PCCP

View Online / Journal Homepage
Dynamic Article Links

Cite this: DOI: 10.1039/c2cp23421a

www.rsc.org/pccp

PAPER

A method for finding the ridge between saddle points applied to rare event rate estimates†

Jón Bergmann Maronsson,[a] Hannes Jónsson[b] and Tejs Vegge*[a]

Received 30th October 2011, Accepted 22nd December 2011
DOI: 10.1039/c2cp23421a

Downloaded by DTU Library on 23 January 2012
Published on 22 December 2011 on http://pubs.rsc.org | doi:10.1039/C2CP23421A

A method is presented for finding the ridge between first order saddle points on a multidimensional surface. For atomic scale systems, such saddle points on the energy surface correspond to atomic rearrangement mechanisms. Information about the ridge can be used to test the validity of the harmonic approximation to transition state theory, in particular to verify that second order saddle points—maxima along the ridge—are high enough compared to the first order saddle points. New minima along the ridge can also be identified during the path optimisation, thereby revealing additional transition mechanisms. The method is based on a string of discretisation points along a path between the first order saddle points and using an iterative optimisation which requires only the force acting on the atoms. At each iteration during the optimisation, the force is inverted along an unstable eigenmode perpendicular to the path. The method is applied to Al adatom diffusion on the Al(100) surface to find the ridge between 2-, 3- and 4-atom concerted displacements and hop mechanisms. A correction to the harmonic approximation of transition state theory was estimated by direct evaluation of the configuration integral along the ridge.

1. Introduction

The rate of rare events in a system coupled to a heat bath can be estimated by evaluating the free energy barriers for the transitions. Transition state theory (TST)[1–5] is the foundation for this approach. In atomic scale systems, for example, the rate of thermally activated rearrangements of the atoms, which, typically, are several orders of magnitude slower than atomic vibrations, can be estimated from the free energy of a dividing surface separating atomic configurations corresponding to the initial state from the atomic configurations corresponding to all final states. Due to the large difference in time scale between atomic vibrations and typical thermally induced processes such as chemical reactions or diffusion, it would require immense computational power to directly simulate dynamical trajectories sufficiently long to include these rare events. TST makes it possible to focus on the rare events and neglect the details of the fast vibrational motion. It provides an approximation which eliminates the time scale problem, while dynamical information and an exact rate constant can be obtained by applying dynamical corrections based on trajectories started at the dividing surface.[6] The greatest challenge in the implementation of TST is to determine and represent a good dividing surface. The better the choice of the dividing surface, the better the estimate of the rate will be (TST gives a variational upper bound) and easier to subsequently evaluate the dynamical corrections.

Given a dividing surface (transition state), ‡, the reaction rate, k^{TST}, out of the initial state, I, can be calculated as the thermally averaged probability of being at the transition state times the average velocity, $v_\perp$, perpendicular to the dividing surface in the direction away from the initial state

$$k^{\mathrm{TST}} = \tfrac{1}{2}\langle |v_\perp| \delta(\mathbf{R} - \mathbf{R}^{\ddagger})\rangle. \tag{1}$$

A hyperplane is a commonly used choice for a dividing surface. The expression for the rate constant then becomes

$$k^{\mathrm{TST}} = \sqrt{\frac{k_{\mathrm{B}}T}{2\pi\mu}}\frac{Z_{\ddagger}}{Z_I}, \tag{2}$$

$$Z_S = \int_S \mathrm{e}^{-E(\mathbf{R})/k_{\mathrm{B}}T}\,\mathrm{d}\mathbf{R}. \tag{3}$$

where Z_S denotes a configuration integral over a subspace S in configuration space. Here, $E(\mathbf{R})$ is the energy of configuration $\mathbf{R}$, k_{B} is the Boltzmann constant, T the temperature and μ the effective mass for motion along the hyperplane normal.

The task of finding and representing a transition state is simplified greatly in the harmonic approximation to TST[7,8] (HTST),

[a] *Department of Energy Conversion and Storage, Technical University of Denmark, Risø Campus, 4000 Roskilde, Denmark. E-mail: jber@risoe.dtu.dk, teve@risoe.dtu.dk*
[b] *Science Institute and Faculty of Science, VR-III, University of Iceland, 107 Reykjavík, Iceland. E-mail: hj@hi.is*
† Electronic supplementary information (ESI) available. See DOI: 10.1039/c2cp23421a

Published on 22 December 2011 on http://pubs.rsc.org | doi:10.1039/C2CP23421A
Downloaded by DTU Library on 23 January 2012

where the transition state is represented as a collection of hyperplanar segments placed at first order saddle points (SP_1s) on the potential energy rim surrounding the initial state and having a normal pointing along the unstable vibrational mode (see, for example ref. 9). The potential energy function is then approximated by a second order Taylor expansion around the SP_1s for the transition state and around the potential energy minimum at the initial state. Each saddle point represents a particular transition mechanism and the energy of the saddle point minus the energy of the initial state minimum gives the corresponding energy barrier, $E_b = E^{\ddagger} - E^{I}$. The expression for the reaction rate constant at this level of approximation becomes

$$k^{\mathrm{HTST}} = \frac{\prod_{i}^{3N} \nu_i^{\mathrm{I}}}{\prod_{i}^{3N-1} \nu_i^{\ddagger}} \mathrm{e}^{-E_b/k_B T}, \tag{4}$$

where N is the number of atoms in the system, $\nu^{\ddagger}$ denotes frequency of vibrational normal modes at the saddle point and ν^{I} at the initial state minimum. HTST is much less computationally demanding than full TST, typically by several orders of magnitude and is therefore frequently used, especially for transitions in solids. It, however, does not apply to soft matter systems where a multitude of saddle points with low energy tend to be present. It is important to identify under what conditions HTST is accurate enough and when full TST is needed.

