
Metastability, conformation dynamics, and transition pathways in complex systems

Weinan E¹ and Eric Vanden-Eijnden²

¹ Department of Mathematics and PACM, Princeton University, Princeton, NJ 08544 weinan@princeton.edu

² Courant Institute, New York University, New York, NY 10012 eve2@cims.nyu.edu

Summary. We present a systematic introduction to the basic concepts and techniques for determining transition pathways and transition rates in systems with multiple metastable states. After discussing the classical transition state theory and its limitations, we derive a new set of equations for the optimal dividing surfaces. We then discuss transition path sampling, which is the most general technique currently available for determining transition regions and rates. This is followed by a discussion on minimal energy path for systems with smooth energy landscapes. For systems with rough energy landscapes, our presentation is centered around the notion of reaction coordinates. We discuss the two related notions of free energies associated with a reaction coordinate, and show that at least in the high friction limit, there does exist an optimal reaction coordinate that gives asymptotically the correct prediction for the transition rates. Variational principles associated with the optimal reaction coordinates are exploited under the assumption that the transition paths are restricted to tubes, and this provides a theoretical justification for the finite temperature string method. Blue moon sampling techniques, metadynamics and a new form of accelerated dynamics are also discussed.

1 Introduction

The evolution of complex systems often involves widely separated time-scales. Well-known examples include nucleation events during phase transition, conformational changes of molecules, and chemical reactions. The appearance of long time-scales is related to the existence of energy barriers or entropic bottlenecks which partition the system into metastable basins. A dynamical trajectory may spend a very long time in one basin before finding its way to another. The separation of time-scales is typically so pronounced that it is difficult to observe a single transition event, let alone gather enough statistical information about the transitions, by using conventional dynamical simulations. Several techniques have been introduced to tackle the numerical difficulty of determining the transition pathways and transition rates in sys-

tems of this kind. The purpose of the present paper is to review some (not all) of these techniques and the underlying theoretical framework. In doing so, we find it sometimes more convenient to take a point of view that is slightly different from that of the original papers. In addition, we will make an attempt to put frequently used notions such as reaction coordinates, optimal reaction coordinates and optimal dividing surfaces on a solid footing.

We shall mostly focus on methods which assume the knowledge of the initial and final states for the transition that we are interested in. The transition may proceed through additional metastable states that are not identified beforehand. We will understand metastability in the following way. Suppose that the system admits a unique equilibrium distribution such as the Gibbs distribution

$$\rho(x) = Z^{-1}e^{-\beta V(x)}, \quad Z = \int_{\mathbb{R}^n} e^{-\beta V(x)} dx \quad (1)$$

where $x \in \mathbb{R}^n$ is the configuration space, $V(x)$ is some given potential, and $\beta = 1/k_B T$ is the inverse temperature. Then the regions $\{B_j\}_{j=1}^J$ will be metastable if

$$\sum_{j=1}^J N_j \approx 1 \quad \text{where } N_j = \int_{B_j} \rho(x) dx, \quad (2)$$

i.e. with very high probability, the system at equilibrium is found in one of the sets B_j . Assuming ergodicity, N_j is also the fraction of time that the system spends in B_j . In addition to (2) one must also require that the regions B_j be well separated so that the transitions between these regions are quasi-Markov events. We shall come back to this requirement later. Note that the Gibbs distribution in (1) is consistent with the Langevin dynamics

$$\begin{cases} \dot{x} = v \\ M\dot{v} = -\nabla V(x) - \gamma v + \sqrt{2\gamma\beta^{-1}} \eta \end{cases} \quad (3)$$

where $\eta(t)$ is a white-noise, γ is the friction coefficient, and M is the diagonal mass matrix. Two limiting cases of (3) are the overdamped (high friction) dynamics obtained as $\gamma \rightarrow \infty$

$$\gamma \dot{x} = -\nabla V(x) + \sqrt{2\gamma\beta^{-1}} \eta \quad (4)$$

and the Hamiltonian dynamics which arises when $\gamma = 0$

$$M\ddot{x} = -\nabla V(x). \quad (5)$$

2 Transition state theory and the Bennett-Chandler procedure

Transition state theory is the oldest attempt to describe metastability in ergodic systems [Wig38, Eyr35]. Consider a system governed by the Hamiltonian

dynamics (5), and let us partition the configuration space into two sets, A and its complement A^c . Define the mean residence time in A and A^c as

$$t_A = \lim_{T \rightarrow \infty} \frac{2}{N_T} \int_0^T \chi_A(x(t)) dt, \quad (6)$$

and similarly for t_{A^c} . Here $x(t)$ is the instantaneous position of a generic trajectory, N_T is the number of times this trajectory crosses the boundary between A and A^c before time T , and χ_A denotes the indicator function of A , i.e. $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ otherwise. (6) can be rewritten as

$$t_A = N_A / \nu^{\text{TST}}, \quad t_{A^c} = N_{A^c} / \nu^{\text{TST}} \quad (7)$$

where

$$N_A = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \chi_A(x(t)) dt = \int_A \rho(x) dx, \quad (8)$$

is the proportion of time that the trajectory $x(t)$ spends in A , i.e. the equilibrium population density in A assuming ergodicity, $N_{A^c} = 1 - N_A$, and

$$\nu^{\text{TST}} = \lim_{T \rightarrow \infty} \frac{N_T}{2T}, \quad (9)$$

is the (half) mean frequency of crossing the boundary ∂A . This mean frequency can be estimated upon noting that $|\dot{\chi}_A(x(t))|$ is a sum of delta functions concentrated at the times when $x(t)$ crosses ∂A . Therefore

$$\begin{aligned} \nu^{\text{TST}} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_0^T |\dot{\chi}_A(x(t))| dt \\ &= \frac{1}{2} \int_{\mathbb{R}^n} \int_{\partial A} |\hat{n} \cdot v| \rho(x, v) d\sigma(x) dv \end{aligned} \quad (10)$$

Here $\rho(x, v)$ is the joint equilibrium density of (x, v) and $d\sigma(x)$ is the surface element on the boundary ∂A . Assuming that the number of degrees of freedom is large, $\rho(x, v)$ can be approximated by the Gibbs distribution $Z_H^{-1} e^{-\beta(|v|^2/2 + V(x))}$, and (10) reduces to

$$\nu^{\text{TST}} = \sqrt{\frac{2}{\pi\beta}} Z^{-1} \int_{\partial A} e^{-\beta V(x)} d\sigma(x). \quad (11)$$

Transition state theory is exact for the computation of the mean residence time. But the mean residence time may not be very relevant as far as description of transition events is concerned. Indeed, since a trajectory that has just crossed the boundary ∂A may have a high probability of re-crossing it right away, the successive transition times between A and A^c are correlated in general. Therefore, the symbolic dynamics on these two sets is usually poorly approximated by a two-state Markov chain with rates

$$k_{AA^c} = t_A^{-1} = \nu^{\text{TST}}/N_A, \quad k_{A^cA} = t_{A^c}^{-1} = \nu^{\text{TST}}/N_{A^c} \quad (12)$$

Quasi-Markov transitions will only arise between sets that are sufficiently separated, i.e. there exists a buffer region between them that has a large volume but negligible probability. How to correct transition state theory to handle this situation was explained by Bennett [Ben77] and Chandler [Chan78] (see also [Yam60]). These authors base their considerations on response functions. Here we take a slightly different viewpoint. Consider a system with two metastable sets, B_1 and B_2 , satisfying (2) with $J = 2$, and define A such that ∂A belongs to the buffer region $(B_1 \cup B_2)^c$ (i.e. $B_1 \subset A$ and $B_2 \subset A^c$). Then the rates between B_1 and B_2 can be expressed as

$$k_{12} = \nu/N_1, \quad k_{21} = \nu/N_2, \quad (13)$$

where $N_1 \approx N_A$ and $N_2 \approx N_{A^c}$ are the population densities in B_1 and B_2 , respectively, and ν is the corrected frequency

$$\nu = \frac{1}{2} \int_{\mathbb{R}^n} \int_{\partial A} |\hat{n} \cdot v| \rho(x, v) \chi(x, v) d\sigma(x) dv. \quad (14)$$

The factor $\chi(x, v)$ is included so that only trajectories passing through (x, v) which correspond to transition events, i.e. they go from ∂B_1 to ∂B_2 , or ∂B_2 to ∂B_1 , are accounted for. Thus, if $N_R(x, v)$ is the number of crossings of ∂A between the times the trajectory exit B_1 and enters B_2 or conversely, one has

$$\chi(x, v) = \frac{1}{2} \left(1 - (-1)^{N_R(x, v)} \right) / N_R(x, v). \quad (15)$$

The first factor $\frac{1}{2}(1 - (-1)^{N_R})$ is 1 if N_R is odd (i.e. if the trajectory connects B_1 to B_2 or B_2 to B_1) and zero otherwise; the second factor N_R^{-1} accounts for over-counting since N_R starting points on ∂A are contained in the same transition trajectory. The ratio ν/ν^{TST} , which is necessarily smaller or equal to one, is called the transmission coefficient κ of the surface ∂A .

The Bennett-Chandler approach in principle offers a way of identifying the transition paths and computing the mean transition times in a two-step procedure. First choose a dividing surface ∂A and compute the expectation in (10) or (11) using e.g. umbrella or blue-moon sampling techniques and thermodynamic integration (see section 8). Then evaluate the transmission coefficient of ∂A using

$$\kappa = \frac{1}{2} \left\langle \left(1 - (-1)^{N_R(x, v)} \right) / N_R(x, v) \right\rangle_{\partial A} \quad (16)$$

where $\langle \cdot \rangle_{\partial A}$ denotes expectation with respect to the density proportional to $|\hat{n} \cdot v| \rho(x, v)$ after proper normalization. Evaluating (16) amounts to initiating trajectories from ∂A based on this density, running them both forward and backward in time, observing if they correspond to transition trajectories and counting the number of times they re-cross ∂A . Unfortunately, this procedure suffers from the following practical difficulties.

1. ∂A might be a poor dividing surface in the sense that very few trajectories initiated on ∂A give rise to switching and, hence, are statistically significant in the Bennett-Chandler procedure. A good dividing surface should have the property that trajectories initiated on it have about half probability to reach either B_1 or B_2 .

2. Even if ∂A is a good dividing surface, it will be difficult to compute accurately the transmission coefficient κ if the transition trajectories tend to re-cross many times ∂A before reaching either B_1 or B_2 . The main reason is that the initial conditions generated on ∂A by sampling are correlated when the number of re-crossings is large. A set of N transition trajectories that on the average re-cross N_R times the surface ∂A contains only N/N_R statistically significant transition trajectories [Str01]. In situations when N_R is large, the transition is said to be diffusive (as opposed to being ballistic when recrossing is rare). This may mean that the transition state is in reality a rather wide region.

3 Optimizing TST

Since TST necessarily overestimate the rate, it is natural to look for the dividing surface with minimum TST rate. Using (7) and (11) this is the dividing surface that minimizes

$$I_0 = \int_{\partial A} e^{-\beta V} d\sigma(x). \quad (17)$$

It is shown in Appendix A that the minimizer of (17) satisfies

$$0 = -\nabla V \cdot \hat{n} + \beta^{-1}\kappa, \quad (18)$$

where $\hat{n}(x)$ is the unit normal to ∂A at point x and $\kappa = \nabla \cdot \hat{n}$ is the local mean curvature of this surface.

