



Transition state optimization of periodic systems using delocalized internal coordinates

Tomáš Bučko^{1,2}

Received: 15 June 2018 / Accepted: 25 September 2018 / Published online: 2 November 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

In this work, we adapt our algorithm for relaxations of periodic systems (Bučko et al. in *J Chem Phys* 122: 124508, 2005) in delocalized internal coordinates of Baker et al. (*J Chem Phys* 105: 192, 1996) for the use in transition state geometry optimizations. The abilities of our algorithm are demonstrated on examples of relaxations of atomic positions and cell geometries of systems with and without additional geometric constraints that include transition states for reactions of molecules in the gas phase, reconnection of H atoms in the one-dimensional periodic chain of H₂ molecules, proton transfer in zeolite chabazite, partial desorption of crotonaldehyde from the MgO surface, and a pure affine shear deformation of Al. A simple approximate initial Hessian is suggested, in which only the matrix elements corresponding to atoms actively participating in reaction of interest are determined accurately at a DFT level, while remaining elements, typically related to inactive atoms and lattice vectors components, are defined on a basis of a simple empirical model. The calculations employing the approximate Hessian are shown to be more effective compared to simulations carried out with exact initial Hessian, in which all elements related to atomic positions are computed at the DFT level.

Keywords Geometry optimization · Internal coordinates · Transition state · Periodic DFT

1 Introduction

Analysis of potential energy surfaces (PES) provides an useful information about properties of a system at hands. The shape of PES in vicinity of important stationary points, such as minima (i.e., stable states) and first-order saddle points [i.e., transition states (TS)], is as a key ingredient for the calculations based on the harmonic transition state theory [1], which is presently the most popular method to study reaction kinetics. Furthermore, the relaxations are often performed as

a first step in the analysis of diverse physical properties of a system at hands, and hence they certainly represent one of the most common tasks executed by computational chemists. A determination of minima is relatively straightforward - in principle, one can just follow direction of negative gradient to descent from arbitrary initial point to some local minimum. In a first-order saddle point optimization, however, energy must be maximized along one and only one direction parallel with reaction coordinate (RC) and minimized along all remaining independent degrees of freedom. The fact that the reaction coordinate is usually unknown a priori is one of the reasons why the TS relaxations are typically much more difficult to perform than the stable states optimizations. Hence the computational algorithms designed for the TS optimization require, one way or the other, an initial guess for the unstable direction. The algorithms such as the quasi-Newton method [2], the partitioned rational function optimization [3] (P-RFO), or the geometrical direct inversion in the iterative subspace (DIIS) method [4], receive this information from the eigenpairs of the matrix of second order partial derivatives of energy with respect to the coordinates of system (Hesse matrix or Hessian), which has to be provided, at least for the initial configuration, by the user.

Published as part of the special collection of articles In Memoriam of János Ángyán.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00214-018-2367-0>) contains supplementary material, which is available to authorized users.

✉ Tomáš Bučko
bucko19@uniba.sk

¹ Comenius university in Bratislava, Ilkovičova 6,
84215 Bratislava, Slovakia

² Institute of Inorganic Chemistry, Slovak Academy
of Sciences, Dúbravská cesta 9, 84236 Bratislava, Slovakia

The initial unstable direction must be explicitly specified also in the dimer method [5], while the nudged elastic band method [6] (NEB) optimizes whole RC connecting the initial and the final stable states and the user must define several intermediate states lying on the guessed initial reaction coordinate. For a detailed review of TS optimization methods, the reader is referred e.g., to Ref. [7].

Many chemical reactions and other activated processes of technological interest, such as surface diffusion and phase transitions, occur in solid state. Due to practical reasons, these processes are commonly simulated using periodic boundary conditions (PBC). In addition to atomic positions, the periodic systems are defined also by the size and shape of their unit cells. The methods allowing for a simultaneous optimization of atomic and lattice degrees of freedom, such as the solid state NEB [8, 9] and the solid state dimer methods [10], are surprisingly scarce. In this work, we extend our algorithm for periodic systems relaxation [11] in delocalized internal coordinates [12], which we briefly review in Sect. 2.1, for the use in geometry optimizations of transition states. The usefulness of internal coordinates in the TS optimizations has already been well documented for molecular systems [13–15]. One of the greatest advantages of methods based on internal coordinates is the availability of simple and reasonable initial Hesse matrix models. In Sect. 2.2, we define the Hesse matrix for the use in periodic TS optimizations and we suggest a simple model allowing for its inexpensive construction. The abilities of our algorithm are demonstrated in Sect. 3, where optimizations of various molecular and periodic systems with and without additional geometric constraints are discussed. Finally, summary and conclusions are presented in Sect. 4.

2 Methods

2.1 Delocalized internal coordinates for optimization of periodic systems

In this section, we briefly review the use of delocalized internal coordinates of Baker et al. [12] in optimization of periodic systems, which we proposed in our previous work [11]. We note that an alternative formulation of the same problem has been reported by Andzelm et al. [16].

Let us consider a periodic system consisting of N_{at} atoms contained within a single unit cell formed by three lattice vectors \mathbf{a}_1 , \mathbf{a}_2 , and \mathbf{a}_3 , which we arrange in the matrix $\mathbf{h} = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]$. It is convenient to define positions of atoms within the unit cell by fractional coordinates $s = \{s_\alpha^a; a = 1, \dots, N_{\text{at}}; \alpha = 1 \equiv a_1, 2 \equiv a_2, 3 \equiv a_3\}$, which, by convention, fulfill the condition $0 \leq s_\alpha^a < 1$, and which are related to Cartesian coordinates of the same set of atoms $\mathbf{r} = \{r_\alpha^a; a = 1, \dots, N_{\text{at}}; \alpha = 1 \equiv x, 2 \equiv y, 3 \equiv z\}$ shifted by the

$\mathbf{L} = (l_1, l_2, l_3)$ multiples of the lattice vectors by the following linear transformation:

$$r_\alpha^{a,L_a} = \sum_{\beta=1}^3 h_{\alpha\beta} (s_\beta^a + l_\beta^a). \quad (1)$$