Finding a SP_1 is challenging especially if the transition mechanism is not known *a priori*. Several methods have been developed for this task, which can be divided into two categories: those where knowledge about only the initial state is input and those where knowledge about both the initial and a final state is input. An example of the first category are minimum mode following methods[10,11] which involves inverting the force component along the eigenmode corresponding to the lowest eigenvalue of the Hessian matrix to locally transform SP_1s into minima. The latter involves finding a minimum energy path (MEP) between the two states, a path for which the component of the gradient in directions perpendicular to the path is zero and corresponding eigenvalues are positive. Any maximum along a MEP is a SP_1 and the highest one gives the best estimate of the energy barrier that needs to be overcome. The nudged elastic band (NEB) method[12–14] is frequently used in this context, where an initial path is discretised by images of the system which are then optimised until each image has converged onto a path with zero perpendicular gradient. Frequently, a NEB calculation finds additional minima along the MEP, thereby revealing possible states of the system that may not have been known beforehand, see for example ref. 15. A review of several SP_1 finding algorithms is presented in ref. 16.

Both methods mentioned above only require the evaluation of the first derivative of the energy surface, even though they are used to find points that are identified by the curvature, *i.e.* properties related to the second derivatives. It is desirable to avoid having to evaluate second derivatives because they usually require a large computational effort. In plane wave based density functional theory calculations, for example, the first derivative of the energy can be obtained without much extra computational effort beyond the potential energy, while the evaluation of the second derivatives, and thus the Hessian matrix, is much more difficult and costly.

As mentioned above, the harmonic approximation only applies to certain types of systems. The multi-dimensional parabolas need to be accurate enough approximations to the energy surface in the statistically relevant regions near the minimum and near the SP_1 at the temperature of interest. In particular, HTST only applies if the SP_1 is sufficiently higher in energy than the initial state (a commonly used rule of thumb is $E_b > 5k_BT$),[5] and sufficiently lower than surrounding second order saddle points, SP_2s (at an Nth order saddle point the Hessian matrix has N negative eigenvalues). Checking these criteria is non-trivial but important. If they are not met, the HTST estimate of the rate may not be satisfactory.

The presence of low energy SP_2s or irregularities in the potential energy ridge near SP_1s can be seen as warning signs that the harmonic approach is inadequate. Furthermore, a rough correction factor for the reaction rate can be evaluated by comparing the configuration integral, $Z_{\ddagger}$, with and without the harmonic approximation.

Characterizing potential energy ridges can also be useful when ensuring that all relevant SP_1s have been found, as the ridge between two SP_1s that are not adjacent on the ridge will go through any intermediate SP_1. This can be helpful when analysing complex systems where it is hard to predict the most efficient transition mechanism.

Finding SP_2s and energy ridges is inherently more difficult than finding SP_1s as maxima along a MEPs. As more constraints need to be applied during the optimisation, several new challenges emerge, as discussed below.

Most rigorously, the energy ridge can be identified by running steepest descent trajectories and finding the boundary between starting points in configuration space that converge to different local minima.[17] Such a procedure is, however, far too computationally tedious for all but the simplest, low dimensional systems.

In the vicinity of a SP_1, the Hessian matrix has one negative eigenvalue, λ, and at any point on the ridge, the gradient of the energy has no component in the direction of the corresponding eigenvector, $\hat{\boldsymbol{\varepsilon}}$,[18]

$$\hat{\boldsymbol{\varepsilon}} \cdot \nabla E = 0 \text{ and } \lambda < 0. \tag{5}$$

As discussed below, this condition does not hold near a SP_2, where the Hessian matrix has two negative eigenvalues.

In this article, we present a method for estimating the location of an energy ridge between two SP_1s, going through one or more SP_2s. It combines elements from the two categories of algorithms for finding SP_1 mentioned above. After presenting the method in the next section, we describe an application to a study of the diffusion of an Al adatom on an Al(100) surface. The results show that some of the low energy processes are not well described by the harmonic approximation even at rather low temperatures and a correction factor for HTST is estimated by evaluating the configuration integral along the ridge.

2. Methodology: finding the ridge

Given an energy surface, $E(\mathbf{R})$, its gradient, $\nabla E(\mathbf{R})$, and two SP_1s, the goal is to identify a path that lies close to the energy

 This journal is © the Owner Societies 2012

Downloaded by DTU Library on 23 January 2012
Published on 22 December 2011 on http://pubs.rsc.org | doi:10.1039/C2CP23421A

ridge between the two SP_1s. The path should, in particular, lie through any intermediate SP_2 so that a comparison of the height of SP_2s with respect to the SP_1s can be made. The method should, furthermore, lead to the identification of previously unknown SP_1(s) on the ridge in between the given end points, should they exist.

The path is at each point characterised by its tangent, $\hat{\boldsymbol{\tau}}$. When the path lies along the ridge, the component of the gradient or, equivalently, the force, $\mathbf{F} \equiv -\nabla E(\mathbf{R})$, that is perpendicular to the tangent,

$$\mathbf{F}^{\perp} \equiv \mathbf{F} - (\mathbf{F}\cdot\hat{\boldsymbol{\tau}})\hat{\boldsymbol{\tau}}, \tag{6}$$

must vanish at each point along the path,

$$\mathbf{F}^{\perp}_{\text{ridge}} = \mathbf{0}. \tag{7}$$

For a given estimate of the path, the optimisation task involves iteratively adjusting its shape and location until $\mathbf{F}^{\perp}$ vanishes, but in order for the path to lie along a ridge rather than a MEP, it is necessary that the energy has a maximum at the path along a perpendicular direction. Close to a SP_1, this direction is given by the eigenmode corresponding to the smallest eigenvalue of the Hessian matrix as illustrated in Fig. 1. Close to a SP_2, however, the Hessian matrix has two negative eigenvalues and the eigenmode corresponding to the lower one may be parallel to the tangent of the ridge and the direction at which a point on the ridge is a maximum corresponding to the second lowest eigenvalue. This is also illustrated in Fig. 1. There can, furthermore, be regions where neither of the two eigenmodes corresponding to negative eigenvalues are perpendicular to the tangent of the ridge as seen in Fig. 1. The latter regions can pose significant stability issues during the optimisation if not properly constrained.