In general, optimizing ∂A by solving (18) will be too formidable a task to be practical. More realistically, one can minimize the functional (17) over restricted classes of surfaces. The simplest choice is to assume that ∂A is planar, in which case it is specified by the equation

$$0 = \hat{n} \cdot x - b \quad (19)$$

where b is a scalar and \hat{n} is the unit normal to the plane. Using the planar ansatz in (17) gives

$$I_0 = \int_{\mathbb{R}^n} e^{-\beta V} \delta(\hat{n} \cdot x - b) dx. \quad (20)$$

As shown in Appendix A, the minimizers of (20) satisfy

$$\begin{cases} 0 = \int_{\mathbb{R}^n} \hat{n} \cdot \nabla V e^{-\beta V} x^\perp \delta(\hat{n} \cdot x - b) dx \\ 0 = - \int_{\mathbb{R}^n} \hat{n} \cdot \nabla V e^{-\beta V} \delta(\hat{n} \cdot x - b) dx, \end{cases} \quad (21)$$

where $x^\perp = x - (x \cdot \hat{n})\hat{n}$ is the in-plane projection of x . These equations are the steady state solutions of the gradient flow

$$\begin{cases} \dot{\hat{n}} = -\langle (\hat{n} \cdot \nabla V) x^\perp \rangle_P \\ \dot{b} = \langle \hat{n} \cdot \nabla V \rangle_P, \end{cases} \quad (22)$$

where $\langle \cdot \rangle_P$ denotes the expectation with respect to $e^{-\beta V} \delta(\hat{n} \cdot x - b)$ properly normalized. Therefore (22) can be solved by evaluating the right hand-sides using blue moon sampling technique discussed in section 8.

Jóhannesson and Jónsson [JJ01] were the first to suggest optimizing the TST dividing surface among planes. However, their derivation is different from ours, and they arrive at different equations for \hat{n} and b .

Whether the optimal TST dividing surface can be approximated by a plane is not *a priori* clear. However a more serious limitation of the procedure is that the optimal TST dividing surface may not solve either one of the two problems in the Bennett-Chandler procedure. Clearly the transition process may be such that the trajectories re-cross the optimal TST surface many times, i.e. the procedure does not solve the second problem in the Bennett-Chandler procedure. But it is important to realize that for such diffusive crossings the procedure may not solve the first problem either: the optimal TST surface may not be a good dividing surface, i.e. trajectories initiated from this surface may not have half-half probability to reach the metastable sets B_1 or B_2 .

This indicates that the very concept of a dividing surface for the transition may be flawed. Therefore there is a need for techniques that, besides evaluating transition rates, give a more global insight about the mechanism of transition than a mere dividing surface. Techniques of this kind are discussed next.

4 Transition path sampling

In [Pra86], Pratt suggested a strategy for dealing with the rarity of transition trajectories by sampling precisely these trajectories via Monte-Carlo techniques in path space. The Transition Path Sampling (TPS) method, developed by Bolhuis, Chandler, Dellago, and Geissler, is a practical way of implementing such a strategy [DBG02, BCDG02].

We first explain TPS in the overdamped limit. The statistical weight of a trajectory $x(t)$ on $t \in [0, T]$ with $x(0) = x_0$ is, after rescaling of time by $t/\gamma \rightarrow t$, proportional to

$$\exp\left(-\frac{1}{4}\beta \int_0^T |\dot{x} + \nabla V|^2 dt\right). \quad (23)$$

In other words, any average of a functional of $x(t)$, $t \in [0, T]$, can in principle be evaluated via the path integral of this functional with the weight (23) over all paths $x(t)$ satisfying $x(0) = x_0$. There are some technical difficulties in properly normalizing (23) but since a Metropolis Monte-Carlo algorithm would only require (23) and not the normalization factor, we will leave these difficulties aside. Suppose that instead of the constraint $x(0) = x_0$ one uses $x(0) \in B_1$, $x(T) \in B_2$, then such a Monte-Carlo scheme will sample precisely the transition paths (or at least the ones that switch in less than T) with the proper weight. How to implement such a procedure in practice after time-discretization of (23) – i.e. how to generate new paths with reasonable acceptance rate in a Metropolis algorithm, etc – is explained in [DBG02, BCDG02].

In principle, TPS removes all pre-assumptions about the mechanism of transition by sampling all the paths that may be involved in such transitions. It is a brute force calculation of the true dynamical trajectories, except that the method cleverly manages to observe these trajectories precisely during the windows of time when a transition happens. Proper analysis of the paths sampled by TPS (which is a nontrivial operation beyond TPS) may lead to better choices for potential dividing surfaces to be used in the Bennett-Chandler procedure. This removes the first difficulty mentioned in section 2. In principle, the second can also be removed by abandoning the Bennett-Chandler procedure altogether and using TPS combined with umbrella sampling in which the volume of B_2 is artificially increased until this set intersects with B_1 to compute the rate directly from TPS. This procedure is necessary because the transition rate is expressed in terms of average over path constrained on their initial position only, $x(0) \in B_1$.

TPS can be easily generalized to dynamics with arbitrary finite friction by replacing (23) by

$$\exp\left(-\frac{1}{4}\beta\gamma^{-1}\int_0^T|M\ddot{x} + \gamma\dot{x} + \nabla V|^2 dt\right). \quad (24)$$

The Hamiltonian situation is more delicate. In this case, the statistical weight on any path connecting B_1 and B_2 is either one or zero; depending on whether this path solves (5), in which case it is as good as any other such path and must be accepted, or it does not solve (5), in which case it must be rejected. Procedures have been developed to sample paths with such a weight [BDC98].

As we mentioned earlier, TPS has the advantage that it makes no assumptions on the mechanism of transition. But it may be very demanding when the transition takes a relatively long time (e.g. for diffusive crossings) or when several well-separated channels for the transition exist, in which case there is also metastability in path space. In addition, analyzing the data provided by TPS, e.g. to determine a few reaction coordinates governing the transition, is nontrivial. The techniques we will discuss next provide insight about these reaction coordinates, often (but not always) at the expense of making additional approximations. These techniques are in general less expensive than TPS.

5 Identifying the minimal energy path for smooth energy landscapes

If the potential energy landscape is smooth, i.e. has no features on the thermal energy scale, and the critical points are isolated, the metastable states are the local minima and the transition states between local minima are the saddle points along the minimum energy paths (MEP) between these minima. In this section, we discuss methods for identifying MEPs.

Analytically a MEP is a heteroclinic orbit, i.e. it is a curve connecting the two minima with the property that the component of ∇V perpendicular to the curve is zero,

$$\nabla V^\perp = 0. \quad (25)$$

One way of finding solutions of (25) is to use the gradient flow of curves in configuration space:

$$v_n = -\nabla V^\perp, \quad (26)$$

where v_n is the normal velocity of the curve. There are several different ways of implementing this dynamics of curves, according to how the curves are parameterized. If $\varphi(\alpha, t)$ is a curve parameterized by $\alpha \in [0, 1]$, then the most direct implementation of (26) (proposed originally in [UE89]),

$$\frac{\partial \varphi}{\partial t} = -\nabla V^\perp(\varphi) = -\nabla V(\varphi) + (\nabla V(\varphi) \cdot \hat{t})\hat{t}, \quad \hat{t} = \varphi_\alpha / |\varphi_\alpha|, \quad (27)$$

will in general not preserve the parameterization of $\varphi(\alpha, t)$ during the evolution. This may lead to computational instabilities if the discretization points along $\varphi(\alpha, t)$ happen to cluster in certain parts of the curve and leave other parts under-resolved. In [JMJ98], Jónnson, Mills, and Jacobson gave a solution to this problem by modifying (27) to

$$\frac{\partial \varphi}{\partial t} = -\nabla V(\varphi) + (\nabla V(\varphi) \cdot \hat{t})\hat{t} + \lambda(\varphi_{\alpha\alpha} \cdot \hat{t})\hat{t}, \quad (28)$$

where $\lambda > 0$ is an adjustable parameter. The new term at the right hand-side is a penalty term which helps to distribute the points uniformly along the curve after it is discretized. Note that both the potential and the spring forces are nudged in (28), respectively in the directions perpendicular and tangential to the curve. For this reason, the method is referred to as the Nudged Elastic Band (NEB) method [JMJ98].

On the other hand, the broader perspective of evolving curves with intrinsic parameterization is useful to remember since it offers possibilities different from (28) which may lead to more efficient numerical schemes. For instance instead of (28), one can use

$$\frac{\partial \varphi}{\partial t} = -\nabla V(\varphi) + (\nabla V(\varphi) \cdot \hat{t})\hat{t} + r\hat{t}, \quad (29)$$

where r is now a Lagrange multiplier for enforcing some specific parameterization of $\varphi(\alpha, t)$. A simple example is to use the equal arclength parameterization, i.e. $|\varphi_\alpha|$ is constant. But other choices are possible. A method based on (29) was introduced in [ERVE02a] under the name of the (zero-temperature) string method. The string method turns out to have several advantages. (29) is free of adjustable parameter and it naturally leads to very flexible time-splitting schemes where φ is evolved by the term $-\nabla V^\perp$ alone for a couple of time steps, then a re-parameterization step is performed to re-enforce proper parameterization. This makes it very simple to use any constraint on the parameterization – such as arc-length weighted by energy or curvature which put more discretization point where needed – or to change the number of discretization points at the re-parameterization step if necessary. Note that for both (28) and (29) upwind scheme for the derivative with respect to α must be used to guarantee stability of the scheme [HJ00, Ren02]. Also, higher order schemes in α can be constructed [Ren03], and convergence rate can be improved by changing the steepest descent dynamics to Broyden-like dynamics [Ren02].

Once the MEPs have been identified, the transition rate can be estimated for instance by

$$k_{12} = (4\pi)^{-1} \left(\sqrt{\gamma^2 + |\lambda_s|} - \gamma \right) \sqrt{|\det(H_m H_s^{-1})|} e^{-\beta \Delta V}, \quad (30)$$

where λ_s is the unstable eigenvalue of the Hessian of $V(x)$ evaluated at the saddle point between the two minima, H_s and H_m are the Hessians of V evaluated at the starting minimum and the saddle point, respectively, and ΔV is the energy difference between the saddle point and the starting minimum. At nonzero friction, these statements follow from the theory of large deviation and extensions thereof to account for the prefactor [FW98]. In the Hamiltonian case, (30) follows by evaluation of (11) in the harmonic approximation.

The assumptions underlying these results are very restrictive since, in many interesting situations, the energy has too many critical points. In addition these critical points are mostly irrelevant for the transitions because the saddle point(s) are not, in general, a fair approximation of the transition state. One may think that this can be checked *a posteriori* by initiating trajectories from these saddle points. But in practice this will be undoable if the number of saddle points is so large that it is impossible to identify them all.

6 Finite temperature string method

Another advantage of the (zero-temperature) string method is that it can be naturally generalized to problems with rough energy landscapes by coupling the dynamics of the string with some sampling procedure [ERVE02b]. This allows to move the string in a thermally averaged potential whose irrelevant

details on the thermal scale have been smoothed out. Instead of (29) it was proposed in [ERVE02b] to use

$$\frac{\partial \varphi}{\partial t} = -\langle \nabla V(\tilde{\varphi}) \rangle + (\langle \nabla V(\tilde{\varphi}) \rangle \cdot \hat{t})\hat{t} + r\hat{t}. \quad (31)$$

Here $\langle \nabla V(\tilde{\varphi}) \rangle$ denotes some suitably defined statistical average. A seamless way of defining such statistical averages and moving the string at the same time is given by

$$\frac{\partial \tilde{\varphi}}{\partial t} = -\nabla V(\tilde{\varphi}) + (\nabla V(\tilde{\varphi}) \cdot \hat{t})\hat{t} + \sqrt{2\beta^{-1}} (\xi - (\xi \cdot \hat{t})\hat{t}) + r\hat{t}. \quad (32)$$

Here $\tilde{\varphi}(\alpha, t)$ is a stochastic process whose mean is the string defined earlier in (31): $\varphi = \langle \tilde{\varphi} \rangle$. \hat{t} is the unit tangent along this curve, r is a Lagrange multiplier term to enforce some constraint on the parameterization of φ , and η is a white-noise satisfying

$$\langle \xi(\alpha, t)\xi(\alpha', t') \rangle = \begin{cases} \delta(t - t') & \text{if } \alpha = \alpha' \\ 0 & \text{otherwise} \end{cases} \quad (33)$$

Because of the resemblance of (32) with (29), the method based on (32) was called the finite temperature string method in [ERVE02b]. Notice that (31) can be obtained from averaging (32).