Primitive internal coordinates (\mathbf{q}) for any connected set of atoms are defined by collecting all interatomic distances within expected bond-lengths. Subsequently, bond-lengths are combined to define bond-angles and dihedral angles. Note that for the optimization of periodic systems, both the intra-cell (i.e., those defined by atoms from the same unit cell) and inter-cell (i.e., those defined by atoms from different cells) coordinates must be taken into account. Due to the translational symmetry, however, only the coordinates which are defined via at least one atom located in a reference cell (i.e., $\mathbf{L} = 0$) have to be considered. Having accumulated the information on connectivity of individual atoms, mutually unconnected molecular fragments can be identified and interconnected using suitable coordinates, such as inverse-power distances [17]. The number of internal coordinates generated in such a way is typically much greater than the number of independent degrees of freedom of the system (e.g., $3N_{\text{at}} + 3$ for an unconstrained three-dimensional periodic system). Baker et al. [12] proposed to replace such a redundant set of internal coordinates by a non-redundant set of delocalized internal coordinates ($\tilde{\mathbf{q}}$), related to \mathbf{q} via the following transformation:

$$\tilde{\mathbf{q}} = \mathbf{U}^t \mathbf{q}, \quad (2)$$

where the transformation matrix \mathbf{U} is formed by eigenvectors associated with nonzero eigenvalues of the matrix $\mathbf{B}\mathbf{B}^t$. In the case of a molecular system, the Wilson's B-matrix [18] \mathbf{B} is defined as the matrix of derivatives of \mathbf{q} with respect to the Cartesian coordinates of all atoms (i.e., $\delta\mathbf{q} = \mathbf{B}\delta\mathbf{r}$). As suggested in our previous work [11], the definition of \mathbf{B} can be generalized for the use in optimizations of periodic systems as follows:

$$(\delta\mathbf{q}) = (\mathbf{B}^{qs} \ \mathbf{B}^{qh}) \begin{pmatrix} \delta\mathbf{s} \\ \delta\mathbf{h} \end{pmatrix}, \quad (3)$$

where the blocks \mathbf{B}^{qs} and \mathbf{B}^{qh} are defined by the following equations:

$$\begin{aligned} B_{i,a_\alpha}^{qs} &= \frac{\partial q_i(r_\alpha^{a,L_a}, r_\beta^{b,L_b}, \dots)}{\partial s_\alpha^a} = \sum_{L_a} \sum_{\beta} \frac{\partial q_i}{\partial r_\beta^{a,L_a}} \frac{\partial r_\beta^{a,L_a}}{\partial s_\alpha^a} \\ &= \sum_{L_a} \sum_{\beta} \frac{\partial q_i}{\partial r_\beta^{a,L_a}} h_{\beta\alpha}, \end{aligned} \quad (4)$$

and

$$\begin{aligned}
 B_{i,\alpha\beta}^{qh} &= \frac{\partial q_i \left(r_\alpha^{a,L_a}, r_\beta^{b,L_b}, \dots \right)}{\partial h_{\alpha\beta}} \\
 &= \sum_{a,L_a} \sum_{\gamma} \frac{\partial q_i}{\partial r_\gamma^{a,L_a}} \frac{\partial r_\gamma^{a,L_a}}{\partial h_{\alpha\beta}} \\
 &= \sum_{a,L_a} \sum_{\gamma} \frac{\partial q_i}{\partial r_\alpha^{a,L_a}} \delta_{\alpha\gamma} \left(s_\beta^a + l_\beta^a \right) \\
 &= \sum_{a,L_a} \frac{\partial q_i}{\partial r_\alpha^{a,L_a}} \left(s_\beta^a + l_\beta^a \right).
 \end{aligned} \tag{5}$$

The gradient vector \mathbf{f} used in optimizations of periodic systems must involve contributions from the energy derivatives with respect to the fractional coordinates:

$$\frac{\partial E}{\partial s_\alpha^a} = \sum_{\beta} \frac{\partial E}{\partial r_\beta^a} h_{\beta\alpha}, \tag{6}$$

as well as those with respect to the lattice vectors components:

$$\frac{\partial E}{\partial h_{\alpha\beta}} = -\Omega \sum_{\mu} \sigma_{\alpha\mu} (h)_{\beta\mu}^{-1}, \tag{7}$$

where $\Omega = |\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)|$ is the volume of unit cell, and σ is the stress tensor. The gradients expressed in terms of delocalized internal coordinates ($\boldsymbol{\varphi}$) can be obtained from \mathbf{f} using the relation:

$$\boldsymbol{\varphi} = \mathbf{U}^t \mathbf{A}^t \mathbf{f}, \tag{8}$$

with \mathbf{A}^t being a Moore–Penrose pseudoinverse of matrix $\mathbf{B} = (\mathbf{B}^{qs} \mathbf{B}^{qh})$:

$$\mathbf{A} = (\mathbf{B}^t \mathbf{B})^{-1} \mathbf{B}^t. \tag{9}$$

Upon determining the geometry optimization step in the space of delocalized internal coordinates by using an appropriate relaxation algorithm, the corresponding set of fractional coordinates and lattice vectors components is obtained in an iterative procedure:

$$(\mathbf{s}^{i+1} \mathbf{h}^{i+1}) = (\mathbf{s}^i \mathbf{h}^i) + \mathbf{A} \mathbf{U} (\tilde{\mathbf{q}}^{opt} - \tilde{\mathbf{q}}^i), \tag{10}$$

where $\tilde{\mathbf{q}}^{opt}$ are the delocalized internal coordinates determined by the optimization algorithm, and $\tilde{\mathbf{q}}^i$ are the delocalized internal coordinates calculated using the coordinates \mathbf{s}^i and \mathbf{h}^i from the iteration i , and the iteration is initialized using the values of coordinates from the previous optimization step.