The reduced Hessian matrix for the subspace excluding the tangent vector has at each point on the ridge one and only one

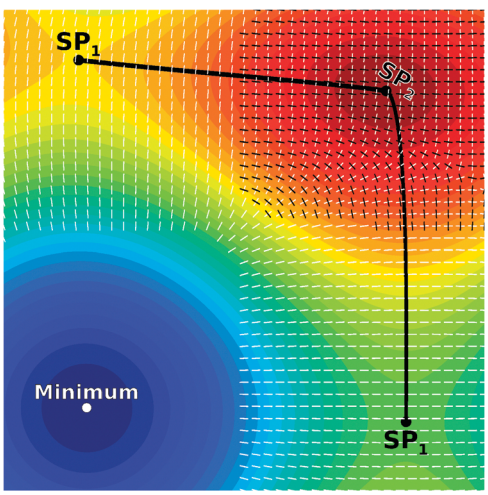

Fig. 1 A schematic, two-dimensional, energy surface illustrating an energy ridge between two SP_1s through a SP_2. The ridge is shown with a black line, which is the dividing surface between starting points of steepest descent paths that lead to different minima (only one of the minima is shown). The directions of eigenmodes corresponding to negative eigenvalues of the Hessian matrix are shown with short line segments, white indicating the one corresponding to the lower eigenvalue and the black corresponding to the higher one.

negative eigenvalue whose corresponding eigenmode, $\hat{\boldsymbol{\varepsilon}}$, is necessarily orthogonal to the ridge. The ridge can be located by maximizing the energy along this direction while minimizing in all other directions perpendicular to the ridge. By transforming the force, $\mathbf{F}^{\perp}$, in such a way that it locally, near the ridge, corresponds to that of a MEP

$$\mathbf{F}^{\text{t}} = \mathbf{F}^{\perp} - 2(\mathbf{F}^{\perp}\cdot\hat{\boldsymbol{\varepsilon}})\hat{\boldsymbol{\varepsilon}}, \tag{8}$$

an iterative displacement of the path in the direction of $\mathbf{F}^{\text{t}}$, *i.e.* a minimisation, can be used to locate the ridge. This mapping of the force makes it possible to employ a method that is similar to a NEB search for a MEP, but by using the transformed force, $\mathbf{F}^{\text{t}}$, the path converges on a ridge instead.

A numerical implementation of the path optimisation requires the introduction of discretisation. The path is represented by a discrete set of configurations, a set of images of the system, with coordinates $[\mathbf{R}_0, \mathbf{R}_1, \ldots, \mathbf{R}_{N-1}, \mathbf{R}_N]$. The energy of each image, i, is $E_i \equiv E(\mathbf{R}_i)$. Since a discrete representation of the path is used, the path's tangent needs to be approximated at each image. In NEB calculations, it has been found to be important, for numerical stability, to use the vector displacement to the higher energy neighbouring image, so as to minimise the formation of kinks in the path.[14] This same tangent estimation was used in the calculations presented here.

In order to perform the force transformation described in eqn (8), it is necessary to be able to find the eigenmode associated with the lowest eigenvalue (hereafter referred to as the minimum mode) of the reduced Hessian matrix. For this purpose we use the dimer method[10,11] since it gives the minimum mode using only the force as input. It is also possible to use the Lanczos method for this purpose.[11] For clarity, we rewrite eqn (8) for each image, i

$$\mathbf{F}^{\text{t}}_i = \mathbf{F}^{\perp}_i - 2(\mathbf{F}^{\perp}_i\cdot\hat{\boldsymbol{\varepsilon}}_i)\hat{\boldsymbol{\varepsilon}}_i, \tag{9}$$

which is applied after the dimer has been rotated subject to a constraint $\hat{\boldsymbol{\varepsilon}}_i\cdot\hat{\boldsymbol{\tau}}_i = 0$, to find the minimum mode of the reduced Hessian matrix.

In order to ensure an even distribution of the images along the path, a spring force, $\mathbf{F}^{\text{S}}_i$, between adjacent images is introduced. The force exerted on image i by the springs is

$$\mathbf{F}^{\text{S}}_i = k[(\mathbf{R}_{i+1} - \mathbf{R}_i) - (\mathbf{R}_i - \mathbf{R}_{i-1})], \tag{10}$$

where k is the spring constant which can be chosen to fit the energy landscape. For numerical convergence, it is best to choose k in such a way that the spring force is of roughly the same magnitude as the force derived from the energy surface, but a wide range of values can be used. Here, the same value of k is used for all pairs of adjacent images, but unequal values can be chosen if an unequal distribution of the images along the ridge is desired, analogous to NEB calculations.[13]

The effective force, $\mathbf{F}^{\text{eff}}_i$, acting on each image can now be written as

$$\mathbf{F}^{\text{eff}}_i = \mathbf{F}^{\text{t}}_i + \mathbf{F}^{\text{S}}_i, \tag{11}$$

where the first term is the transformed force from eqn (9) and the second term is the spring force from eqn (10) that controls the distribution of the images along the path and increases

Downloaded by DTU Library on 23 January 2012
Published on 22 December 2011 on http://pubs.rsc.org | doi:10.1039/C2CP23421A