It is a simple matter to show that, for each α , the equilibrium density function for (32) is

$$\rho(x, \alpha) = Z^{-1}(\alpha) e^{-\beta V} \delta(\hat{t} \cdot (x - \varphi)), \quad (34)$$

i.e. it is the restriction of $e^{-\beta V}$ in the plane perpendicular to the curve φ containing the point $x = \varphi(\alpha)$, which can be parameterized as $0 = \hat{t} \cdot (x - \varphi)$. We will refer to this plane as $P(\alpha)$; note that by definition the unit normal \hat{n} to this plane coincide with the unit tangent along the string, i.e.

$$\hat{n} = \hat{t}. \quad (35)$$

From (34) it follows that at statistical steady state, the string φ satisfies

$$\varphi(\alpha) = \langle x \rangle_{P(\alpha)}, \quad (36)$$

where $\langle \cdot \rangle_{P(\alpha)}$ denotes the expectation with respect to (34). Together with (35), (36) specifies completely the string φ and the associated planes P . These equations will actually be derived in section 12 in the context of adaptive sampling techniques where it will be shown that the family of planes defined by (36) emerges as a reaction coordinate to describe the transition which is optimal within a certain class. Here let us simply note that (36) is a natural finite temperature generalization of the concept of MEP. Note first that, letting $\beta \rightarrow \infty$ in (36) and assuming that the potential V is smooth, this equation reduces to

$$\varphi(\alpha) = \arg \min V(x) \text{ in } P(\alpha). \quad (37)$$

This equation is equivalent to (25), i.e. φ converges to a MEP as $\beta \rightarrow \infty$. At finite temperature, the path satisfying (36) can be thought of as a MEP in some thermally averaged potential in which the small features of the energy below the thermal scale have been smoothed out. This is apparent from (31). In fact, φ defines the center of a tube in configuration space whose width may for instance be characterized by the covariance matrix of $\tilde{\varphi}$:

$$\text{cov}(\tilde{\varphi}) = \langle (x - \varphi)(x - \varphi)^T \rangle_{P(\alpha)}. \quad (38)$$

The finite temperature string method averages over details of the potential within this tube and thereby accounts for entropic effects. The free energy along the path can be defined as

$$\begin{aligned} F(\alpha) &= -\beta^{-1} \log \int_{P(\alpha)} e^{-\beta V} d\sigma(x) \\ &= -\beta^{-1} \log \int_{\mathbb{R}^n} e^{-\beta V} \delta(\hat{t} \cdot (x - \varphi)) dx. \end{aligned} \quad (39)$$

How to interpret this free energy and derive from it the rates for the transition will be explained in sections 11 and 12.

In practice, (32) is solved similarly as (29) except that a collection of replica of $\tilde{\varphi}$ must be evolved on each plane $P(\alpha)$, and the string φ has to be approximated by arithmetic average over these replicas. These calculations can be easily parallelized. Finally, note that in (32) the average force on the string is computed using the restricted Gibbs ensemble on the hyperplanes normal to the string. One may think of other ways of computing the averaged forces.

7 Reaction coordinates and free energy

The traditional belief in the study of rare events is that a transition between two metastable sets, B_1 and B_2 , can be described by a well-chosen reaction coordinate and the free energy associated with it. In this section we give a first discussion of these concepts and we will revisit them in section 11.

A reaction coordinate is a function $q(x)$ whose level sets $q(x) = cst$ foliate the configuration space and specify the advancement of the transitions between B_1 and B_2 . In the simplest setting $q(x)$ is a scalar-valued function and we will focus on this case first. The generalization to vector-valued functions is straightforward and will be discussed in the next sections. Given a reaction coordinate $q(x)$, there are two natural ways to define a free energy associated with it. The first is based on the marginal probability density in the variable q . This is the most standard free energy, and it is the one that will prove to have the right dynamical content provided that the right reaction coordinate

is chosen. The second free energy is based on the probability density of the surface $q(x) = q$. This free energy is more relevant in the context of TST. Since it is important to understand the differences between them, we discuss both.

Given $q(x)$, the first natural way to define a free energy is as

$$F(q) = -\beta^{-1} \log \bar{\rho}(q), \quad (40)$$

where

$$\bar{\rho}(q) = Z^{-1} \int_{\mathbb{R}^n} e^{-\beta V(x)} \delta(q(x) - q) dx. \quad (41)$$

$\bar{\rho}$ is the marginal of the equilibrium density ρ in the variable q and (40) implies that it can also be written as $\bar{\rho}(q) = e^{-\beta F(q)}$. The introduction of F and $\bar{\rho}$ allows us to factorize the expectation of $A(x)$ with respect to $Z^{-1}e^{-\beta V}$. Indeed, $\langle A \rangle$ can be expressed as

$$\begin{aligned} \langle A \rangle &= Z^{-1} \int_{\mathbb{R}^n} A(x) e^{-\beta V(x)} dx \\ &= \int_{\mathbb{R}} \langle A \rangle_{q(x)=q} e^{-\beta F(q)} dq, \end{aligned} \quad (42)$$

where $\langle A \rangle_{q(x)=q}$ is average of $A(x)$ in the surface $q(x) = q$, with respect to $e^{-\beta V} \delta(q(x) - q)$ properly normalized. Quite remarkably, the derivative of F can be expressed in terms of an average over the density $Z^{-1}e^{-\beta V(x)}$ restricted in the surface $q(x) = q$. To see this differentiate (40) with respect to q , use $|\nabla q(x)|^2 \delta'(q(x) - q) = \nabla q \cdot \nabla \delta(q(x) - q)$, and integrate by parts in x to obtain

$$F'(q) = \left\langle \frac{\nabla q \cdot \nabla V}{|\nabla q|^2} - \beta^{-1} \nabla \cdot \left(\frac{\nabla q}{|\nabla q|^2} \right) \right\rangle_{q(x)=q}. \quad (43)$$

$F'(q)$ is usually referred to as the mean force.

We will see below that the free energy in (40) is the one with the right dynamical content provided that one uses the right reaction coordinate $q(x)$. Indeed if $x(t)$ is a trajectory in the system, the dynamics of the coarse variable $q(x(t))$ is driven by the mean force plus appropriate thermal noise. We stress however that this interpretation of $F(q)$ is different from the most direct (and also most common) one where F is mostly used to identify a good dividing surface as the surface $q(x) = q_*$ where $F(q)$ reaches a maximum (i.e. such that the mean force vanishes on it, $F'(q_*) = 0$). This direct interpretation is in fact incorrect, as can be seen as follows. Suppose that $q(x)$ is a reaction coordinate and let $f(z)$ be a scalar function, strictly monotonous, $f'(z) \neq 0$. Then

$$\bar{q}(x) = f(q(x)) \quad (44)$$

is as good a reaction coordinate as $q(x)$ since the level sets of both functions coincide – we shall refer to (44) as a gauge transformation. However, if one denotes by $\bar{F}(\bar{q})$ the free energy in \bar{q} , it is easy to see that

$$\begin{aligned}
\bar{F}(f(q)) &= -\beta^{-1} \log \int_{\mathbb{R}^n} e^{-\beta V(x)} \delta(\bar{q}(x) - f(q)) dx + \beta^{-1} \log Z \\
&= -\beta^{-1} \log \int_{\mathbb{R}^n} e^{-\beta V(x)} (f'(q(x)))^{-1} \delta(q(x) - q) dx + \beta^{-1} \log Z \quad (45) \\
&= F(q) + \beta^{-1} \log f'(q),
\end{aligned}$$

i.e. $\bar{F}(f(q)) \neq F(q)$, the free energy is not left invariant by the gauge transformation in (44). In particular, if q_* is a critical point (minimum or maximum) of $F(q)$, i.e. $F'(q_*) = 0$, one has

$$\bar{F}'(f(q_*)) = \beta^{-1} f''(q_*) / (f'(q_*))^2 \neq 0, \quad (46)$$

which is not equal to zero in general. This is a serious flaw of the argument which identifies the critical points of the free energy in (40) as metastable sets or transition states.

The second natural way of defining a free energy is as follows (compare (40))

$$\begin{aligned}
G(q) &= -\beta^{-1} \log Z^{-1} \int_{q(x)=q} e^{-\beta V(x)} d\sigma(x) \\
&= -\beta^{-1} \log Z^{-1} \int_{\mathbb{R}^n} e^{-\beta V(x)} |\nabla q(x)| \delta(q(x) - q) dx. \quad (47)
\end{aligned}$$

With this definition $e^{-\beta G(q)}$ is the probability density of the surface $q(x) = q$, and it is easy to see that $G(q)$ is gauge invariant: $\bar{G}(f(q)) = G(q)$ if $\bar{q} = f(q)$ and \bar{G} is the free energy in \bar{q} . $G(q)$ is related to the standard free energy $F(q)$ defined in (40) via the relation

$$e^{-\beta G(q)} = \langle |\nabla q| \rangle_{q(x)=q} e^{-\beta F(q)} \quad (48)$$

Therefore one has (compare (42))

$$\langle A \rangle = \int \frac{\langle A \rangle_{q(x)=q}}{\langle |\nabla q| \rangle_{q(x)=q}} e^{-\beta G(q)} dq. \quad (49)$$

Also the mean force associated with G can be expressed as the following conditional average:

$$G'(q) = \langle \hat{n} \cdot \nabla V - \beta^{-1} \kappa \rangle_{q(x)=q} \langle |\nabla q| \rangle_{q(x)=q}^{-1}, \quad (50)$$

where $\hat{n} = \nabla q / |\nabla q|$ is the unit normal to the surface $q(x) = q$, and $\kappa = \nabla \cdot \hat{n}$ is the local mean curvature of this surface. Quite interestingly, this formula implies that $G'(q_*) = 0$ if

$$0 = \langle \hat{n} \cdot \nabla V - \beta^{-1} \kappa \rangle_{q(x)=q_*} \quad (51)$$

This equality is an averaged version of equation (18) for the optimal TST surface. In fact it is easy to show that if one optimizes the object function

in (17) over the level sets of a given reaction coordinate $q(x)$, then the optimal TST dividing surface is the level set of $q(x)$ where (51) is satisfied, i.e. it is a critical point of the free energy $G(q)$ (actually it can be shown that it is a maximum by checking that $G''(q_*) < 0$). This indicates that the definition in (47) for the free energy is the natural one in the context of TST.

We conclude this section by a severe warning. As mentioned before, one usually hopes that the reaction coordinate describes the advancement of the transitions between B_1 and B_2 , i.e. that it has some dynamical meaning which can be deduced by analyzing the properties of the free energy $F(q)$ or $G(q)$. But clearly this requires to pick the right reaction coordinate since the characteristics of $F(q)$ and $G(q)$ depend sensitively on $q(x)$. The results in this section leave completely open the question of how to choose $q(x)$. This question is highly nontrivial, particularly since the order parameter used to distinguish the metastable basins may be a bad reaction coordinate for the transition between these states. It is also particularly important since knowing the right reaction coordinate often means that we have a good intuitive understanding of how the reaction proceeds. How to choose $q(x)$ will be discussed in section 11.