The geometric constraints can be implemented in a straightforward way within the framework of delocalized internal coordinates [12]. The essential point is that the

matrix $\mathbf{B} = (\mathbf{B}^{qs} \mathbf{B}^{qh})$ must be modified in such a way that the vectors \mathbf{B}_j corresponding to coordinates free to relax are made orthogonal to all vectors \mathbf{B}_c corresponding to constrained coordinates, i.e., the vectors \mathbf{B}_j must be modified according to formula:

$$\bar{\mathbf{B}}_j = \mathbf{B}_j - \sum_c \frac{(\mathbf{B}_j \cdot \mathbf{B}_c)}{|\mathbf{B}_c|} \frac{\mathbf{B}_c}{|\mathbf{B}_c|}, \tag{11}$$

where the summation is over all constrained coordinates. The delocalized internal coordinates generated using the modified matrix $\bar{\mathbf{B}}$ correspond to linear combinations of either constrained or active coordinates but the two types are never mixed in a definition of individual \tilde{q} . The energy gradients for \tilde{q} defined via constrained coordinates are then set to zero and the relaxation proceeds as in an unconstrained case. In principle, any coordinate with well defined first order derivatives with respect to atomic positions and lattice vectors components (e.g., bonds, angles, torsions, fractional, and Cartesian coordinates, lengths of lattice vectors and angles between them, cell volume, lattice vector components...) can be constrained in such a way.

2.2 Hesse matrix for transition state optimization

In our optimization program GADGET [11], the optimization methods that employ the Hesse matrix (\mathbf{H}) are implemented. The matrix \mathbf{H} used in the TS optimizations should have one and only one negative eigenvalue with the corresponding eigenvector being approximately parallel with the reaction coordinate. Ideally, the Hesse matrix should be computed in every optimization step but such a procedure would be too time-consuming when performed for an extended system at ab initio level. When a finite-differences algorithm is used (as is often the case when a periodic DFT code is used in simulations), for instance, at least $3N_{\text{at}}$ additional gradient evaluations are needed for the \mathbf{H} calculation, and the related computational cost increases the total cost of optimization. A commonly used strategy to tackle this problem [14, 15] is to determine \mathbf{H} only for the initial structure and update it for the use in the next relaxation steps by employing the information on positions and gradients accumulated during the optimization. Out of the Hesse matrix updating schemes suitable for the transition state optimizations, we chose the scheme proposed by Boffill [19], which is based on a weighted combination of Murtagh–Sargent [20] and Powell-symmetric-Broyden [21] formulae.

In our TS optimizations, we define the initial Hesse matrix as follows. First, we approximate the Hesse matrix in primitive internal coordinates (\mathbf{H}^p) by a simple diagonal matrix [12] in which the force-constants for all bonds, angles, and torsions are set to 0.5 a.u., 0.2 a.u., and 0.1 a.u.,

respectively. We note that more sophisticated models, such as those by Lindh [22] or Fischer [23], can also be used for this purpose. Next, the \mathbf{H}^q is transformed into the $\{s, \mathbf{h}\}$ coordinates using the approximate relation that neglects the contribution of terms involving the second order derivatives of primitive internal coordinates with respect to the atomic positions [14] and the lattice vectors components:

$$\begin{pmatrix} \mathbf{H}^s & \mathbf{H}^{sh} \\ (\mathbf{H}^{sh})^t & \mathbf{H}^h \end{pmatrix} \approx \mathbf{B}^t \mathbf{H}^q \mathbf{B}, \quad (12)$$

where \mathbf{H}^s and \mathbf{H}^h , are the blocks corresponding to second order derivatives of energy with respect to fractional coordinates and lattice vectors components, respectively, and \mathbf{H}^{sh} is the block with mixed second order derivatives with respect to components of s and \mathbf{h} . Next, either the whole blocks \mathbf{H}^s and \mathbf{H}^{sh} or their rows and columns involving coordinates of 'active' atoms (a), i.e., those with presumably significant contribution to reaction coordinate, are replaced by the corresponding terms computed from the second order derivatives with respect to Cartesian coordinates determined at the DFT level (e.g., via finite differences):

$$\frac{\partial E}{\partial s_\alpha^a \partial s_\beta^b} = \sum_\mu \sum_\nu \frac{\partial E}{\partial r_\mu^a \partial r_\nu^b} h_{\mu\alpha} h_{\nu\beta}, \quad (13)$$

and

$$\frac{\partial E}{\partial h_{\alpha\beta} \partial s_\gamma^a} = \sum_\mu \sum_\nu \frac{\partial \sigma_{\alpha\mu}}{r_\nu^a} (h)_{\beta\mu}^{-1} h_{\nu\gamma}. \quad (14)$$

We note that the idea of treating contributions of active and inactive atoms to the Hesse matrix at different levels of theory has been reinvented by several different authors [24–27]. It is only in special cases of processes with reaction coordinate determined solely by the lattice vectors components (such as the martensitic transformation discussed in Sect. 3.6), when the components of block \mathbf{H}^h :

$$\frac{\partial E}{\partial h_{\alpha\beta} \partial h_{\gamma\delta}} = \sum_\mu \frac{\partial \sigma_{\alpha\mu}}{h_{\gamma\delta}} (h)_{\beta\mu}^{-1} \quad (15)$$

need to be determined accurately. In all other cases discussed in this work, the block \mathbf{H}^h is approximated by the simple model described above.

The correct eigenvalue spectrum of the Hesse matrix is then ensured by its spectral decomposition and replacing the eigenvalues (λ_i) by their absolute values for all modes (\mathbf{u}_i) except of the one representing the unstable mode (\mathbf{u}_1) that is approximately parallel with the reaction coordinate:

$$\mathbf{H} \rightarrow \lambda_1 \mathbf{u}_1 \mathbf{u}_1^t + \sum_{i=2}^{3N_{\text{at}}+9} |\lambda_i| \mathbf{u}_i \mathbf{u}_i^t. \quad (16)$$

In the cases where the initial \mathbf{H} matrix has multiple negative eigenvalues, the appropriate unstable mode is selected by a visual inspection of the corresponding vibrational modes. Finally, the matrix $\mathbf{H}^{\tilde{q}}$ used in the relaxation in delocalized internal coordinates is obtained by the following transformation:

$$\mathbf{H}^{\tilde{q}} = \mathbf{U}^t \mathbf{A}^t \begin{pmatrix} \mathbf{H}^s & \mathbf{H}^{sh} \\ (\mathbf{H}^{sh})^t & \mathbf{H}^h \end{pmatrix} \mathbf{A} \mathbf{U}. \quad (17)$$

The procedure of Eq. 16 is applied after every \mathbf{H} update performed during the relaxation. We note that a similar treatment has been used also in the previous work [25, 27].