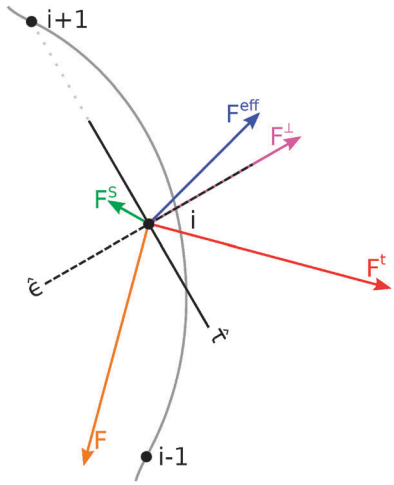

Fig. 2 The construction of the effective force, $\mathbf{F}^{\mathrm{eff}}$, which acts on image i of the path and is used in the iterative optimisation. The solid grey line indicates the ridge, the black filled circles represent the current location of three adjacent images, the black solid line shows the tangent estimate, $\hat{\boldsymbol{\tau}}$ and the black dashed line shows the minimum mode estimate, $\hat{\boldsymbol{\varepsilon}}$, at image i. The orange arrow shows the original force, $\mathbf{F} = -\nabla E$. The red arrow shows the transformed force, $\mathbf{F}^{\mathrm{t}}$, obtained by inverting $\mathbf{F}$ in the direction of the minimum mode, $\hat{\boldsymbol{\varepsilon}}$ (see eqn (9)). The purple arrow shows $\mathbf{F}^{\perp}$, the component of the transformed force that is perpendicular to the tangent. The green arrow shows $\mathbf{F}^{\mathrm{S}}$, the spring force given by eqn (10). The blue arrow shows $\mathbf{F}^{\mathrm{eff}}$, as given by eqn (11).

numerical stability. The construction of the effective force is illustrated in Fig. 2.

If the ridge is curved, the spring force tends to shorten the path by allowing the component perpendicular to the path to pull it off the ridge. The position of the images, then, converges to an equilibrium between the components of the spring force and the transformed force, that are perpendicular to the tangent,

$$\mathbf{F}_i^{\mathrm{t}} - (\mathbf{F}_i^{\mathrm{t}}\cdot\hat{\boldsymbol{\tau}}_i)\hat{\boldsymbol{\tau}}_i = -(\mathbf{F}_i^{\mathrm{S}} - (\mathbf{F}_i^{\mathrm{S}}\cdot\hat{\boldsymbol{\tau}}_i)\hat{\boldsymbol{\tau}}_i). \qquad (12)$$

This equilibrium position will systematically be slightly of the ridge. Equivalently, this is referred to as corner-cutting (see, for example ref. 12). The perpendicular component of the spring force can be projected out, as is often done in the NEB method, but here it is retained in order to improve the stability of the iterative optimisation.

The climbing image algorithm[13] can be used on the highest energy image of the path to ensure exact convergence to the highest energy SP_2 along the ridge. This is accomplished by decoupling the highest energy image from the springs and inverting the force along the tangent and the minimum mode in a manner similar to eqn (9),

$$\mathbf{F}_{i_{\max}}^{\mathrm{eff}} = \mathbf{F}_{i_{\max}} - 2(\mathbf{F}_{i_{\max}}\cdot\hat{\boldsymbol{\tau}}_{i_{\max}})\hat{\boldsymbol{\tau}}_{i_{\max}} - 2(\mathbf{F}_{i_{\max}}\cdot\hat{\boldsymbol{\varepsilon}}_{i_{\max}})\hat{\boldsymbol{\varepsilon}}_{i_{\max}}, \quad (13)$$

where $i_{\max}$ refers to the image with the highest energy. This decoupling achieves two things. It allows the highest energy image to converge onto the SP_2 exactly without significant increase in computational power and it allows the highest energy image to be decoupled from the spring force which, in turn, will allow it to overcome any tendency for corner-cutting.

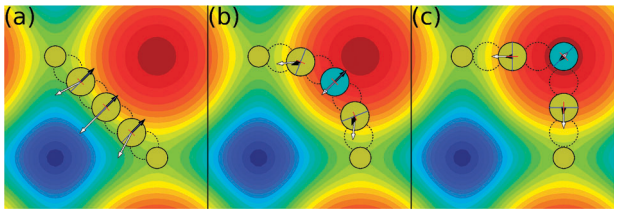

Fig. 3 Snapshots from an optimisation of a path between two neighbouring adatom hop SP_1s on a $2 \times 2 \times 1$ Al(100) slab. An animation is included in the ESI.† For simplicity, the Al atoms in the slab are kept fixed in this illustrative test problem. The energy surface is generated by minimizing the energy of the adatom along the normal to the surface plane but keeping the in-plane coordinates fixed. The circles represent images of the path. The cyan circle is the climbing image. For clarity, less information is shown for the dotted circles. The arrows represent the in-plane force acting on each image. (white) The force derived from the energy surface, $-\nabla E$. (black) The effective force, given by eqn (11) or (13), which is used in the iterative optimisation. The red lines represent the tangent and the blue lines represent the minimum mode estimate. The red areas of the surface represent high potential energy and the blue areas low potential energy. The immobile substrate atoms are located at the centre of the high potential areas. (a) Initial, straight line interpolation between the SP_1s, (b) after 19 optimisation steps, converged to 0.05 eV Å^{-1}, without climbing image, (c) after 70 steps, converged to 0.001 eV Å^{-1} with the climbing image algorithm.

