. R.; Forde, N. R.; Izhaky, D. Mechanical processes in Biochemistry. *Ann. Rev. Biochem.* **2004**, *73*, 705 – 748.

(4) Ong, M. T.; Leiding, J.; Tao, H.; Virshup, A. M.; Martínez, T. J. First Principles Dynamics and Minimum Energy Pathways for Mechanochemical Ring Opening of Cyclobutene. *J. Am. Chem. Soc.* **2009**, *131*, 6377– 6379.

(5) Ribas-Ariño, J.; Marx, D. Covalent Mechanochemistry: Theoretical Concepts and Computational Tools with Applications to Molecular Nanomechanics. *Chem. Rev.* **2012**, *112*, 5412–5487.

(6) Wolinski, K.; Baker, J. Theoretical predictions of enforced structural changes in molecules. *Molec. Phys.* **2009**, *107*, 2403 – 2417.

(7) Konda, S. S. M.; Avdoshenko, S. M.; Makarov, D. E. Exploring the Topography of the Stress-modified Energy Landscapes of Mechanosensitive Molecules. *J. Chem. Phys* **2014**, *140*, 104114.

(8) Avdoshenko, S. M.; Makarov, D. E. Reaction Coordinates and Pathways of Mechanochemical Transformations. *J. Phys. Chem. B* **2015**, *120*, 1537 – 1545.

(9) Subramanian, G.; Mathew, N.; Leiding, J. A generalized force-modified potential energy surface for mechanochemical simulations. *J.Chem.Phys.* **2015**, *143*, 134109.

(10) Avdoshenko, S. M.; Makarov, D. E. Finding Mechanochemical Pathways and Barriers without Transition State Search. *J. Chem. Phys* **2015**, *142*, 174106.

(11) Makarov, D. E. Perspective: Mechanochemistry of Biological and Synthetic Molecules. *J. Chem. Phys.* **2016**, *144*, 030901.

(12) Quapp, W.; Bofill, J. M.; Ribas-Ariño, J. Analysis of the Acting Forces in a Theory of Catalysis and Mechanochemistry. *J.Phys.Chem.A* **2017**, *121*, 2820 – 2838.

(13) Quapp, W.; Bofill, J. M. A contribution to a theory of mechanochemical pathways by means of Newton trajectories. *Theor. Chem. Acc.* **2016**, *135*, 113.

(14) Quapp, W.; Bofill, J. M. Reaction Rates in a Theory of Mechanochemical Pathways. *J. Comput. Chem.* **2016**, *37*, 2467 – 2478.

(15) Ribas-Ariño, J.; Shiga, M.; Marx, D. Understanding Covalent Mechanochemistry. *Angew. Chem. Int. Ed.* **2009**, *48*, 4190 – 4193.

(16) Branin, F. H. Widely Convergent Methods for Finding Multiple Solutions of Simultaneous Nonlinear Equations. *IBM J. Res. Develop.* **1972**, *16*, 504–522.

(17) Quapp, W.; Hirsch, M.; Heidrich, D. An approch to reaction path branching using valley-ridge-inflection points of Potential Energy surfaces. *Theor. Chem. Acc.* **2004**, *112*, 40 – 51.

(18) Bofill, J. M.; Quapp, W. Analysis of the Valley-Ridge inflection points through the partitioning technique of the Hessian eigenvalue equation. *J. Math. Chem.* **2013**, *51*, 1099 – 1115.

(19) Bofill, J. M.; Ribas-Ariño, J.; García, S. P.; Quapp, W. An Algorithm to Locate Optimal Bond Breaking Points on a Potential Energy Surface. *J.Chem.Phys.* **2017**, *147*, 152710–10.

(20) Fletcher, R. *Practical Methods of Optimization*; Wiley-Interscience, New York, 1987.

(21) Simons, J.; Jørgensen, P.; Taylor, H.; Ozment, J. Walking on Potential Energy Surfaces. *J. Phys. Chem.* **1983**, *87*, 2745 – 2753.

(22) Banerjee, A.; Adams, N.; Simons, J.; Shepard, R. Search for Stationary Points on Surfaces. *J. Phys. Chem.* **1985**, *89*, 52 – 57.

(23) Anglada, J. M.; Bofill, J. M. On the Restricted Step Method Coupled with the Augmented Hessian for the Search of Stationary Points of Any Continuous Function. *Int. J. Quantum Chem.* **1997**, *62*, 153 – 165.

(24) Besalú, E.; Bofill, J. M. On the automatic restricted-step rational-function-optimization method. *Theor. Chem. Acc.* **1998**, *100*, 265 – 274.

(25) Golub, G. H. Some Modified Matrix Eigenvalue Problems. *SIAM Rev.* **1973**, *15*, 318 – 334.