3 Numerical tests

3.1 Simulation details

The energy and force calculations have been performed using the periodic DFT code VASP [28–31]. The Kohn–Sham equations have been solved variationally in a plane-wave basis set using the projector-augmented-wave (PAW) method of Blöchl [32], as adapted by Kresse and Joubert [33]. The PBE exchange–correlation functional in the generalized gradient approximation proposed by Perdew et al. [34] was used. For the purposes of our tests, high-quality forces were essential. Hence the calculations have been carried out with dense FFT grids (set automatically via the input parameter PREC=Accurate), and projection operators have been evaluated in reciprocal space. In order to reduce the Pulay stress possibly biasing the cell geometry optimizations, large plane-wave cutoffs have been used in calculations of periodic systems, see Table 1. In each self-consistent field cycle, the electronic wave function was converged to 1×10^{-7} eV/cell. In order to reduce the effect of periodic boundary conditions in relaxations of molecular systems (see Sect. 3.2), the size of each unit cell has been chosen such as to ensure that the minimal distance between any two atoms located each at a different periodic image of the molecule was at least 10 Å.

Table 1 Summary of important simulation parameters used in this study

| System | k -Point mesh | Plane-wave cutoff (eV) |
|-------------------------------------|--------------------------|------------------------|
| Molecular systems | $1 \times 1 \times 1$ | 400 |
| 1D H ₂ -chain | $50 \times 1 \times 1$ | 1000 |
| H transfer in CHAB | $2 \times 2 \times 2$ | 1000 |
| Crotonaldehyde desorption from MgO | $1 \times 1 \times 1$ | 1000 |
| Pure affine shear deformation of Al | $31 \times 25 \times 15$ | 600 |

The geometry optimizations were carried out using the external optimizer GADGET [11], which reads the geometry, energy, and forces from the VASP output, and, until convergence, sets up internal coordinates, estimates an optimal move, calculates the new set of lattice parameters and fractional coordinates, and starts a new VASP calculation. In our tests, the P-RFO method [3] has been used but a similar performance has been achieved also by using the geometrical DIIS method [4] or a quasi-Newton method [2]. The maximal size of optimization step performed in primitive internal coordinate space (i.e., bonds, angles, and torsions) has been limited to 0.1 a.u. The optimizations have been considered to be converged when the maximal Cartesian force acting on each atom and, in the case of the cell geometry optimizations, also the maximal force acting on the lattice vectors, was smaller than 5×10^{-3} eV/Å. Except of the test described in Sect. 3.6, two types of initial Hesse matrices have been used in simulations, both constructed as described in Sect. 2.2. In 'exact' initial Hessian, all elements of blocks H^s and H^{sh} were computed at the DFT level via finite differences, while only rows and columns corresponding to the elements related to a relatively small number of active atoms (i.e., those with presumably significant contribution to RC) were determined at the DFT level in 'approximate' H . The input and selected output files from all calculations presented in this work are provided as electronic supplementary material (see Online Resource 1).

3.2 Molecular systems

In this section, we compare the performance of exact and approximate initial model Hesse matrices (see Sect. 3.1) in relaxation of atomic positions at fixed cell geometry in transition states of selected gas-phase reactions from the benchmark set of Baker and Chan [13]. Clearly, such a test

is meaningful only for sufficiently large systems, where a significant fraction of atoms with negligible contribution to RC can be identified. For this reason, only systems consisting of ten or more atoms have been considered, whereby the number of active atoms ranged from 2 to 8 (see Table 2). The active atoms have been selected on the basis of visual inspection of unstable vibrational modes.

The numbers of optimization steps (i.e., gradient calculations) needed to fulfil the relaxation criterion are compiled in Table 2. We note that relaxations performed for the same system using different methods converged to the same states with the energies that were identical within 0.1 meV (see the data in Online Resource 1). As evident, the optimization is strongly affected by the quality of initial H and the number of steps needed to achieve convergence in all seven systems is almost doubled (257 vs. 130) when exact H is replaced by the approximate Hessian. Importantly, however, the number of gradient evaluations needed to construct the approximate H is significantly smaller than that for the exact Hessian and this fact should be also considered when comparing the computational cost of these two sets of simulations. Taking all gradient evaluations into account, it is evident that the optimizations with the approximate H (total of 349 gradient evaluations) are actually slightly more efficient than those employing the exact Hesse matrix (373 evaluations). One can expect that the use of approximate H will be even more effective in optimizations of large systems where the ratio of active to inactive atoms is small. This should hold true especially in the cases where the initial configuration is very different from the relaxed structure and hence the Hessian can be expected to change significantly in the course of optimization. Our results presented in Sect. 3.4 and 3.5 support this expectation.

In order to illustrate the efficiency of the optimization algorithm used in this work, additional calculations have

Table 2 The number of steps needed to relax the atomic positions in transition states of selected reactions from the benchmark set of Baker and Chan [13] using the method employed in this work (P-RFO) in combination with exact (exact) and approximate (approx.) initial Hesse matrices. The numbers in parentheses indicate the total number of gradient calculations performed in optimizations and

Hesse matrix calculations. The number of atoms (N_{tot}) in each system and the number of active atoms (N_{act}) used to compute the approximate initial Hesse matrix are also indicated. For sake of comparison, the number of relaxation steps needed to achieve convergence using the improved dimer method [5, 35] (IDM) is also shown

| Reaction | Relaxation steps | | | N_{tot} | N_{act} |
|--|------------------|-----------------|---------------|------------------|------------------|
| | P-RFO (exact) | P-RFO (approx.) | IDM | | |
| Ring opening bicyclo[1.1.0.]butane TS 1 | 11 (41) | 48 (60) | 277 (307) | 10 | 4 |
| Ring opening bicyclo[1.1.0.]butane TS 2 | 44 (74) | 55 (70) | Failed | 10 | 5 |
| 1,2-migration -(formyloxy) ethyl | 9 (39) | 29 (35) | 344 (374) | 10 | 2 |
| Butadiene + ethylene \leftrightarrow cyclohexene | 30 (78) | 40 (58) | 645 (693) | 16 | 6 |
| Trans-butadiene \leftrightarrow cis-butadiene | 14 (44) | 19 (42) | Failed | 10 | 8 |
| CH ₂ CHCH ₂ -O-CHCH ₂ \leftrightarrow CH ₂ CHCH ₂ CH ₂ CHO | 11 (53) | 26 (38) | 449 (491) | 14 | 4 |
| SiH ₂ + CH ₃ CH ₃ \leftrightarrow SiH ₃ CH ₂ CH ₃ | 11 (44) | 40 (46) | 1008 (1041) | 11 | 2 |
| Total | 130 (373) | 257 (349) | > 2723 (2906) | | |