The ridge calculations are different from MEP calculations in that the energy may not have a maximum along the path during the optimisation. The variation of the energy along the path can be monotonic or even an inverted barrier, *e.g.* when the initial path lies close to a minimum. This can lead to instability in the iterative optimisation and in order to increase the stability, the full spring force is used. Furthermore, complications can arise when turning on the climbing image algorithm as either of the immobile end points might, in fact, have the highest energy. In such cases it may prove beneficial for convergence to assign an image that is not the highest energy one as the climbing image. The images closest to the end images are partially constrained by the immobility of the end images, thus an image even further in, $\mathbf{R}_2$ or $\mathbf{R}_{N-2}$, is a better choice for the climbing image. When the path is near the ridge, the energy will have a maximum along the path and the highest energy image can be used as the climbing image. The lack of an intrinsic barrier makes climbing image ridge calculations risky to run from the initial interpolation. It is, therefore, better to, first, carry out enough iterations without the climbing image for an energy maximum to appear along the path and then turn on the climbing image (Fig. 3).

The method described here has been implemented and tested using the Atomic Simulation Environment (ASE)[19,20] using both analytical potential energy functions and density functional theory (DFT) to evaluate the atomic force.

3. Application: Al adatom diffusion on an Al(100) surface

An Al adatom on the Al(100) surface provides an interesting system to study because several different diffusion mechanisms have been found, including various concerted displacements of

 This journal is © the Owner Societies 2012

Downloaded by DTU Library on 23 January 2012
Published on 22 December 2011 on http://pubs.rsc.org | doi:10.1039/C2CP23421A

two or more atoms, in addition to the, more intuitive, hop mechanism.[10,21,22] An embedded atom method potential (EAM)[23] is used here since it has been shown to accurately describe the system and requires much less computational effort than DFT calculations. The simulated cell was a slab of 6 layers, each of which was 8 × 8 atoms, with one adatom, totalling 385 atoms. The two bottom layers were kept fixed at bulk positions with a lattice parameter of 4.038 Å. Initially, traditional NEB calculations were carried out to find the relevant SP_1s which were then used as end points in the ridge calculations. The spring constant was set at $k = 5.0$ eV Å^{-1}. The two images in the dimer had a fixed separation of 0.0001 Å and were allowed to rotate only once for each iteration in the path optimisation. The initial minimum mode guesses were taken from a Gaussian distribution. The convergence criteria for the maximum effective force component were set at 0.01 eV Å^{-1} and 0.001 eV Å^{-1} for the regular and climbing image calculations respectively. The FIRE algorithm[24] was used for the path optimisation.

Several low energy transition mechanisms for adatom diffusion on this surface have been found previously using the dimer method.[10] The mechanism with the lowest energy barrier is a two atoms concerted displacement, $E_b = 0.227$ eV. The second lowest is the simple hop of the adatom from one site to an adjacent site, $E_b = 0.372$ eV, but then three and four atoms concerted displacements are only slightly higher in energy $E_b = 0.426$ eV and $E_b = 0.413$ eV.

The potential energy ridges and SP_2s were calculated between each pair of SP_1s and the results are shown in Fig. 4 for the three concerted displacement mechanisms. Low energy SP_2s were found near the concerted 3- and 4-atom displacement SP_1s, with energies ~0.005 eV and ~0.012 eV. The energy of these SP_2s is less than thermal energy at room temperature, $k_BT = 0.025$ eV, over the adjacent SP_1s, which means that HTST is likely not a good approximation for these mechanisms.

In principle, knowledge of the ridge and the SP_2s can be used to improve on the HTST approximation. Here, a rough estimate of a correction factor, Γ, will be evaluated by calculating the ratio of the configuration integrals of the harmonic approximation to that calculated from the potential energy along the ridge,

$$\Gamma = \frac{Z_{\ddagger}^{\text{ridge}}}{Z_{\ddagger}^{\text{harm.}}} = \frac{\int_{\text{ridge}} \exp[-E(x)/k_BT]\,\mathrm{d}x}{\int_{-\infty}^{\infty} \exp[-\alpha x^2/k_BT]\,\mathrm{d}x}, \qquad (14)$$

where x is the displacement along the ridge and α is the curvature of the one-dimensional parabola obtained by performing least squares analysis of the 4 images closest to the SP_1s. The ratios obtained with eqn (14) are shown as a function of temperature in Fig. 5. As expected, the harmonic approximation works well for the concerted 2-atom process as the SP_1 is much lower in energy than the adjacent SP_2s in that case. On the other hand the concerted 3- and 4-atom displacements have a significant correction factor (~0.5 for the concerted 4-atom process at 350 K) as can be seen from the figure. It should be noted that the ratio increases with temperature due to the limited range of the ridge integral as compared with the infinite limits of the harmonic one. This is particularly

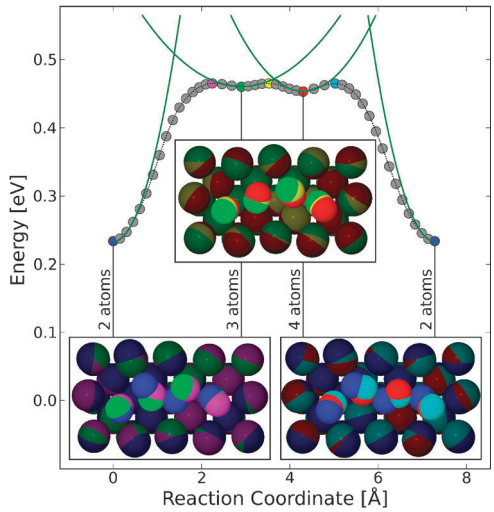

Fig. 4 The energy ridge going through the SP_1s of 2-atom, 3-atom, 4-atom and, then the same, 2-atom concerted displacement for an Al adatom on an Al(100) surface. The circles represent the position of images in the optimised paths, the SP_1s and the SP_2s being coloured differently but the rest coloured grey. The green curves represent harmonic approximations to the energy surface at each SP_1. The insets show an overlay of three configurations, two adjacent SP_1s and the intermediate SP_2. The atom colours correspond to the coloured circles of the energy ridge. The SP_2s adjacent to the 3-atom and 4-atom concerted displacements are low and the harmonic approximation to TST is less accurate for these mechanisms than the 2-atom concerted displacement.