(26) Tisseur, F.; Meerbergen, K. The Quadratic Eigenvalue Problem. *SIAM Rev.* **2001**, *43*, 235 – 286.

(27) Broyden, C. G. A class of methods for solving nonlinear simultaneous equations. *Math. Comp.* **1965**, *19*, 577 – 593.

(28) Barnes, J. G. P. An algorithm for solving nonlinear equations based on the secant method. *Comp. J.* **1965**, *8*, 66 – 72.

(29) Fletcher, R. A Technique for Orthogonalization. *J. Inst. Math. Appl.* **1969**, *5*, 162 – 166.

(30) Ben-Israel, A. An Iterative Method for Computing the Generalized Inverse of an Arbitrary Matrix. *Math. Comp.* **1965**, *19*, 452 – 455.

(31) Braun, O. M.; Kivshar, Y. S. Nonlinear dynamics of the Frenkel-Kontorova model. *Phys. Rep.* **1998**, *306*, 1–108.

(32) Tekić, J.; Mali, P. *The ac driven Frenkel-Kontorova model*; University of Novi Sad: Novi Sad, 2015.

(33) Quapp, W.; Bofill, J. M. Newton Trajectories for the Frenkel-Kontorova model. *Molec. Phys.* **2019**, *117*, 1541–1558.

(34) Quapp, W.; Bofill, J. M. A Model for a Driven Frenkel-Kontorova Chain. *European Phys. J. B* **2019**, *92*, 95–117.

(35) Quapp, W.; Bofill, J. M. Sliding Paths for Series of Frenkel-Kontorova Models - A Contribution to the Concept of 1D-superlubricity. *European Phys. J. B* **2019**, *92*, 193.

(36) Savage, T. J.; Hristova, M. K.; Croteau, R. In *Physiology, Biochemistry and Molecular Biology of Plant Lipids*; Williams, J. P., Khan, M. U., Lem, N. W., Eds.; Springer: Dordrecht, 1997; pp 51–53.

(37) Hirsch, M.; Quapp, W. Reaction Channels of the Potential Energy Surface: Application of Newton Trajectories. *J. Molec. Struct., THEOCHEM* **2004**, *683*, 1–13.

(38) Bofill, J. M.; Ribas-Ariño, J.; Valero, R.; Albareda, G.; Moreira, I. d. P. R.; Quapp, W. Interplay between the Gentlest Ascent Dynamics Method and Conjugate Directions to Locate Transition States. *Journal of Chemical Theory and Computation* **2019**, *15*, 5426–5439.

(39) Haslam, E. *Shikimic Acid: Metabolism and Metabolites*; John Wiley & Sons: New York, 1993.

(40) Štrajbl, M.; Shurki, A.; Kato, M.; Warshel, A. Apparent NAC Effect in Chorismate Mutase Reflects Electrostatic Transition State Stabilization. *J. Am. Chem. Soc.* **2003**, *125*, 10228–10237.

(41) Brickel, S.; Meuwly, M. Molecular Determinants for Rate Acceleration in the Claisen Rearrangement Reaction. *J. Phys. Chem. B* **2019**, *123*, 448–456.

(42) TURBOMOLE V7.3 2018, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007. `http://www.turbomole.com`.

(43) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kälmel, C. Electronic structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.* **1989**, *162*, 165–169.

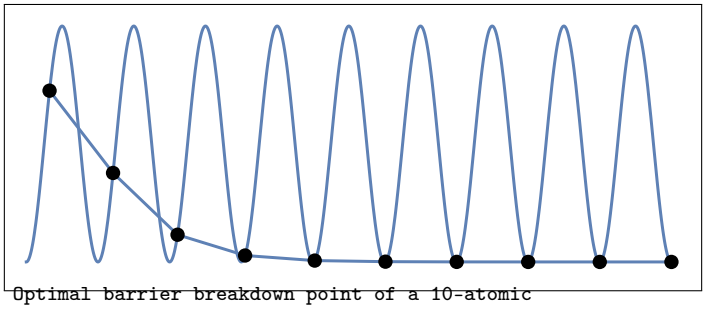
(44) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(45) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

(46) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. Transition structure selectivity in enzyme catalysis: a QM/MM study of chorismate mutase. *Theor. Chem. Acc.* **2001**, *105*, 207–212.

(47) Freindorf, M.; Tao, Y.; Sethio, D.; Cremer, D.; Kraka, E. New mechanistic insights into the Claisen rearrangement of chorismate - a Unified Reaction Valley Approach study. *Mol. Phys.* **2019**, *117*, 1172–1192.

Graphical TOC Entry



Optimal barrier breakdown point of a 10-atomic Frenkel-Kontorova chain. The optimal external force moves already here the chain over the first saddle point.