been performed using the improved dimer method (IDM) [5, 35], which is one of the standard TS optimization methods implemented in VASP. A default computational setting for the IDM calculations has been used and the initial dimer vector was defined as the unstable eigenvector of the exact Hesse matrix. Unlike the P-RFO method, the IDM failed to converge in two out of seven systems (see Table 2). The performance of IDM in optimization of remaining systems (as measured in number of steps needed to achieve convergence) is significantly worse than that of the P-RFO combined with the delocalized internal coordinates. This result is, however, not surprising as the dimer method in its present implementation does not make use of the full information encoded in the Hesse matrix, and a large part of steps performed by the IDM corresponds to auxiliary calculations, such as the determination of curvature along the unstable direction or rotation of the dimer axis, rather than to actual optimization.

3.3 1D H₂-chain

As a first example of TS optimization of periodic systems, a simple reaction of one-dimensional chain of H₂ molecules is considered, in which the atoms are reconnected as shown in Fig. 1. The unit cell contains only two atoms and the system can be fully described by just two geometric parameters, e.g., the difference in distances between an arbitrary H atom and its two neighbors ($r_1 - r_2$), and the length of the lattice vector parallel with the chain (a) shown in Fig. 1. As dictated by symmetry, the structure of TS corresponds to an infinite chain of H atoms with equidistant separation between nearest neighbours (i.e., $r_1 - r_2 = 0$).

The geometry of simulation cell in the initial structure was given by the lattice parameters $a = 2.7$ Å, $b = 10.0$ Å, $c = 10.0$ Å, $\alpha = \beta = \gamma = 90^\circ$, and the value of the coordinate $r_1 - r_2$ was set to 0.54 Å. The geometry was allowed to relax only in direction parallel with the lattice vector a , while the lattice vectors b and c as well as the corresponding fractional coordinates of atoms were fixed.

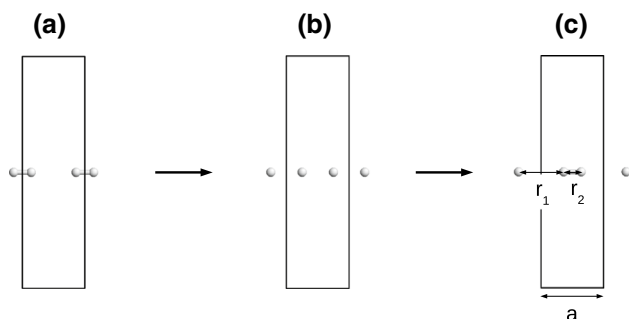


Fig. 1 Initial (a), transition (b), and final (c) states of reconnection reaction of atoms in one-dimensional H₂ chain. The solid lines represent the unit cell used in calculations

Maximal Cartesian component of gradient computed for the initial structure was 3.94 eV/Å and the internal pressure of 0.9 GPa indicated that the initial structure was expanded with respect to the equilibrium structure. Owing to the dimensionality of this system, only relaxation with exact initial H matrix has been performed (see Sect. 3.1). The geometry optimization of atomic positions at fixed cell geometry converged in 9 steps and the correct TS geometry (i.e., $r_1 - r_2 = 0$ Å) has been obtained. We note that the relatively large number of optimization steps in this truly one-dimensional problem was caused partly by the step limitation (see Sect. 3.1) consistently applied in all calculations discussed in this work. The simultaneous relaxation of atomic positions and lattice parameter a converged in 16 steps and the geometry with $r_1 - r_2 = 0$ Å and $a = 1.977$ Å has been obtained. As the value of $r_1 - r_2$ in TS follows from symmetry, the predicted TS geometry can be easily checked by a set of single-point calculations with varied value of a and fixed $r_1 - r_2 = 0$ Å. As shown in Fig. 2, the position of relaxed TS fits the minimum on the energy versus a plot perfectly, as it should.

3.4 Proton transfer in chabazite

In the next example, we discuss the transition state optimization of an unconstrained 3D system. The reaction that we consider is a proton shift between two oxygen atoms next to aluminum in zeolite chabazite (see Fig. 3). This reaction has been studied theoretically [36, 37] in the context of heterogeneous catalysis. The unit cell with 37 atoms and the following initial values of lattice parameters has been used in calculations: $a = b = c = 9.291$ Å, $\alpha = \beta = \gamma = 93.9^\circ$, $V = 796.1$ Å³. We note that the rhombohedral symmetry of cell is valid for highly siliceous chabazite or for a structure with a random distribution of Al and H atoms [38] and the use of a relatively small simulation cell creating an artificial

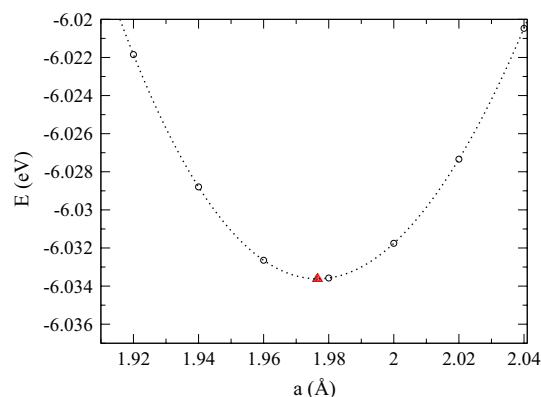


Fig. 2 Energy of the one-dimensional chain of H atoms with $r_1 - r_2 = 0$ Å as a function of the cell parameter a (see Fig. 1b). Red triangle represents the result of TS optimization with variable a

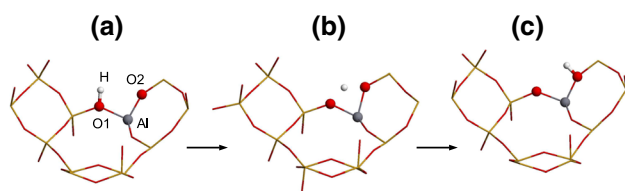


Fig. 3 Initial (a), transition (b), and final (c) states of the proton transfer in acid chabazite. Atoms H, O1, and O2 were considered as active in construction of approximate initial Hesse matrix

regular pattern in the Al and H positions can be expected to break this symmetry. The maximum component of Cartesian gradient determined for the initial structure was as large as 7.71 eV/\AA and the internal pressure was negative (-1.9 GPa) indicating that this structure was compressed.