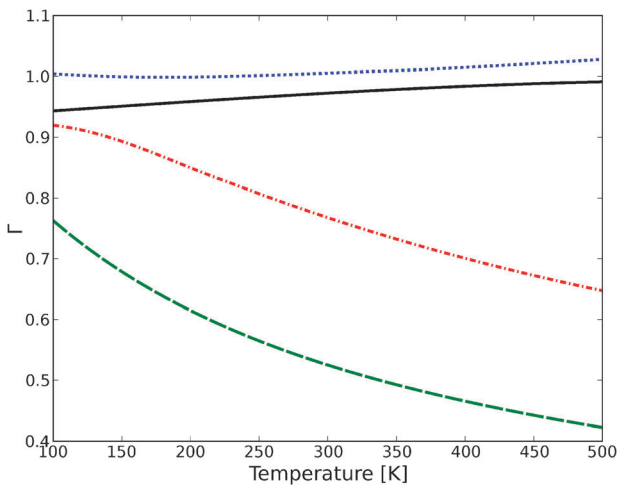

Fig. 5 The ratio, Γ, defined in eqn (14), between the configuration integrals of the potential energy ridge shown in Fig. 4 and the corresponding harmonic approximations. The black, solid, line is the ratio for the full integral, including all three concerted displacement processes. The blue, dotted, line is the ratio when only considering the 2-atom concerted displacement. The green, dashed, line is the ratio when only considering the 3-atom concerted displacement. The red, dash-dotted, line is the ratio when only considering the 4-atom concerted displacement. For the individual processes, the end points of the ridge integral are the adjacent SP_2s, while the full integral is done for the whole ridge. While the harmonic approximation gives a good approximation for the total configuration integral over the whole temperature range shown, because it is dominated by the 2-atom displacement, the estimate for each of the 3-atom and 4-atom displacements is poor unless the temperature is very low.

Downloaded by DTU Library on 23 January 2012
Published on 22 December 2011 on http://pubs.rsc.org | doi:10.1039/C2CP23421A

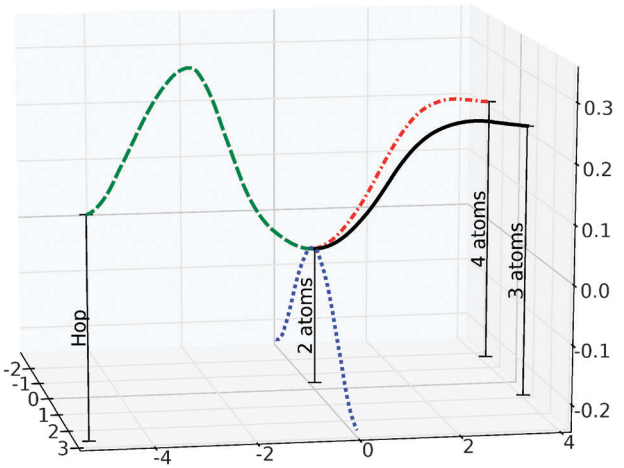

Fig. 6 A schematic view (due to the high dimensionality of the system) of some of the ridges lying through the SP_1 for 2-atom concerted displacement (the MEP is shown with blue dotted line). The three ridges shown connect to the SP_1s of the hop and the 2-atom and 3-atom displacement mechanisms. The vertical bars represent the height of each of the SP_1s. The figure illustrates that a SP_1 typically lies on several energy ridges for a complex system.

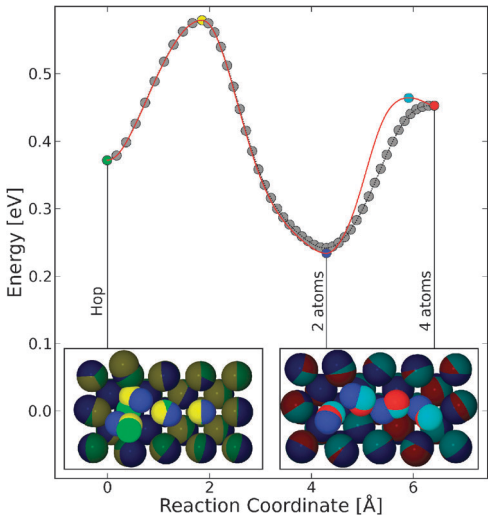

Fig. 7 Calculated path at the ridge between the SP_1s for a hop and concerted 4-atom displacement. The circles show the position of converged images (coloured circles for SP_1s and SP_2s, but gray for the rest). These SP_1s turned out not to be adjacent on a ridge and the path optimisation reveals an intermediate SP_1, the one for the concerted 2-atom displacement. This illustrates how a ridge calculation could reveal new and possibly unknown transition mechanisms. The long path is not able to accurately locate the intermediate SP_1 and the lower energy SP_2 due to finite resolution in the discretisation and corner-cutting. The exact configuration of the SP_1 can be found using a SP_1 finding algorithm starting with the approximation obtained from the optimised path. Then, a calculation of a shorter path, between the SP_1s of the 2-atom and 4-atom displacements, locates the intermediate SP_2 accurately (cyan circle). The insets show an overlay of three configurations, two adjacent SP_1s and the intermediate SP_2. The atom colours correspond to the coloured circles of the energy ridge.

prominent for the 2-atom concerted displacement where the ratio even goes above 1.0. In high dimensional systems, SP_1s lie on multiple ridges, as can be seen in Fig. 6, where 3 of the different ridges in which the concerted 2-atom displacement SP_1 lies on, are shown. Thus it may be necessary to perform corrections such as those above for more than one ridge for any given SP_1.