8 Blue-moon sampling technique

It is remarkable that one can actually compute rather efficiently the free energy in (40) and the conditional expectation in (42) by using constrained simulations. This was first noted by Carter, Ciccotti, Hynes, and Kapral in [CCHK89] (see also [SC98]) where the blue-moon sampling technique was introduced. The idea is to use ergodicity and evaluate the conditional average in (43) for $F'(q)$ by time-averaging over a trajectory whose dynamics is constrained such that $q(x) = q$ and has as its equilibrium density $Z^{-1}e^{-\beta V(x)}$ restricted in this surface. F can then be estimated by integration of F' – a step referred to as thermodynamic integration [FS01]. In [CCHK89], it was proposed to compute the expectation in (43) via Hamiltonian dynamics simulations subject to proper constraint. Another possibility is to use

$$\dot{z} = -(\nabla V(z))^\perp - \beta^{-1} \nabla \cdot (\hat{n} \otimes \hat{n}) + \sqrt{2\beta^{-1}} \eta^\perp \quad (52)$$

where η is a white-noise, $a^\perp = a - (a \cdot \hat{n})\hat{n}$ with $\hat{n} = \nabla q / |\nabla q|$ denotes the projection of a into the hypersurface defined by $q(x) = q$. The term $-\beta^{-1} \nabla \cdot (\hat{n} \otimes \hat{n})$ is a spurious drift term that arises since the products in η^\perp are interpreted in Itô sense. It is easy to check that (52) has the right equilibrium density, i.e. if $q(z(0)) = q$, then

$$\langle A \rangle_{q(x)=q} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(z(t)) dt, \quad (53)$$

offers a practical way to evaluate the conditional expectations in (42) and (43). (52) can be solved using a time-splitting scheme such as

$$\begin{aligned} z^* &= z^n - (\nabla V(z^n))^\perp \Delta t + \sqrt{2\beta^{-1}} (\eta^n)^\perp \\ z^{n+1} &= z^* + \lambda \nabla \cdot (\hat{n}(z^n) \otimes \hat{n}(z^n)). \end{aligned} \quad (54)$$

Here Δt is the time-step, the η^n 's are independent identically distributed Gaussian random variables with mean zero and variance Δt , and λ is a scalar quantity such that

$$q(z^{n+1}) = q(z^n). \quad (55)$$

A scheme based on (54) and (55) is very much in the spirit of the method SHAKE introduced in [RCB77] (see also [WCH01]).

It is also interesting to note that the blue moon procedure can be generalized to multi-dimensional reaction coordinates where one uses a set of function $q_1(x), \dots, q_m(x)$ instead of a single $q(x)$. The associated free energy $F(q_1, \dots, q_m)$ is the multidimensional analog of (40)

$$F(q_1, \dots, q_m) = -\beta^{-1} \log Z^{-1} \int_{\mathbb{R}^n} e^{\beta V(x)} \delta(q_1(x) - q_1) \cdots \delta(q_m(x) - q_m) dx \quad (56)$$

If the functions $q_1(x), \dots, q_m(x)$ satisfy

$$\nabla q_j \cdot \nabla q_k = 0 \quad \text{if } j \neq k, \quad (57)$$

i.e. they form an orthogonal set of curvilinear coordinates, then the gradient of the free energy in q_1, \dots, q_m can be expressed as

$$\frac{\partial F}{\partial q_j} = \left\langle \frac{\nabla q_j \cdot \nabla V}{|\nabla q_j|^2} - \beta^{-1} \nabla \cdot \left(\frac{\nabla q_j}{|\nabla q_j|^2} \right) \right\rangle_{q_1(x)=q_1, \dots, q_m(x)=q_m}. \quad (58)$$

This average can be obtained via constrained simulations, e.g. similar to (52). Of course, retrieving $F(q_1, \dots, q_m)$ from its gradient (58) will be more complicated than in the one-dimensional case. One possibility is to use methods such as NEB or the zero-temperature string method that only requires (58) as its input to at least determine saddle points and heteroclinic orbits on the free energy surface (that is, the minimum free energy paths, MFEPs). Since the free energy surface will in general be much smoother than the original potential, it will have much less critical points, and MFEPs may be relevant even in situations where MEPs are not.

9 Metadynamics

At the end of section 8 we suggested how to determine the MFEPs in a two-step procedure that would combine the blue moon sampling technique with methods like NEB or the zero temperature string method. Recently in [ILP03] (see also [LP02]), Iannuzzi, Laio, and Parrinello proposed an alternative technique, termed metadynamics, which permits to sample in a seamless way the

free energy surface in a set of prescribed reaction coordinates $q_1(x), \dots, q_m(x)$. The idea is to extend the phase-space so as to include the reaction coordinates as additional dynamical variables. [ILP03] works in the Hamiltonian context, but we will consider the case of overdamped dynamics which is better suited for analysis.

Metadynamics amounts to considering the thermally perturbed gradient flow on the extended energy

$$U(x, Q) = V(x) + \frac{1}{2}\mu|Q - q(x)|^2, \quad (59)$$

where $\mu > 0$ is a parameter to be prescribed later, and $Q = (Q_1, \dots, Q_m)$ are the additional variables corresponding to the reaction coordinates. Associated with (59) we take the following equations of motion:

$$\begin{cases} \dot{x} = -\delta\nabla V(x) + \delta\mu(Q - q(x))\nabla q(x) + \sqrt{2\beta^{-1}\delta}\eta, \\ \dot{Q} = -\mu(Q - q(x)) + \sqrt{2\beta^{-1}}\eta_q, \end{cases} \quad (60)$$

where η_q is a white-noise independent of η , and $\delta > 0$ is another adjustable parameter. We will consider the case when $\delta > 1$ corresponding to situations where there is more friction on $Q(t)$ than on $x(t)$ – i.e. $Q(t)$ are slower variables than $x(t)$.

For all $\delta > 0$, the equilibrium density function for (60) is

$$\rho_\mu(x, Q) = Z^{-1}(\mu\beta/2\pi)^{m/2} e^{-\beta V(x) - \frac{1}{2}\mu\beta|Q - q(x)|^2}, \quad (61)$$

where Z is the normalization factor of $e^{-\beta V(x)}$ alone, $Z = \int_{\mathbb{R}^n} e^{-\beta V(x)} dx$. The corresponding marginal density on Q is

$$\bar{\rho}_\mu(Q) = Z^{-1}(\mu\beta/2\pi)^{m/2} \int_{\mathbb{R}^n} e^{-\beta V(x) - \frac{1}{2}\mu\beta|Q - q(x)|^2} dx \quad (62)$$

Viewed as a function of Q , $(\mu\beta/2\pi)^{m/2} e^{-\frac{1}{2}\mu\beta|Q - q(x)|^2}$ converges weakly towards $\delta(Q - q(x))$ as $\mu \rightarrow \infty$. Therefore

$$\bar{\rho}_\mu(q) \rightarrow \bar{\rho}(q) \quad \text{as } \mu \rightarrow \infty, \quad (63)$$

where $\bar{\rho}(q)$ is the multi-dimensional analog of the marginal density in the variables q that we defined in (41). A similar argument shows that (62) is also the reduced density for the extended Hamiltonian system

$$\begin{cases} M\ddot{x} = -\nabla V(x) + \mu(Q - q(x))\nabla q(x) \\ M_q\ddot{Q} = -\mu(Q - q(x)), \end{cases} \quad (64)$$

corresponding to the extended Lagrangian

$$L(x, Q) = \frac{1}{2}(\dot{x}, M\dot{x}) - V(x) + \frac{1}{2}(\dot{Q}, M_q\dot{Q}) - \frac{1}{2}\mu|Q - q(x)|^2. \quad (65)$$

Here M_q is the mass tensor associated with the new variable Q . [ILP03] works in the Hamiltonian context and use (64), but we will stick to (60).

The idea in [ILP03] is to use (64) at large μ and large M_Q (with μ/M_Q small) to compute the free energy associated with $\bar{\rho}(q)$. Here we show that this free energy can also be computed from (60) at large μ and large δ with μ/δ small. When δ is large, the variables $Q(t)$ evolves much more slowly than $x(t)$ and only feel the average effect of the latter. The proper conditional probability density with which to average the right hand-side of (60) is (compare (61))

$$\rho_\mu(x|Q) = Z_\mu^{-1}(Q) e^{-\beta V(x) - \frac{1}{2}\mu\beta|Q-q(x)|^2}, \quad (66)$$

where

$$Z_\mu(Q) = \int_{\mathbb{R}^n} e^{-\beta V(x) - \frac{1}{2}\mu\beta|Q-q(x)|^2} dx, \quad (67)$$

is a normalization factor which guarantees that $\int_{\mathbb{R}^n} \rho_\mu(x|Q) dx = 1$ for all Q . Since

$$\begin{aligned} & -Z_\mu^{-1}(Q) \int_{\mathbb{R}^n} \mu(Q - q(x)) e^{-\beta V(x) - \frac{1}{2}\mu\beta|Q-q(x)|^2} dx \\ &= \beta^{-1} Z_\mu^{-1}(Q) \nabla_Q Z_\mu(Q) \\ &= \beta^{-1} \nabla_Q \log Z_\mu(Q), \end{aligned} \quad (68)$$

the limiting equations for $Q(t)$ (60) as $\delta \rightarrow \infty$ can be written as

$$\dot{Q} = -\nabla_Q F_\mu + \sqrt{2\beta^{-1}} \eta_q, \quad (69)$$

provided one defines

$$F_\mu(Q) = -\beta^{-1} \nabla_Q \log Z_\mu(Q). \quad (70)$$

This is a mollified version of the free energy since (67) and (70) imply that

$$F_\mu(Q) \rightarrow F(Q) \quad (71)$$

as $\mu \rightarrow \infty$ and therefore the limiting equation for $Q(t)$ as $\delta \rightarrow \infty$ and $\mu \rightarrow \infty$ (in this order) is

$$\dot{Q} = -\nabla_Q F + \sqrt{2\beta^{-1}} \eta_q, \quad (72)$$

Thus, by simulating (60) at large δ , μ , and monitoring the evolution of $Q(t)$ one can indeed sample the free energy landscape in the variables $q(x)$. But there is still one difficulty: (60) will display metastability if the original dynamics does, so this equation may not be practical. To fix this problem, Iannuzzi, Laio, and Parrinello suggest to use a technique introduced in [HTV94] and [WL01] and further modify the dynamics by including in (60) (or rather (64) in the original paper) an additional non-Markovian term which discourages the trajectory from going back to regions that it has already visited. For instance, one may modify (60) as

$$\begin{cases} \dot{x} = -\delta \nabla V(x) + \delta \mu (Q - q(x)) \nabla q(x) + \sqrt{2\beta^{-1}\delta} \eta, \\ \dot{Q} = -\mu(Q - q(x)) + \sqrt{2\beta^{-1}} \eta_q \\ \quad + A \int_0^t (Q(t) - Q(t')) e^{-|Q(t) - Q(t')|^2 / \Delta q^2} dt' \end{cases} \quad (73)$$

where A and Δq are adjustable parameters. Proceeding as before, it is easy to see that the limiting equation for Q as $\delta \rightarrow \infty$ and $\mu \rightarrow \infty$ is

$$\dot{Q} = -\nabla_Q F + A \int_0^t (Q(t) - Q(t')) e^{-|Q(t) - Q(t')|^2 / \Delta q^2} dt' + \sqrt{2\beta^{-1}} \eta_q, \quad (74)$$

Therefore, the memory term added in (73) is a term that fills up the potential well that the trajectory has already visited. In particular, if

$$U(q, t) = \frac{1}{2} A \Delta q^2 \int_0^t e^{-|q - Q(t')|^2 / \Delta q^2} dt', \quad (75)$$

then, as $t \rightarrow \infty$, $U(q, t) - U(q', t)$ converges to an estimate of $F(q') - F(q)$ when $\delta \gg \mu \gg 1$. The parameters A and Δq control the accuracy on the resolution of the free energy: as they are decreased, the resolution improves, but the convergence rate in time deteriorates. Given some accuracy requirement, estimating the optimal choice of parameters δ , μ , A , and Δq in a metadynamics calculation is a nontrivial question which we will leave aside.