The relaxation of atomic degrees of freedom at fixed lattice geometry converged in 88 steps when the exact initial Hessian was used. With the same \mathbf{H} , the full relaxation of atomic and lattice degrees of freedom converged in 67 steps, yielding the structure with the following lattice parameters: $a = 9.367 \text{ \AA}$, $b = 9.480 \text{ \AA}$, $c = 9.206 \text{ \AA}$, $\alpha = 92.0^\circ$, $\beta = 93.2^\circ$, $\gamma = 93.8^\circ$, $V = 813.8 \text{ \AA}^3$. Despite the relatively large expansion and deformation of lattice in the relaxation, the internal degrees of freedom contributing to the RC most significantly, namely the distances H-O1 and H-O2 (see Fig. 3), differ only by $< 0.01 \text{ \AA}$ from the values determined in relaxation of atomic positions only. This result is a consequence of the fact that this TS geometry is primarily determined by strong Coulomb interactions of H^+ with the neighboring O atoms possessing a partial negative charge. The stabilization of TS due to the lattice relaxation is relatively modest (0.153 eV/cell).

In the next test, we examine the performance of the approximate Hessian (see Sect. 3.1) whereby only the atoms H, O1, and O2 (see Fig. 3) are considered as active. Both the atomic only and the full relaxations performed with the approximate initial \mathbf{H} converged to the same structures as the relaxations performed with the exact initial \mathbf{H} . The number of relaxation steps needed to achieve convergence increased significantly (132 steps for the relaxation of atoms and 98 steps for the full relaxation) but taking into account the minimal number of gradient evaluations needed for the \mathbf{H} calculation (111 for the exact and only 9 for the approximate Hessian), the relaxation with the approximate \mathbf{H} turns out to be more efficient than that employing the exact initial Hessian.

3.5 Partial desorption of crotonaldehyde from the MgO surface

In this section, we consider partial desorption of crotonaldehyde from the MgO surface, which is a part of a catalytic

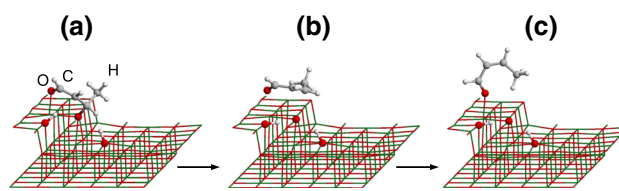


Fig. 4 Initial (a), transition (b), and final (c) states of partial desorption of crotonaldehyde from the MgO surface with a stepped kink. The atoms represented by spheres were considered as active in construction of approximate initial Hesse matrix

transformation of ethanol to 1,3-butadiene, recently studied by Taifan et al. [39]. As shown in Fig. 4, the reaction consists of rotation of a relatively large molecular fragment with respect to the substrate (see also Online Resource 2 showing animation of the full transformation path). As it is typical for this kind of motion, the imaginary frequency of the unstable vibrational mode is relatively soft (173 i cm^{-1} in the initial and 93 i cm^{-1} in the final structures, respectively) making the determination of this TS rather challenging. Following Ref. [39], the initial geometry of simulation cell was defined by the lattice parameters $a = 12.746 \text{ \AA}$, $b = 17.020 \text{ \AA}$, $c = 19.255 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$, $V = 4183.1 \text{ \AA}^3$, and the fractional coordinates of atoms of the bottommost layer were fixed at their bulk values in all simulations discussed here, i.e., 82 out of the total of 130 atoms were allowed to relax. In relaxations involving the lattice components, additional constraints fixing the length of the lattice vector \mathbf{c} , which is perpendicular to the slab, and all lattice angles were introduced as described in Sect. 2.1. The maximal Cartesian component of gradient computed for the initial structure was 2.22 eV/\AA and the stress-tensor components σ_{xx} , and σ_{yy} (supposed to be relaxed out in the lattice optimization) were -2.5 GPa , and -3.7 GPa , respectively.

The relaxations of atomic positions performed with the exact initial Hessian converged in 50 steps. The stress-tensor components σ_{xx} and σ_{yy} determined for the final structure (-2.4 GPa and -3.7 GPa , respectively) changed only little compared to those computed for the initial configuration and their numerical values indicate that the structure is significantly stretched in directions parallel with the surface plane. Indeed, the lengths of lattice vectors \mathbf{a} and \mathbf{b} shrank in the full relaxation to 12.426 \AA and 16.314 \AA , respectively, decreasing thus the energy by 3.018 eV/cell . Due to the relatively large cell geometry relaxation, the full optimization required significantly larger number of steps (84) than the optimization of atomic positions only. As mentioned above, the reaction involves rotation of whole crotonaldehyde molecule with respect to the MgO surface, hence all 11 atoms forming the molecule plus 4 substrate atoms have been considered as active (displayed as spheres in Fig. 4) when constructing the approximate \mathbf{H} matrix for the initial structure.

With this approximate \mathbf{H} , the atomic only and the full relaxations converged in 95 and 193 steps, respectively, yielding the same structures as those obtained using the exact Hessian. When the minimal number of gradient evaluations needed for the \mathbf{H} calculation is taken into account (246 for the exact and 45 for the approximate Hessian), the relaxation with the approximate \mathbf{H} is, again, found to be more efficient than that employing the exact initial Hessian.