In the insets of Fig. 4, a comparison of the atom coordinates at two adjacent SP_1 and the intermediate SP_2 can be seen. In particular, the difference between the concerted 3- and 4-atom SP_1s is shown. The coordinates of the two left-most atoms only change slightly while the two right-most coordinates change more, as is to be expected as the atom furthest to the right is not directly involved in the concerted 3-atom process. The similarities in coordinates and the small SP_2s separating the 3- and 4-atom displacement processes indicate that a trajectory passing through the vicinity of either SP_1 could easily end up in the product state corresponding to the other. Dynamical trajectories started at the ridge would be needed to determine the probability of each of the product states.

When finding the ridge between the concerted 4-atom displacement and the hop SP_1s, shown in Fig. 7, it became apparent that a ridge does not directly connect the two. Instead, the ridge passes through the concerted 2-atom displacement SP_1. The path passes exactly through the highest SP_2, as can be verified from the calculated force, given that the climbing image algorithm is employed. However, due to the possible corner cutting, there is no guarantee that other, lower energy, SP_2s along the ridge will be found exactly. Nevertheless, if a sufficient number of images is used, the path will give good approximation for any SP_1 and SP_2 along the ridge. Here, the image closest to the concerted 2-atom displacement SP_1 is found to be at a distance of 0.004 Å per atom from the exact SP_1. Using these coordinates in a SP_1 searching algorithm quickly yields the exact SP_1. As for the lower SP_2, a second ridge calculation can be performed with the adjacent SP_1s as endpoints to focus on a shorter segment of the ridge with only one intermediate SP_2, thereby enabling the climbing image to converge exactly on the lower SP_2. This is shown in Fig. 7, where the discovered SP_1 is used as an endpoint in a subsequent optimisation of a shorter path and, thus, the lower SP_2 is found accurately.

3.1. Performance

Convergence depends on a number of factors; the accuracy of the forces, the number of atoms in the system, the number of images in the calculation and the algorithm used for optimisation. Furthermore, the specific parameters of the minimum mode algorithm, the stiffness of the springs and the criterion for non-climbing image convergence influence the ridge convergence. Although no comprehensive convergence study for the plethora of parameters was performed, Table 1 shows convergence data for the processes presented here. In order to avoid any spurious complication during the path optimisation, the simple, but robust, FIRE algorithm[24] was used to converge each 29 image path until the maximum effective force component was smaller than 0.01 eV $Å^{-1}$ and 0.001 eV $Å^{-1}$ for the regular and climbing image calculations respectively. The spring constant was set at 5.0 eV $Å^{-1}$ and the dimer was allowed to rotate only once per optimisation step.

 This journal is © the Owner Societies 2012

Downloaded by DTU Library on 23 January 2012
Published on 22 December 2011 on http://pubs.rsc.org | doi:10.1039/C2CP23421A

Table 1 Performance of ridge calculations. The average number of optimisation steps, force calculations and a comparison with the NEB calculations of the endpoints (and corresponding MEPs), of 50 path optimisations for each of the concerted multiple atom diffusion mechanisms using 29 images. Where applicable the exact same parameters were used for both the MEP and ridge calculations. It should be noted that the SP_2 is generally found much faster than the ridge in full

	Optimisation steps		Force evaluations	
Diffusion mechanism	Ridge	Ridge/MEP	Ridge	Ridge/MEP
Concerted 2 ↔ 3 atom	398	2.5	23 456	5.5
Concerted 2 ↔ 4 atom	400	2.5	23 558	5.4
Concerted 3 ↔ 4 atom	132	0.8	6709	1.6

The concerted 3 and 4 atom mechanisms are rather hard for the NEB to find as they correspond to shallow minima along the energy ridge. As a result, the ridge calculations can take fewer optimisation steps than the corresponding NEB calculations. On the other hand, ridge calculations have difficulty in converging the paths which have a SP_2 close to one SP_1 but far from the other, such as the ridges between the concerted 2 and 3 and concerted 2 and 4 atom mechanisms. Near these SP_2s, the dimer will sometimes find a soft eigenmode, *e.g.* rearrangements of the full system, which can result in failed optimisations or ones that are not useful.

A general comparison of the efficiency of NEB and ridge calculations is not feasible as the methods are used for different types of paths. The data in Table 1, nevertheless, give some indication of the resources required to perform a ridge calculation.

4. Summary

A method for finding an energy ridge and second order saddle points between first order saddle points is presented and used to assess the applicability of the harmonic approximation to transition state theory in a test system involving the diffusion of an Al adatom on an Al(100) surface. The method is based on locally transforming the gradient near an energy ridge to that of a MEP using an inversion of the gradient along the direction of the minimum mode perpendicular to the path. The method can be regarded as an extension of the nudged elastic band method to find an energy ridge instead of a MEP.

The application of the method to an Al adatom diffusion on an Al(100) surface shows that the harmonic approximation can be expected to perform well over a wide range of temperature for the lowest energy process, the concerted 2-atom displacement. On the other hand, for the 3-atom and 4-atom concerted displacement mechanisms, the harmonic approximation is not accurate because the SP_2s adjacent to the SP_1s for these processes are comparatively low in energy. After the energy ridge has been located, it is possible to estimate a correction factor for the reaction rate given by HTST by comparing the configuration integrals of the ridge and the harmonic approximations. Finally, calculation of the ridge between SP_1s that are not adjacent can reveal additional, intermediate, SP_1s on the ridge. This can be useful when dealing with complicated energy landscapes.

In the implementation described here, we have included the full spring force that distributes images along the path in order to improve stability of the optimisation. This can lead to corner-cutting in the path where it does not follow the ridge exactly. However, by using the climbing image algorithm, an exact convergence to the highest SP_2 can be ensured.