10 Another type of metadynamics: the accelerated dynamics

The metadynamics calculation of section 9 gives the free energy of the set of reaction coordinates $q(x) = (q_1(x), \dots, q_m(x))$ in the specific gauge fixed by this choice. We wish now to propose an alternative type of metadynamics – the accelerated dynamics – which shed some light on the issue of choice of gauge. Given a set of reaction coordinates $q(x)$, consider

$$\begin{aligned} \dot{x} = & \delta P(-\nabla V(x) + \sqrt{2\beta^{-1}} \eta) \\ & + (1 - P)(-\nabla V(x) + \sqrt{2\beta^{-1}} \eta) + (\delta - 1)\beta^{-1} \nabla \cdot P, \end{aligned} \quad (76)$$

where

$$P = 1 - \sum_{j=1}^m \hat{n}_j \otimes \hat{n}_j, \quad \hat{n}_j = \nabla q_j / |\nabla q_j|. \quad (77)$$

P is the projector into the surface where $q(x) = cst$ (here for simplicity we assume that (57) holds). When $\delta = 1$, (76) reduces to the overdamped equation (4) (in time units where $\gamma = 1$). When $\delta > 1$, (76) corresponds to a dynamics where one has artificially accelerated the dynamics on $q(x) = cst$ while

keeping its original speed in the direction perpendicular to these surfaces. The term $(\delta - 1)\beta^{-1}\nabla \cdot P$ is again a spurious drift term from the interpretation of the noise terms in (76) in Itô sense. A set of equations analogous to (76) can be written in the Hamiltonian setting,

$$\begin{cases} \dot{x} = \delta P v + (1 - P)v \\ M\dot{v} = -\delta P \nabla V(x) - (1 - P)\nabla V(x), \end{cases} \quad (78)$$

but we will again stick to (76) as it is better suited for analysis. The equilibrium density function of both (76) and (78) is

$$\rho(x) = Z^{-1} e^{-\beta V(x)} \quad (79)$$

for all values of $\delta > 0$.

Let $Q(t) = q(x(t))$. It is easy to see that $Q(t)$ satisfies

$$\dot{Q} = -\nabla V \cdot \nabla q + \beta^{-1} \Delta q + \sqrt{2\beta^{-1}} (\eta \cdot \nabla) q. \quad (80)$$

(80) is not closed. But if we now assume that $\delta \gg 1$, the motion on the surface $q(x) = cst$ can be averaged out. It is shown in Appendix B that in the limit as $\delta \rightarrow \infty$ (80) reduces to the following closed equation for $Q(t)$:

$$\dot{Q}_j = -a_j(Q) \frac{\partial F}{\partial Q_j} + \beta^{-1} \frac{\partial a_j}{\partial Q_j} + \sqrt{2\beta^{-1} a_j(Q)} \eta_j, \quad j = 1, \dots, m. \quad (81)$$

Here

$$F(q) = -\beta^{-1} \log Z^{-1} \int_{R^n} e^{-\beta V} \delta(q_1 - q_1(x)) \dots \delta(q_m - q_m(x)) dx, \quad (82)$$

is the free energy in the variables q and

$$a_j(q) = \langle |\nabla q|^2 \rangle_{q(x)=q} \quad (83)$$

where $\langle \cdot \rangle_{q(x)=q}$ denotes expectation in $q(x) = q$ with respect to $e^{-\beta V}$ properly normalized. Since the forward operator associated with (81) is

$$L_q \rho(q) = \sum_{j=1}^m \frac{\partial}{\partial q_j} \left(a_j(q) \frac{\partial F}{\partial q_j} \rho + \beta^{-1} a_j(q) \frac{\partial \rho}{\partial q_j} \right) \quad (84)$$

it is easy to see that $e^{-\beta F}$ is the equilibrium density for (81).

In general, (81) is not in the standard form (72) since $a_j \neq 1$. But in contrast to what happens with (72), the dynamics associated with (81) is left invariant by a gauge transformation. Consider the multidimensional analog of (44)

$$\bar{q}_j(x) = f_j(q_j(x)), \quad j = 1, \dots, m. \quad (85)$$

$\bar{q}(x)$ and $q(x)$ have the same level sets, i.e. they define the same reaction coordinate. And it is easy to see that the limiting equation similar to (81) one obtains for $\bar{Q}(t) = \bar{q}(x(t))$ guarantees that

$$\bar{Q}_j(t) = f_j(Q_j(t)). \quad (86)$$

This is in fact obvious since (76) is gauge invariant. In addition if

$$a_j(q) = 1, \quad (87)$$

then (81) takes the standard form

$$\dot{Q}_j = -\frac{\partial F}{\partial Q_j} + \sqrt{2\beta^{-1}} \eta_j, \quad j = 1, \dots, m. \quad (88)$$

(87) can always be satisfied with suitable gauge transformation (85) on $q(x)$ such that

$$\sqrt{a_j(q)} f'_j(q) = 1, \quad j = 1, \dots, m. \quad (89)$$

These considerations suggest that given the appropriate set of reaction coordinates $q(x) = (q_1(x), \dots, q_m(x))$ (but regardless of the choice of gauge), the limiting accelerated dynamics in (81) may still capture some feature of the original, unaccelerated dynamics in (4). In section 11 we will show that this is indeed the case and that the transition rates between metastable basins are independent of the value of $\delta \geq 1$ – and in particular are captured by the limiting equation in (81) – provided that the right reaction coordinates are used.

Finally, we note that (76) can in principle be used to compute the free energy by using a flooding procedure similar to one used in [LP02, ILP03]. The idea is to modify (76) as

$$\begin{aligned} \dot{x} = & \delta P(-\nabla V(x) + \sqrt{2\beta^{-1}} \eta) \\ & + (1 - P)(-\nabla V(x) + \sqrt{2\beta^{-1}} \eta) + (\delta - 1)\beta^{-1} \nabla \cdot P \\ & + A \sum_{j=1}^m \nabla q_j(x) \int_0^t (q_j(x(t)) - q_j(x(t'))) e^{-|q(x(t)) - q(x(t'))|^2 / \Delta q^2} dt', \end{aligned} \quad (90)$$

where A and Δq are adjustable parameters similar to the ones in (74). As $\delta \rightarrow \infty$, (90) leads to the following limiting equation for $Q(t)$

$$\begin{aligned} \dot{Q}_j = & -a_j(Q) \frac{\partial F}{\partial Q_j} + \beta^{-1} \frac{\partial a_j}{\partial Q_j} + \sqrt{2\beta^{-1} a_j(Q)} \eta_j \\ & + A a_j(Q) \int_0^t (Q_j(t) - Q_j(t')) e^{-|Q(t) - Q(t')|^2 / \Delta q^2} dt', \end{aligned} \quad (91)$$

or, in the gauge where $a_j(q) = 1$,

$$\dot{Q}_j = -\frac{\partial F}{\partial Q_j} + \sqrt{2\beta^{-1}} \eta_j + A \int_0^t (Q_j(t) - Q_j(t')) e^{-|Q(t) - Q(t')|^2 \Delta q^2} dt'. \quad (92)$$

It follows that we can estimate the free energy by monitoring the evolution of

$$U(q, t) = \frac{1}{2} A \Delta q^2 \int_0^t e^{-|q - q(x(t'))|^2 / \Delta q^2} dt', \quad (93)$$

since $U(q, t) - U(q', t)$ converges as $t \rightarrow \infty$, to an estimate of $F(q') - F(q)$ when $\delta \gg 1$.

11 Choosing the right reaction coordinates

The main limitation of the blue moon sampling technique discussed in section 8 or the method based on metadynamics discussed in sections 9 and 10 is that the reaction coordinate $q(x)$ must be specified beforehand. Here we discuss how to get around this problem and how to determine “optimal” reaction coordinates. We consider first the overdamped dynamics in (4) and then we will indicate how to generalize the concepts to other types of dynamics.

We start by giving a more precise definition of metastability. A system will be metastable with respect to the dynamics in (4) if the eigenvalues of the backward operator associated with this equation contain a spectral gap, i.e. $\lambda_1/\lambda_2 \ll 1$ where λ_1 is the smallest nonzero eigenvalue of $L = -\nabla V \cdot \nabla + \beta^{-1} \Delta$ and λ_2 is the second smallest eigenvalue. Clearly, this definition encompasses the concept of separation of time scales between the relaxation time within the sets B_1 and B_2 , and the transition time between these sets. In fact it is well known that λ_1 is a good approximation of the relaxation rate,

$$\lambda_1 \approx k_{12} + k_{21}. \quad (94)$$

The error in this expression is $O(\lambda_1/\lambda_2)$. (94) combined with the relation $N_1/N_2 = k_{21}/k_{12}$ allows us to determine the rates.

It is also known [HMS02, SH02, BEGK03a, BEGK03b] that, if $\lambda_1/\lambda_2 \ll 1$, we have

$$\lambda_1 \approx Z^{-1} \frac{N_1 + N_2}{N_1 N_2} \int_{\Omega} |\nabla q|^2 e^{-\beta V} dx \quad (95)$$

where $\Omega = \mathbb{R}^n / (B_1 \cup B_2)$ and $q(x)$ is the solution of the backward equation

$$0 = -\nabla V \cdot \nabla q + \beta^{-1} \Delta q \quad (96)$$

in Ω with Dirichlet boundary conditions $q|_{\partial B_1} = 0$, $q|_{\partial B_2} = 1$. (95) holds because the eigenfunction $\varphi_1(x)$ associated with $\lambda_1 = \int_{\mathbb{R}^n} |\nabla \varphi_1|^2 e^{-\beta V} dx$ is approximately constant over regions that contain B_1 and B_2 and can be approximated by q appropriately rescaled as in (95).

In addition to (95) the solution q of (96) has the following remarkable properties which qualify it as an “optimal” reaction coordinate. First the

level sets of $q(x)$ are the isoprobability surfaces for the transition between B_1 and B_2 because

$$q(x) = \text{Prob}\{X(t, x) \text{ reaches } B_2 \text{ before } B_1\}, \quad (97)$$

where $X(t, x)$ denotes the solution of (4) with initial condition $X(0, x) = x$. Since a reaction coordinate is supposed to describe the advancement of the transition between B_1 and B_2 at the coarse-grained level where all the information within the level set $q(x) = cst$ is averaged out, the function $q(x)$ satisfying (97) is arguably as good as it can get with any reaction coordinate.

Because they are isoprobability surfaces, the level sets of $q(x)$ have also the property that the probability density of the transition paths between B_1 and B_2 restricted to the surface $q(x) = q$ is precisely the equilibrium density $e^{-\beta V(x)}$ restricted to this surface. Assume that the equilibrium density on each of these surfaces is localized, and let T be a tube in configuration space such that the intersection of this tube with each surface $q(x) = q$ supports most of the probability on this surface, i.e. such that

$$\int_T e^{-\beta V(x)} |\nabla q(x)| \delta(q(x) - q) dx \approx \int_{\mathbb{R}^n} e^{-\beta V(x)} |\nabla q(x)| \delta(q(x) - q) dx \quad (98)$$

for each q . Then with probability close to one the transition paths between B_1 and B_2 stay in T . In fact, both the level set of $q(x)$ and the tube T should be used to characterize the reaction coordinate of the transitions between B_1 and B_2 . Note that multiple tubes instead of a single one may have to be introduced when there are more than one channel for the transition. The generalization to this case is straightforward.

Another nice property of the solution of (96) offers possible ways of generalizing the concept to other type of dynamics, or to use more than one reaction coordinate. As shown in Appendix C, the relaxation rate in (95) is precisely the one which one obtains from the limiting accelerated equation in (81). But since the solution of (96) is also the minimizer of

$$I_1 = \int_{\Omega} |\nabla q|^2 e^{-\beta V} dx, \quad (99)$$

this indicates that, among all $q(x)$, the solution of (96) is the one that minimizes the transition rate of the accelerated dynamics (in which case it is in fact the transition rate of the original, un-accelerated dynamics to order $O(\lambda_1/\lambda_2)$). This property should be generic and independent of the type of dynamics or the number of reaction coordinates one chooses to describe the transition (though, in the present case, (95) clearly indicates that a scalar-valued $q(x)$ is enough). We believe it could be taken as a starting point for determining adaptively optimal reaction coordinates for arbitrary dynamics by generalizing the numerical procedure for identifying $q(x)$ and T which we discuss next.