3.6 Pure affine shear deformation of Al

As our final example, we discuss a transformation taking place entirely in the space of lattice vectors components $h_{i,\mu}$, namely the $\langle 11\bar{2} \rangle \{111\}$ pure affine shear deformation of Al, previously studied by Jahnátek et al. [40]. As shown in Fig. 5, this transformation is realized via sliding the Al layers parallel with the (111) plane and the final structure is symmetry equivalent to the initial one. The deformation causes a variation of the lattice degrees of freedom while the fractional coordinates of atoms remain intact because the corresponding net forces are zero by symmetry. As this process is reversible (implying that the transformation path must be continuous), the maximum energy structure can be found by using our TS optimization technique. Alternatively, the same structure can be identified as the configuration corresponding to the maximum on the energy versus α plot obtained in a series of constrained relaxations, where α stands for the angle between two lattice vectors parallel with the $[11\bar{2}]$ and the $[111]$ directions defined with respect to the conventional body-centered cubic cell of Al [40].

Our simulation setup was similar to that reported in Ref. [40], albeit the plane-wave cutoff was increased in order to reduce the biasing effect of Pulay stress in the full TS relaxation. A minimal unit cell with lattice vectors parallel with the $[11\bar{2}]$, $[1\bar{1}0]$, and $[111]$ directions has been chosen containing six Al atoms (i.e., three independent Al layers, see Fig. 5). The following cell geometry has been used in the initial guess for the maximum energy structure: $a = 2.854 \text{ \AA}$,

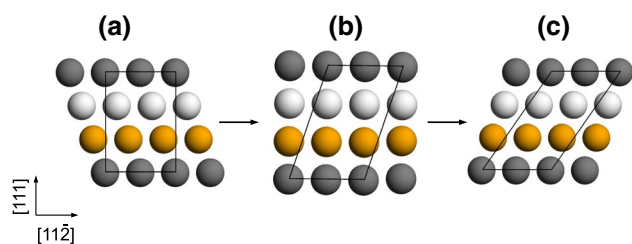


Fig. 5 Initial (a), transition (b), and final (c) states of the transformation of Al under the $\langle 11\bar{2} \rangle \{111\}$ pure affine shear deformation. Full lines represent the unit cell employed in calculations and different colors are used to distinguish different layers of Al atoms parallel with the (111) plane. Note that the structures **a** and **c** are equivalent by symmetry

$b = 4.949 \text{ \AA}$, $c = 7.282 \text{ \AA}$, $\alpha = 74.1^\circ$, $\beta = 90^\circ$, $\gamma = 90^\circ$, $V = 98.9 \text{ \AA}^3$. The components of the initial Hesse matrix corresponding to the lattice degrees of freedom (i.e., the block \mathbf{H}^h) have been determined numerically at the DFT level while all remaining blocks were approximated by the model described in Sect. 2.2. The TS geometry converged in just nine steps and the structure with the following cell geometry has been obtained: $a = 2.809 \text{ \AA}$, $b = 4.924 \text{ \AA}$, $c = 7.781 \text{ \AA}$, $\alpha = 71.6^\circ$, $\beta = 90.0^\circ$, $\gamma = 90.0^\circ$, $V = 102.1 \text{ \AA}^3$. The internal pressure of 2.3 GPa determined for the initial structure has been eliminated entirely and all stress-tensor components vanished upon relaxation. The computed structure is perfectly consistent with that corresponding to the maximum on the energy vs. α profile (see Fig. 6) determined in a series of constrained relaxations.

4 Conclusions

In this work, our algorithm [11] for structural optimizations of periodic systems in delocalized internal coordinates [12] has been adapted for the use in relaxations of transition states. The presented method allows for simultaneous relaxations of atomic positions and cell geometry with or without additional geometric constraints. The performance of the method has been demonstrated on several different real-world examples covering the most important cases occurring in practice. In particular, a full unconstrained TS relaxation of a system with three-dimensional periodicity has been demonstrated on example of the proton exchange reaction in zeolite chabazite. As examples of relaxations with additional geometric constraints including fractional coordinates

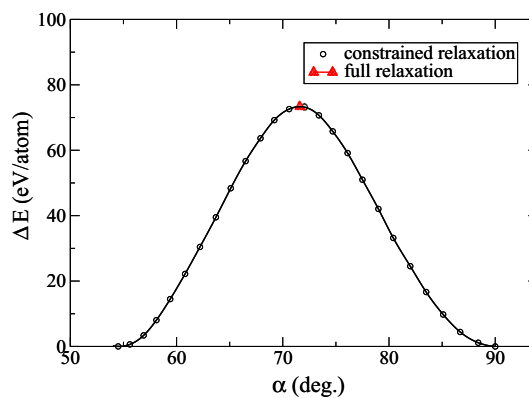


Fig. 6 Change in energy of Al with variation of angle (α) between lattice vectors parallel with the $[11\bar{2}]$ and the $[111]$ directions defined with respect to the conventional body-centered cubic cell of Al (cf. Fig. 5). The result of the transition state optimization (full relaxation) is shown to coincide with maximum of the curve obtained from a series of constrained relaxations with fixed α

of selected atoms and lattice degrees of freedom (such as the lengths of lattice vectors and the lattice angles), the TS optimizations of one-dimensional chain of H₂ molecules and the partial desorption of crotonaldehyde from MgO surface have been discussed. Finally, a relaxation of transition state defined entirely in the space of lattice vectors components has been demonstrated on example of the pure affine shear deformation of Al. The full TS relaxation involving both the atomic positions and the lattice vectors components was found to represent a relatively modest overhead compared to usual TS relaxations involving atomic positions only. The performance of the TS relaxations initialized using an approximate Hesse matrix was examined, in which a relatively small fraction of matrix elements corresponding to active atoms directly participating in the reaction of interest were determined accurately at the DFT level, while a major part of elements, typically related to inactive atoms and lattice vectors components, was defined on the basis of a simple empirical model. When a total number of gradient evaluations used in relaxations and Hessian calculations was taken into account, the TS geometry optimizations initialized with the approximate Hesse matrix was found to outperform the simulations carried out with exact initial Hessian, in which all elements related to atomic positions were computed at the DFT level.

Acknowledgements TB is deeply indebted to Dr. János Gábor Ángyán for introducing him into the subject discussed in this work and for a long, fruitful, and always very pleasant collaboration on 'anything that was interesting'. This work was supported by the Slovak Research and Development Agency under the Contract No. APVV-15-0105. Calculations were performed using the supercomputing infrastructure of Computing Center of the Slovak Academy of Sciences acquired in projects ITMS 26230120002 and 26210120002 supported by the Research and Development Operational Program funded by the ERDF.