The ability to identify energy ridges and second order saddle points could become a valuable tool in computational materials design. One example is the study of catalytic selectivity in heterogeneous or electro-catalytic processes such as methane *vs.* higher alcohol formation in electrochemical CO_2 fixation[25] or ammonia *vs.* hydrogen formation in electrochemical N_2 fixation.[26] The method could, thereby, help estimate Faradaic losses in electrochemical processes.

While the discussion here has focused on a classical, over-the-barrier description of the transition mechanism, similar considerations can apply to transitions where quantum mechanical tunnelling takes place. A harmonic quantum transition state theory can be applied in such cases (see for example ref. 27–29) and similar considerations about the height of SP_2s over SP_1s on an extended, quantum mechanical energy surface, describing the energy of closed Feynman paths, may apply. The energy ridge of the extended energy surface could in principle be found by an extension of the method presented here.

Even though the main focus here is on the application to rate theory and atomic simulations, the method can be used to find ridges and SP_2s for any function of multiple variables where the first derivative is readily available. In fact, the method can easily be extended to find saddle points of order N, SP_Ns, by ensuring that there are $N - 1$ negative eigenvalues for the reduced Hessian. This can be accomplished by having $N - 1$ constrained eigenmode searchers perpendicular to each other and the tangent of the path and applying eqn (8) to each one.

Acknowledgements

The authors would like to acknowledge support from the European Graduate School on Sustainable Energy Technology, the Icelandic Science Foundation and Danish Center for Scientific Computing for supercomputer access. The Center for Atomic Scale Material Design (CAMD) is funded by the Lundbeck foundation and the Catalysis for Sustainable Energy (CASE) initiative is funded by the Danish Ministry of Science, Technology and Innovation.

References

1 E. Wigner, *Trans. Faraday Soc.*, 1938, **34**, 29.
2 H. Eyring, *J. Chem. Phys.*, 1935, **3**, 107.
3 M. G. Evans and M. Polanyi, *Trans. Faraday Soc.*, 1935, **31**, 875.
4 P. Hänggi, P. Talkner and M. Borkovec, *Rev. Mod. Phys.*, 1990, **62**, 251.
5 E. Pollack and P. Talkner, *Chaos*, 2005, **15**, 026116.
6 J. C. Keck, *J. Chem. Phys.*, 1967, **13**, 85.
7 C. Wert and C. Zener, *Phys. Rev.*, 1949, **76**, 1169.
8 G. H. Vineyard, *J. Phys. Chem. Solids*, 1957, **3**, 121.
9 H. Jónsson, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 944.
10 G. Henkelman and H. Jónsson, *J. Chem. Phys.*, 1999, **111**, 7010.
11 R. A. Olsen, G. J. Kroes, G. Henkelman, A. Arnaldsson and H. Jónsson, *J. Chem. Phys.*, 2004, **121**, 9776.
12 H. Jónsson, G. Mills and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, ed. B. J. Berne, G. Ciccotti and D. F. Coker, World Scientific, 1998, p. 385.

13 G. Henkelman, B. P. Uberuaga and H. Jónsson, *J. Chem. Phys.*, 2000, **113**, 9901.
14 G. Henkelman and H. Jónsson, *J. Chem. Phys.*, 2000, **113**, 9978.
15 G. Henkelman, A. Arnaldsson and H. Jónsson, *J. Chem. Phys.*, 2006, **124**, 044706.
16 G. Henkelman, G. Jóhannesson and H. Jónsson, in *Theoretical Methods in Condensed Phase Chemistry*, vol. 5 *Progress in Theoretical Chemistry and Physics*, ed. Steven D. Schwartz, Springer, Netherlands, 2002, p. 269.
17 J.-Q. Sun and K. Ruedenberg, *J. Chem. Phys.*, 1993, **98**, 9707.
18 E. M. Sevick, A. T. Bell and D. N. Theodorou, *J. Chem. Phys.*, 1993, **98**, 3196.
19 S. R. Bahn and K. W. Jacobsen, *Comput. Sci. Eng.*, 2002, **4**, 56.
20 https://wiki.fysik.dtu.dk/ase.
21 P. J. Feibelman, *Phys. Rev. Lett.*, 1990, **65**, 729.
22 G. Jóhannesson and H. Jónsson, *J. Chem. Phys.*, 2001, **115**, 9644.
23 A. F. Voter and S. P. Chen, *MRS Proceedings of 'Symposium I Characterization of Defects in Materials'*, 1986, **82**, 175.
24 E. Bitzek, P. Koskinen, F. Gähler, M. Moseler and P. Gumbsch, *Phys. Rev. Lett.*, 2006, **97**, 170201.
25 A. Peterson, F. Abild-Pedersen, F. Studt, J. Rossmeisl and J. Nørskov, *Energy Environ. Sci.*, 2010, **3**, 1311.
26 E. Skúlason, T. Bligaard, S. Gumundsdóttir, F. Studt, J. Rossmeisl, F. Abild-Pedersen, T. Vegge, H. Jónsson and J. K. Nørskov, *Phys. Chem. Chem. Phys.*, 2012, **14**, 1235–1245.
27 S. Andersson, G. Nyman, A. Arnaldsson, U. Manthe and H. Jónsson, *J. Phys. Chem. A*, 2009, **113**, 4468.
28 G. Mills, G. K. Schenter, D. Makarov and H. Jónsson, *Chem. Phys. Lett.*, 1997, **278**, 91.
29 G. Mills, G. K. Schenter, D. Makarov and H. Jónsson, *RAW Quantum Transition State Theory*, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, ed. B. J. Berne, G. Ciccotti and D. F. Coker, World Scientific, 1998, p. 405.

Downloaded by DTU Library on 23 January 2012
Published on 22 December 2011 on http://pubs.rsc.org | doi:10.1039/C2CP23421A

This journal is © the Owner Societies 2012