12 Adaptive sampling techniques

We now turn the observations of section 11 into practical tools to determine the “optimal” reaction coordinate and the transition tube. These can be viewed as a way to do adaptive blue moon sampling.

Besides being the minimizer of I_1 , q is also the maximizer of

$$\begin{aligned} I_2 &= \int_0^1 \left(\int_{q(x)=q} |\nabla q| e^{-\beta V} d\sigma(x) \right)^{-1} dq \\ &= \int_0^1 \left(\int_{\mathbb{R}^n} |\nabla q|^2 e^{-\beta V} \delta(q(x) - q) dx \right)^{-1} dq, \end{aligned} \quad (100)$$

with $\sup_q I_2 = (\inf_q I_1)^{-1}$. This is shown in Appendix C. (100) has the advantage of being invariant under the gauge transformation in (44), $q(x) \rightarrow \bar{q}(x) = f(q(x))$. Note also that in the gauge when $\langle |\nabla q|^2 \rangle_{q(x)=q} = 1$ (i.e. when $a = 1$ and the alternative metadynamics limiting equation is in the standard form (88)), (100) can be written as

$$I_2 = \int_0^1 e^{\beta F(q)} dq. \quad (101)$$

To turn these observations into a computational procedure for determining the optimal reaction coordinate, we can maximize the functional I_2 over specific classes of $q(x)$. The simplest choice is to assume that the level sets of $q(x)$ are locally planar in the tube T specified by (98). It is convenient to specify these planes by a parameterized curve – i.e. a string – $\varphi(\alpha)$, with $\alpha \in [0, 1]$, such that (i) the plane labeled by α , which we will denote by $\bar{P}(\alpha)$ contains the point $x = \varphi(\alpha)$, (ii) the unit normal of $\bar{P}(\alpha)$ is $\hat{n}(\alpha)$, and (iii) the point $\varphi(\alpha)$ is the mean position in the plane, i.e.

$$\varphi(\alpha) = \langle x \rangle_{\bar{P}(\alpha)}, \quad (102)$$

where $\langle \cdot \rangle_{\bar{P}(\alpha)}$ denotes the expectation with respect to $e^{-\beta V(x)} \delta(\hat{n} \cdot (x - \varphi))$ properly normalized. The localization approximation requires that

$$(\hat{n} \cdot \varphi_\alpha)^2 \gg \langle (\hat{n}_\alpha \cdot (x - \varphi))^2 \rangle_{\bar{P}(\alpha)} \quad (103)$$

(103) relates the width of the transition tube T to the local curvature of the string and it is required since otherwise the probability on the planes would include regions where these planes intersect, thereby invalidating the local planar assumption for the level sets of $q(x)$. When (103) is satisfied, it is shown in Appendix C that (100) reduces to

$$\begin{aligned} I_2 &= \int_0^1 \left(\int_{\bar{P}(\alpha)} e^{-\beta V} d\sigma(x) \right)^{-1} \hat{n} \cdot \varphi_\alpha d\alpha \\ &= \int_0^1 \left(\int_{\mathbb{R}^n} e^{-\beta V} \delta(\hat{n} \cdot (x - \varphi)) dx \right)^{-1} \hat{n} \cdot \varphi_\alpha d\alpha, \end{aligned} \quad (104)$$

subject to (102). Furthermore, the curve maximizing this functional satisfies

$$0 = \langle (\hat{n} \cdot \nabla V)(x - \varphi) \rangle_{\bar{P}(\alpha)} \quad (105)$$

Comparing (105) with (22) we see that the plane along the string where $\langle \hat{n} \cdot V \rangle_{\bar{P}} = 0$ coincides with the optimal TST plane. Furthermore, within the approximation in (103), it is shown in Appendix C that (105) is equivalent to

$$\varphi_\alpha / |\varphi_\alpha| = \hat{n}, \quad (106)$$

(106) indicates that the unit tangent $\hat{t} = \varphi_\alpha / |\varphi_\alpha|$ along the string coincides with the unit normal \hat{n} to the planes. It is remarkable that (106) and (102) are the same set of conditions as (36) and (35). In other words, under the localization assumption, the planes perpendicular to the string determined in a seamless way by the finite temperature string method coincide with the level sets of the optimal reaction coordinate determined by optimization of (100) within the class of $q(x)$ with (locally) planar level sets.

13 Concluding remarks

We have reviewed a variety of techniques for determining transition pathways and rates in complex systems. We left aside many important techniques such as hyperdynamics [SV00, Vot97], stochastic difference equation [EGC02, ECGS03], or clustering algorithms based on spectral analysis of the Perron-Frobenius operator [FSDC02, SHD02]. Most of the techniques we discussed require prior knowledge of the final states for the transition, but we have also included techniques such as metadynamics which does not require such information. Some of these techniques such as NEB and the zero-temperature string method are predicated on the notion of MEP, and therefore are only useful when the transition proceeds through isolated saddle points of the potential energy landscape. Others, such as TPS, blue moon sampling and the finite-temperature string method, should in principle work even when the energy landscape is rather rough with a dense set of critical points, for which the notion of MEP is no longer relevant. Some of the techniques such as blue moon sampling and metadynamics require specifying the reaction coordinates beforehand. Others, such as TPS, the finite-temperature string method and adaptive sampling techniques, do not require knowing the reaction coordinates beforehand.

From a conceptual point of view, the classical TST is based on one dividing surface. Consequently it is not able to give an accurate description of diffusive barrier crossing. The Bennett-Chandler procedure introduces two more surfaces, ∂B_1 and ∂B_2 which are the boundaries of the metastable sets. This gives a well-defined notion of transmission coefficient but the transmission coefficient may be hard to evaluate in practice because of the difficulties

mentioned in section 2. These difficulties can in principle be overcome by introducing the optimal foliation between the two metastable sets. We believe that practical techniques can be developed using this framework, along the lines discussed in section 11.

We have seen that a fairly coherent theoretical framework can be developed for the high friction limit. The main remaining theoretical difficulty is associated with pure Hamiltonian dynamics. This is the case of most practical interest. We hope that some of the notions reviewed here can at least serve as a starting point for developing approximations for transition pathways and transition rates in that case.

Acknowledgments

In our understanding of the various techniques we reviewed we benefited from discussions with the authors of these techniques. We thank them all. We are especially grateful to Giovanni Ciccotti and Ray Kapral for their comments about the material in sections 3 and 11. The work of E is supported in part by ONR grant N00014-01-1-0674. The work of Vanden-Eijnden is supported in part by NSF grants DMS01-01439, DMS02-09959 and DMS02-39625.

A Derivation of (18) and (21)

Suppose that the dividing surface ∂A is parameterized as $0 = q(x)$, i.e. it is the zero level set of some function $q(x)$. Then (17) can be written as

$$I_0 = \int_{\mathbb{R}^n} |\nabla q| e^{-\beta V} \delta(q(x)) dx. \quad (107)$$

Denote by δq the variation in q , the first variation of the above functional with respect to $q(x)$ is

$$\delta I_0 = \int_{\mathbb{R}^n} e^{-\beta V(x)} \hat{n} \cdot \nabla \delta q \delta(q(x)) dx + \int_{\mathbb{R}^n} e^{-\beta V(x)} |\nabla q| \delta q \delta'(q(x)) dx. \quad (108)$$

where $\hat{n} = \nabla q / |\nabla q|$ is the unit normal of $q(x) = 0$. Since

$$\delta'(q(x)) = |\nabla q|^{-1} \hat{n} \cdot \nabla \delta(q(x)), \quad (109)$$

we obtain after integration by parts for the second integral in (108)

$$\begin{aligned} \delta I_0 &= \int_{\mathbb{R}^n} e^{-\beta V(x)} \hat{n} \cdot \nabla \delta q \delta(q(x)) dx + \int_{\mathbb{R}^n} e^{-\beta V(x)} \delta q \hat{n} \cdot \nabla \delta(q(x)) dx \\ &= \int_{\mathbb{R}^n} e^{-\beta V(x)} \hat{n} \cdot \nabla \delta q \delta(q(x)) dx \\ &\quad + \int_{\mathbb{R}^n} e^{-\beta V(x)} \left((\beta \nabla V \cdot \hat{n} - \nabla \cdot \hat{n}) \delta q - \hat{n} \cdot \nabla \delta q \right) \delta(q(x)) dx \\ &= \int_{\mathbb{R}^n} e^{-\beta V(x)} (\beta \nabla V \cdot \hat{n} - \nabla \cdot \hat{n}) \delta q \delta(q(x)) dx. \end{aligned} \quad (110)$$

Therefore, on $q(x) = 0$, the minimizer of I_0 satisfies (18).

The derivation of (21) from (20) is similar. The first variation of this functional with respect to \hat{n} and b gives

$$\begin{aligned}\delta I &= \int_{\mathbb{R}^n} e^{-\beta V} (\delta \hat{n}^\perp \cdot x - \delta b) \delta'(\hat{n} \cdot x - b) dx \\ &= \int_{\mathbb{R}^n} e^{-\beta V} (\delta \hat{n} \cdot x^\perp - \delta b) \hat{n} \cdot \nabla \delta(\hat{n} \cdot x - b) dx \\ &= \beta \int_{\mathbb{R}^n} \hat{n} \cdot \nabla V e^{-\beta V} (\delta \hat{n} \cdot x^\perp - \delta b) \delta(\hat{n} \cdot x - b) dx,\end{aligned}\tag{111}$$

where for any $z \in \mathbb{R}^n$, $z^\perp = z - (z \cdot \hat{n})\hat{n}$ is the in-plane projection of z , and we used the property that the admissible variations of \hat{n} need to preserve the normalization $|\hat{n}| = 1$, i.e. they need to be perpendicular to \hat{n} . Collecting the terms proportional to $\delta \hat{n}$ and δb in (111) we arrive at (21).

B Derivation of (81)

It is a standard result [Pap76] that in the limit as $\delta \rightarrow 0$, (80) converges to

$$\dot{Q}_j = b_j(Q) + \sqrt{2\beta^{-1}a_j(Q)} \eta_j, \quad j = 1, \dots, m.\tag{112}$$

where $a_j(q)$ is given by (83) and

$$b_j(q) = \langle -\nabla V \cdot \nabla q_j + \beta^{-1} \Delta q_j \rangle_{q(x)=q}.\tag{113}$$

Since

$$\begin{aligned}& \int_{\mathbb{R}^n} (-\nabla V \nabla q_j + \beta^{-1} \Delta q_j) e^{-V} \delta(q_1 - q_1(x)) \cdots \delta(q_m - q_m(x)) dx \\ &= \beta^{-1} \int_{\mathbb{R}^n} \nabla \cdot (e^{-\beta V} \nabla q_j) \delta(q_1 - q_1(x)) \cdots \delta(q_m - q_m(x)) dx \\ &= -\beta^{-1} \int_{\mathbb{R}^n} e^{-\beta V} \nabla q_j \cdot \nabla (\delta(q_1 - q_1(x)) \cdots \delta(q_m - q_m(x))) dx \\ &= \beta^{-1} \frac{\partial}{\partial q_j} \int_{\mathbb{R}^n} e^{-\beta V} |\nabla q_j|^2 \delta(q_1 - q_1(x)) \cdots \delta(q_m - q_m(x)) dx\end{aligned}\tag{114}$$

where we used (57), one has

$$b_j(q) e^{-\beta F} = \beta^{-1} \frac{\partial}{\partial q_j} (a_j(q) e^{-\beta F}),\tag{115}$$

or, equivalently,

$$b_j(q) = -a_j(q) \frac{\partial F_j}{\partial q_j} + \beta^{-1} \frac{\partial a_j}{\partial q_j}.\tag{116}$$

Inserting this equality in (112) gives (81).