References

- Jensen F (1997) Introduction to computational chemistry. Wiley, Chichester
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical recipes: the art of scientific computing, 3rd edn. Cambridge University Press, New York, NY
- Banerjee A, Adams N, Simons J, Shepard R (1985) Search for stationary-points on surface. *J Phys Chem* 89(1):52–57
- Császár P, Pulay P (1984) Geometry optimization by direct inversion in the iterative subspace. *J Mol Struct* 114:31–34
- Henkelman G, Jónsson H (1999) A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J Chem Phys* 111(15):7010
- Henkelman G, Jónsson H (2000) Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J Chem Phys* 113(22):9978–9985
- Henkelman G, Jóhannesson G, H J (2005) Methods for finding saddle points and minimum energy paths. In: Schwartz SD (ed) *Theoretical methods in condensed phase chemistry*. Progress in theoretical chemistry and physics, Springer, Dordrecht, vol 5, pp 269–302
- Caspersen KJ, Carter EA (2005) Finding transition states for crystalline solid–solid phase transformations. *Proc Natl Acad Sci* 102(19):6738–6743. <https://doi.org/10.1073/pnas.0408127102>
- Sheppard D, Xiao P, Chemelewski W, Johnson DD, Henkelman G (2012) A generalized solid-state nudged elastic band method. *J Chem Phys* 136(7):074103
- Xiao P, Sheppard D, Rogal J, Henkelman G (2014) Solid-state dimer method for calculating solid–solid phase transitions. *J Chem Phys* 140(17):174104
- Bucko T, Hafner J, Angyan JG (2005) Geometry optimization of periodic systems using internal coordinates. *J Chem Phys* 122(12):124508
- Baker J, Kessi A, Delley B (1996) The generation and use of delocalized internal coordinates in geometry optimization. *J Chem Phys* 105(1):192–212
- Baker J, Chan FR (1996) The location of transition states: a comparison of cartesian, z-matrix, and natural internal coordinates. *J Comput Chem* 17(7):888–904
- Peng C, Ayala PY, Schlegel HB, Frisch MJ (1996) Using redundant internal coordinates to optimize equilibrium geometries and transition states. *J Comput Chem* 17(1):49–56
- Bakken V, Helgaker T (2002) The efficient optimization of molecular geometries using redundant internal coordinates. *J Chem Phys* 117(20):9160–9174
- Andzelm J, King-Smith RD, Fitzgerald G (2001) Geometry optimization of solids using delocalized internal coordinates. *Chem Phys Lett* 335(3):321–326
- Baker J, Pulay P (1996) Geometry optimization of atomic microclusters using inverse-power distance coordinates. *J Chem Phys* 105(24):11100–11107
- Wilson EBJ, Decius JC, Cross PC (1955) *Molecular vibrations. The theory of infrared and Raman vibrational spectra*. Dover, New York
- Bofill JM (1994) Updated Hessian matrix and the restricted step method for locating transition structures. *J Comput Chem* 15(1):1–11
- Murtagh BA, Sargent RWH (1972) *Comput J* 13:185
- Powell MJD (1971) *Math Prog* 1:26
- Lindh R, Bernhardsson A, Karlstrom G, Malmquist PA (1995) On the use of a hessian model function in molecular-geometry optimizations. *Chem Phys Lett* 241(4):423–428
- Fischer TH, Almlöf J (1992) General methods for geometry and wave function optimization. *J Phys Chem* 96(24):9768–9774
- Bučko T (2004) Structure, acidity, and chemical reactivity of zeolites. In: Thesis, University of Vienna, Vienna
- Bucko T, Hafner J, Benco L (2004) Active sites for the vapor phase beckmann rearrangement over mordenite: an ab initio study. *J Phys Chem A* 108(51):11388–11397
- Pu J, Truhlar DG (2005) Use of block Hessians for the optimization of molecular geometries. *J Chem Theory Comput* 1(1):54–60
- Rabi S (2014) Transition-state optimization methods using internal coordinates. In: Thesis, McMaster University
- Kresse G, Hafner J (1993) Ab-initio molecular-dynamics for open-shell transition-metals. *Phys Rev B* 48(17):13115–13118
- Kresse G, Hafner J (1994) Norm-conserving and ultrasoft pseudopotentials for first-row and transition-elements. *J Phys Condens Matter* 6(40):8245–8257
- Kresse G, Furthmüller J (1996) Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput Mater Sci* 6(1):15–50
- Kresse G, Furthmüller J (1996) Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys Rev B* 54(16):11169–11186

32. Blöchl PE (1994) Projector augmented-wave method. *Phys Rev B* 50:17953–17979
33. Kresse G, Joubert D (1999) From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys Rev B* 59(3):1758–1775
34. Perdew JP, Burke K, Ernzerhof M (1996) Generalized gradient approximation made simple. *Phys Rev Lett* 77:3865–3868
35. Heyden A, Bell A, Keil F (2005) Efficient methods for finding transition states in chemical reactions: comparison of improved dimer method and partitioned rational function optimization method. *J Chem Phys* 123(22):224101
36. Fermann JT, Auerbach S (2000) Modeling proton mobility in acidic zeolite clusters: II. Room temperature tunneling effects from semiclassical rate theory. *J Chem Phys* 112(15):6787–6794
37. Sierka M, Sauer J (2001) Proton mobility in chabazite, faujasite, and ZSM-5 zeolite catalysts. comparison based on ab initio calculations. *J Phys Chem B* 105(8):1603–1613
38. Smith LJ, Davidson A, Cheetham AK (1997) A neutron diffraction and infrared spectroscopy study of the acid form of the aluminosilicate zeolite, chabazite (H-SSZ-13). *Catal Lett* 49(3):143–146
39. Taifan WE, Bučko T, Baltrusaitis J (2017) Catalytic conversion of ethanol to 1,3-butadiene on MgO: a comprehensive mechanism elucidation using DFT calculations. *J Catal* 346:78–91
40. Jahnatek M, Hafner J, Krajci M (2009) Shear deformation, ideal strength, and stacking fault formation of FCC metals: a density-functional study of Al and Cu. *Phys Rev B* 79:224103