C The calculations of section 11

To show that the relaxation rate in (95) is precisely the one obtained from the accelerated dynamics equation in (81), note that since

$$\int_{\mathbb{R}^n} |\nabla q|^2 e^{-\beta V} \delta(q - q(x)) dx = \int_{q(x)=q} |\nabla q| e^{-\beta V} d\sigma(x) = cst \quad (117)$$

for the solution of (96), it follows that the relaxation rate in (95) can also be expressed as

$$\begin{aligned} \lambda_1 &= Z^{-1} \frac{N_1 + N_2}{N_1 N_2} \int_0^1 \left(\int_{q(x)=q} |\nabla q| e^{-\beta V} d\sigma(x) \right)^{-1} dq \\ &= \frac{N_1 + N_2}{N_1 N_2} \int_0^1 a^{-1}(q) e^{\beta F(q)} dq \end{aligned} \quad (118)$$

where $a(q)$ is given by (83). (118) is precisely the asymptotic expression for the relaxation rate associated with (81) when B_1 and B_2 are metastable (see e.g. [Gar89]).

To show that the minimizer of I_1 is a maximizer of I_2 , we compute the first variation of I_2 with respect to $q(x)$. Starting from the following expression equivalent to (100)

$$I_2 = \int_0^1 \left(\int_{\mathbb{R}^n} |\nabla q|^2 e^{-\beta V} \delta(q(x) - q) dx \right)^{-1} dq, \quad (119)$$

we obtain

$$\begin{aligned} \delta I_2 &= -2 \int_0^1 A^{-2}(q) \int_{\mathbb{R}^n} \nabla \delta q \cdot \nabla q e^{-\beta V} \delta(q(x) - q) dx dq \\ &\quad - \int_0^1 A^{-2}(q) \int_{\mathbb{R}^n} |\nabla q|^2 e^{-\beta V} \delta q \delta'(q(x) - q) dx dq \end{aligned} \quad (120)$$

where

$$A(q) = \int_{\mathbb{R}^n} |\nabla q|^2 e^{-\beta V} \delta(q(x) - q) dx = \int_{q(x)=q} |\nabla q| e^{-\beta V} d\sigma(x). \quad (121)$$

By integration by parts of the first integral in (120) we obtain

$$\begin{aligned} \delta I_2 &= 2 \int_0^1 A^{-2}(q) \int_{\mathbb{R}^n} \delta q \nabla \cdot (e^{-\beta V} \nabla q) \delta(q(x) - q) dx dq \\ &\quad + \int_0^1 A^{-2}(q) \int_{\mathbb{R}^n} |\nabla q|^2 e^{-\beta V} \delta q \delta'(q(x) - q) dx dq \end{aligned} \quad (122)$$

This can also be written as

$$\begin{aligned} \delta I_2 &= 2 \int_0^1 A^{-2}(q) \int_{\mathbb{R}^n} \delta q \nabla \cdot \left(e^{-\beta V} \nabla q \right) \delta(q(x) - q) dx dq \\ &\quad - \int_0^1 A^{-2}(q) \frac{\partial}{\partial q} \int_{\mathbb{R}^n} |\nabla q|^2 e^{-\beta V} \delta q \delta(q(x) - q) dx dq, \end{aligned} \quad (123)$$

which after integration by parts in q of the second integral gives

$$\begin{aligned} \delta I_2 &= 2 \int_0^1 A^{-2}(q) \int_{\mathbb{R}^n} \delta q \nabla \cdot \left(e^{-\beta V} \nabla q \right) \delta(q(x) - q) dx dq \\ &\quad - 2 \int_0^1 A^{-3}(q) \frac{\partial A}{\partial q} \int_{\mathbb{R}^n} |\nabla q|^2 e^{-\beta V} \delta q \delta(q(x) - q) dx dq. \end{aligned} \quad (124)$$

Since

$$\begin{aligned} \frac{\partial A}{\partial q} &= - \int_{\mathbb{R}^n} |\nabla q|^2 e^{-\beta V} \delta'(q(x) - q) dx \\ &= - \int_{\mathbb{R}^n} e^{-\beta V} \nabla q \cdot \nabla \delta(q(x) - q) dx \\ &= \int_{\mathbb{R}^n} \nabla \cdot \left(e^{-\beta V} \nabla q \right) \delta(q(x) - q) dx. \end{aligned} \quad (125)$$

it follows that if

$$\nabla \cdot \left(e^{-\beta V} \nabla q \right) = e^{-\beta V} |\nabla q|^2 H(q) \quad (126)$$

for an arbitrary H , then

$$\begin{aligned} \delta I_2 &= 2 \int_0^1 A^{-2}(q) H(q) \int_{\mathbb{R}^n} \delta q |\nabla q|^2 \delta(q(x) - q) dx dq \\ &\quad - 2 \int_0^1 A^{-2}(q) H(q) \int_{\mathbb{R}^n} |\nabla q|^2 e^{-\beta V} \delta q \delta(q(x) - q) dx dq \\ &= 0. \end{aligned} \quad (127)$$

Note that (126) with $H(q) = -\beta^{-1} f''(q)/f'(q)$ is precisely the equation obtained from (96) after the gauge transformation $q \rightarrow \bar{q} = f(q)$. By computing the second order variation of I_2 around the solution of (126), one can actually show that this $q(x)$ maximizes I_2 .

Finally we derive (104), then (106) by maximizing (104). By assumption, $q(x) = q(\varphi(\alpha))$ provided that $\hat{n} \cdot (x - \varphi) = 0$. It follows that

$$|\nabla q(x)| = g'(\alpha) (\hat{n} \cdot \varphi_\alpha - \hat{n}_\alpha \cdot (x - \varphi))^{-1}, \quad (128)$$

where $g(\alpha) = q(\varphi(\alpha))$. Because of (102) and (103), the term $\hat{n}_\alpha \cdot (x - \varphi)$ is a small correction to $\hat{n} \cdot \varphi_\alpha$ and we will approximate (128) by

$$|\nabla q(x)| = g'(\alpha) (\hat{n} \cdot \varphi_\alpha)^{-1}, \quad (129)$$

This is a localization approximation which is valid provided that the transition tube T defined in (98) is not too wide. Using (129) and noting that $dq =$

$g'(\alpha)d\alpha$ if α instead of q is used as integration variable, (100) can be written as

$$\begin{aligned} I_2 &= \int_0^1 \left(\int_{\mathbb{R}^n} e^{-\beta V} g'(\alpha) (\hat{n} \cdot \varphi_\alpha)^{-1} \delta(\hat{n} \cdot (x - \varphi)) dx \right)^{-1} g'(\alpha) d\alpha, \\ &= \int_0^1 \left(\int_{\mathbb{R}^n} e^{-\beta V} \delta(\hat{n} \cdot (x - \varphi)) dx \right)^{-1} \hat{n} \cdot \varphi_\alpha d\alpha, \end{aligned} \quad (130)$$

(130) is identical to (104), and it can also be written as

$$I_2 = \int_0^1 e^{\beta F(\alpha)} \hat{n} \cdot \varphi_\alpha d\alpha \quad (131)$$

where

$$F(\alpha) = -\beta^{-1} \log \int_{\mathbb{R}^n} e^{-\beta V} \delta(\hat{n} \cdot (x - \varphi)) dx \quad (132)$$

We must maximize (130) with respect to \hat{n} subject to the constraint (102). This constraint is explicitly

$$0 = \int_{\mathbb{R}^n} (x - \varphi) e^{-\beta V(x)} \delta(\hat{n} \cdot (x - \varphi)) dx, \quad (133)$$

and we enforce it by means of adding a Lagrange multiplier term in (130):

$$\begin{aligned} I_2 &= \int_0^1 \left(\int_{\mathbb{R}^n} e^{-\beta V} \delta(\hat{n} \cdot (x - \varphi)) dx \right)^{-1} \hat{n} \cdot \varphi_\alpha d\alpha \\ &\quad + \int_0^1 \int_{\mathbb{R}^n} \lambda \cdot (x - \varphi) e^{-\beta V(x)} \delta(\hat{n} \cdot (x - \varphi)) dx d\alpha. \end{aligned} \quad (134)$$

where $\lambda = \lambda(\alpha)$ is the Lagrange multiplier to be determined later. To obtain the Euler-Lagrange equations associated with (134), we compute the first variation of this functional with respect to \hat{n} and φ . Since $|\hat{n}| = 1$, we can restrict ourselves variations of the type $\hat{n} \rightarrow \hat{n} + \delta\hat{n}^\perp$ in \hat{n} , which gives

$$\begin{aligned}
\delta_{\hat{n}} I_2 &= \int_0^1 e^{\beta F} \delta \hat{n}^\perp \cdot \varphi_\alpha d\alpha \\
&\quad - \int_0^1 e^{2\beta F} \left(\int_{\mathbb{R}^n} e^{-\beta V} \delta \hat{n}^\perp \cdot (x - \varphi) \delta'(\hat{n} \cdot (x - \varphi)) dx \right) \hat{n} \cdot \varphi_\alpha d\alpha \\
&\quad + \int_0^1 \int_{\mathbb{R}^n} \lambda \cdot (x - \varphi) e^{-\beta V(x)} \delta \hat{n}^\perp \cdot (x - \varphi) \delta'(\hat{n} \cdot (x - \varphi)) dx d\alpha \\
&= \int_0^1 e^{\beta F} \delta \hat{n}^\perp \cdot \varphi_\alpha d\alpha \\
&\quad - \int_0^1 e^{2\beta F} \left(\int_{\mathbb{R}^n} e^{-\beta V} \delta \hat{n}^\perp \cdot (x - \varphi) \hat{n} \cdot \nabla \delta(\hat{n} \cdot (x - \varphi)) dx \right) \hat{n} \cdot \varphi_\alpha d\alpha \\
&\quad + \int_0^1 \int_{\mathbb{R}^n} \lambda \cdot (x - \varphi) e^{-\beta V(x)} \delta \hat{n}^\perp \cdot (x - \varphi) \hat{n} \cdot \nabla \delta(\hat{n} \cdot (x - \varphi)) dx d\alpha \\
&= \int_0^1 e^{\beta F} \delta \hat{n}^\perp \cdot \varphi_\alpha d\alpha \\
&\quad - \beta \int_0^1 e^{2\beta F} \left(\int_{\mathbb{R}^n} e^{-\beta V} (\hat{n} \cdot \nabla V) \delta \hat{n}^\perp \cdot (x - \varphi) \delta(\hat{n} \cdot (x - \varphi)) dx \right) \hat{n} \cdot \varphi_\alpha d\alpha \\
&\quad + \beta \int_0^1 \int_{\mathbb{R}^n} \lambda \cdot (x - \varphi) e^{-\beta V(x)} (\hat{n} \cdot \nabla V) \delta \hat{n}^\perp \cdot (x - \varphi) \delta(\hat{n} \cdot (x - \varphi)) dx d\alpha \\
&= \int_0^1 e^{\beta F} \delta \hat{n}^\perp \cdot \varphi_\alpha d\alpha \\
&\quad - \beta \int_0^1 e^{\beta F} \langle (\hat{n} \cdot \nabla V) \delta \hat{n}^\perp \cdot (x - \varphi) \rangle_{\bar{P}(\alpha)} \hat{n} \cdot \varphi_\alpha d\alpha \\
&\quad + \beta \int_0^1 e^{-\beta F} \langle \lambda \cdot (x - \varphi) (\hat{n} \cdot \nabla V) \delta \hat{n}^\perp \cdot (x - \varphi) \rangle_{\bar{P}(\alpha)} d\alpha
\end{aligned} \tag{135}$$

where we used (132) and (133) repeatedly. It follows that the maximizer of (134) satisfies

$$\begin{aligned}
(\varphi_\alpha)^\perp &= \beta \langle (x - \varphi) (\hat{n} \cdot \nabla V) \rangle_{\bar{P}(\alpha)} \hat{n} \cdot \varphi_\alpha \\
&\quad - \beta e^{-2\beta F} \langle (x - \varphi) \lambda \cdot (x - \varphi) (\hat{n} \cdot \nabla V) \rangle_{\bar{P}(\alpha)}.
\end{aligned} \tag{136}$$

Similarly, by computing the variation of (134) with respect to φ and restricting ourselves to variations of the type $\hat{\varphi} \rightarrow \hat{n} + \delta \hat{\varphi}^\perp$ (since we are looking to the new position of $\varphi(\alpha)$ in $\bar{P}(\alpha)$ and therefore $\varphi(\alpha)$ needs to stay in this plane by definition) we obtain

$$\begin{aligned}
\delta_\varphi I_2 &= \int_0^1 e^{\beta F} \hat{n} \cdot (\delta\varphi^\perp)_\alpha d\alpha \\
&\quad - \int_0^1 \int_{\mathbb{R}^n} \lambda(\alpha) \cdot \delta\varphi^\perp e^{-\beta V(x)} \delta(\hat{n} \cdot (x - \varphi)) dx d\alpha \\
&= - \int_0^1 e^{\beta F} \hat{n}_\alpha \cdot \delta\varphi^\perp d\alpha - \int_0^1 \lambda^\perp \cdot \delta\varphi e^{-\beta F} d\alpha.
\end{aligned} \tag{137}$$

Therefore the maximizer of (134) must also satisfy

$$0 = e^{\beta F} \hat{n}_\alpha + \lambda^\perp e^{-\beta F} \tag{138}$$

Solving this equation in λ^\perp and inserting the result in (136), we arrive at

$$\begin{aligned}
(\varphi_\alpha)^\perp &= \beta \langle (x - \varphi)(\hat{n} \cdot \nabla V) \rangle_{\bar{P}(\alpha)} \hat{n} \cdot \varphi_\alpha \\
&\quad + \beta \langle (x - \varphi)\hat{n}_\alpha \cdot (x - \varphi)(\hat{n} \cdot \nabla V) \rangle_{\bar{P}(\alpha)}.
\end{aligned} \tag{139}$$

Combined with (102) this equation specifies \hat{n} and φ . Next we show that (139) is strictly equivalent to (105). First notice that

$$\langle (x - \varphi)(\hat{n} \cdot \nabla V)\hat{n}_\alpha \cdot (x - \varphi) \rangle_{\bar{P}(\alpha)} = \langle (x - \varphi)(\varphi_\alpha \cdot \nabla V) \rangle_{\bar{P}(\alpha)}. \tag{140}$$

Indeed

$$\begin{aligned}
&\beta \langle (x - \varphi)(\hat{n} \cdot \nabla V)\hat{n}_\alpha \cdot (x - \varphi) \rangle_{\bar{P}(\alpha)} \\
&= \beta e^{-\beta F} \int_{\mathbb{R}^n} (x - \varphi)\hat{n}_\alpha \cdot (x - \varphi)(\hat{n} \cdot \nabla V) e^{-\beta V} \delta(\hat{n} \cdot (x - \varphi)) dx \\
&= -e^{-\beta F} \int_{\mathbb{R}^n} \hat{n} \cdot \nabla \left((x - \varphi)\hat{n}_\alpha \cdot (x - \varphi) e^{-\beta V} \right) \delta(\hat{n} \cdot (x - \varphi)) dx \\
&= e^{-\beta F} \int_{\mathbb{R}^n} (x - \varphi)\hat{n}_\alpha \cdot (x - \varphi) e^{-\beta V} \delta'(\hat{n} \cdot (x - \varphi)) dx
\end{aligned} \tag{141}$$

where we used (133) to get the second equality. Since

$$\begin{aligned}
&\hat{n}_\alpha \cdot (x - \varphi) \delta'(\hat{n} \cdot (x - \varphi)) \\
&= \frac{\partial}{\partial \alpha} \delta(\hat{n} \cdot (x - \varphi)) + \hat{n} \cdot \varphi_\alpha \delta'(\hat{n} \cdot (x - \varphi)) \\
&= \frac{\partial}{\partial \alpha} \delta(\hat{n} \cdot (x - \varphi)) + \varphi_\alpha \cdot \nabla \delta(\hat{n} \cdot (x - \varphi)),
\end{aligned} \tag{142}$$

one can continue as

$$\begin{aligned}
& \beta \langle (x - \varphi)(\hat{n} \cdot \nabla V) \hat{n}_\alpha \cdot (x - \varphi) \rangle_{\bar{P}(\alpha)} \\
&= e^{-\beta F} \int_{\mathbb{R}^n} (x - \varphi) e^{-\beta V} \frac{\partial}{\partial \alpha} \delta(\hat{n} \cdot (x - \varphi)) dx \\
&\quad + e^{-\beta F} \int_{\mathbb{R}^n} (x - \varphi) e^{-\beta V} \varphi_\alpha \cdot \nabla \delta(\hat{n} \cdot (x - \varphi)) dx \\
&= e^{-\beta F} \frac{\partial}{\partial \alpha} \int_{\mathbb{R}^n} (x - \varphi) e^{-\beta V} \delta(\hat{n} \cdot (x - \varphi)) dx + \varphi_\alpha \\
&\quad + e^{-\beta F} \int_{\mathbb{R}^n} (x - \varphi) e^{-\beta V} \varphi_\alpha \cdot \nabla \delta(\hat{n} \cdot (x - \varphi)) dx \\
&= \beta e^{-\beta F} \int_{\mathbb{R}^n} (x - \varphi) (\varphi_\alpha \cdot \nabla V) e^{-\beta V} \delta(\hat{n} \cdot (x - \varphi)) dx \\
&= \beta \langle (x - \varphi) (\varphi_\alpha \cdot \nabla V) \rangle_{\bar{P}(\alpha)}
\end{aligned} \tag{143}$$

and arrive at (140). Second note that

$$(\varphi_\alpha)^\perp = \beta \langle ((\varphi_\alpha)^\perp \cdot \nabla V)(x - \varphi) \rangle. \tag{144}$$

To see this we start from

$$J = e^{-\beta F} \int_{\mathbb{R}^n} e^{-\beta V} (x - \varphi) (\hat{n} \cdot \varphi_\alpha) \delta'(\hat{n} \cdot (x - \varphi)) dx \tag{145}$$

and use either

$$(\hat{n} \cdot \varphi_\alpha) \delta'(\hat{n} \cdot (x - \varphi)) = (\hat{n} \cdot \varphi_\alpha) \hat{n} \cdot \nabla \delta'(\hat{n} \cdot (x - \varphi)). \tag{146}$$

or

$$(\hat{n} \cdot \varphi_\alpha) \delta'(\hat{n} \cdot (x - \varphi)) = \varphi_\alpha \cdot \nabla \delta'(\hat{n} \cdot (x - \varphi)). \tag{147}$$

(146) gives

$$\begin{aligned}
J &= e^{\beta F} \int_{\mathbb{R}^n} e^{-\beta V} (x - \varphi) (\hat{n} \cdot \varphi_\alpha) \hat{n} \cdot \nabla \delta(\hat{n} \cdot (x - \varphi)) \\
&= -\hat{n}(\hat{n} \cdot \varphi_\alpha) + \beta \langle (x - \varphi) (\hat{n} \cdot \nabla V) \rangle_{\bar{P}(\alpha)} (\hat{n} \cdot \varphi_\alpha)
\end{aligned} \tag{148}$$

If we use (147) instead, we get

$$\begin{aligned}
J &= e^{\beta F} \int_{\mathbb{R}^n} e^{-\beta V} (x - \varphi) (\hat{n} \cdot \varphi_\alpha) \hat{n} \cdot \nabla \delta(\hat{n} \cdot (x - \varphi)) \\
&= \varphi_\alpha + \beta \langle (x - \varphi) (\varphi_\alpha \cdot \nabla V) \rangle_{\bar{P}(\alpha)}
\end{aligned} \tag{149}$$

(148) and (149) gives (144). Inserting (140) and (144) in (139) we deduce that this equation is indeed equivalent to (105).

Finally to derive (106), we recall that (130) is equivalent to the original object function in (100) only within the localization assumption in (103) which allows us to approximate (128) by (129). But (103) implies that

$$\frac{\partial}{\partial \alpha} \delta(\hat{n} \cdot (x - \varphi)) \approx -\hat{n} \cdot \varphi_\alpha \delta'(\hat{n} \cdot (x - \varphi)), \quad (150)$$

since the term $\hat{n}_\alpha \cdot (x - \varphi) \delta'(\hat{n} \cdot (x - \varphi))$ in (142) is a small correction. Using (150), we have

$$\begin{aligned} 0 &= \beta \langle (\hat{n} \cdot \nabla V)(x - \varphi) \rangle_{\bar{P}(\alpha)} \\ &= \beta e^{-\beta F} \int_{\mathbb{R}^n} (\hat{n} \cdot \nabla V) e^{-\beta V}(x - \varphi) \delta(\hat{n} \cdot (x - \varphi)) dx \\ &= \hat{n} - e^{-\beta F} \int_{\mathbb{R}^n} \hat{n} \cdot \nabla \left(e^{-\beta V}(x - \varphi) \right) \delta(\hat{n} \cdot (x - \varphi)) dx \\ &= \hat{n} + e^{-\beta F} \int_{\mathbb{R}^n} e^{-\beta V}(x - \varphi) \delta'(\hat{n} \cdot (x - \varphi)) dx \\ &\approx \hat{n} - e^{-\beta F} (\hat{n} \cdot \varphi_\alpha)^{-1} \int_{\mathbb{R}^n} e^{-\beta V}(x - \varphi) \frac{\partial}{\partial \alpha} \delta(\hat{n} \cdot (x - \varphi)) dx \\ &= \hat{n} - e^{-\beta F} (\hat{n} \cdot \varphi_\alpha)^{-1} \frac{\partial}{\partial \alpha} \int_{\mathbb{R}^n} e^{-\beta V}(x - \varphi) \delta(\hat{n} \cdot (x - \varphi)) dx - (\hat{n} \cdot \varphi_\alpha)^{-1} \varphi_\alpha \\ &= \hat{n} - (\hat{n} \cdot \varphi_\alpha)^{-1} \varphi_\alpha. \end{aligned} \quad (151)$$

This is (106).

References

- [Ben77] C. H. Bennett, “Exact defect calculations in model substances” in A. S. Nowick and J. J. Burton, editors, *Algorithms for Chemical Computation*, ACS Symposium Series No. 46, **63** (1977).
- [BEGK03a] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, “Metastability in reversible diffusion processes 1. Sharp estimates for capacities and exit times,” WIAS-preprint 767 (2003).
- [BEGK03b] A. Bovier, V. Gayrard, and M. Klein, “Metastability in reversible diffusion processes. 2. Precise estimates for small eigenvalues,” WIAS-preprint 768 (2003).
- [BCDG02] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, “Transition Path Sampling: Throwing Ropes over Mountain Passes in the Dark,” *Annual Review in Physical Chemistry* **53**, 291 (2002).
- [BDC98] P. G. Bolhuis, C. Dellago and D. Chandler, “Sampling ensembles of deterministic transition pathways,” *Faraday Discussion* **110**, 421 (1998).
- [BDC00] P. G. Bolhuis, C. Dellago, and D. Chandler, “Reaction Coordinates of Biomolecular Isomerization,” *Proc. of Nat. Acad. Sci. USA* **97**, 5877 (2000).
- [CCHK89] E. A. Carter, G. Ciccotti, J. T. Hynes, and R. Kapral’ “Constrained reaction coordinate dynamics for the simulation of rare events,” *Chem. Phys. Lett.* **156**, 472 (1989).
- [Chan78] D. Chandler, *J. Chem. Phys.* **68**, 2959 (1978).
- [DBG02] C. Dellago, P. G. Bolhuis, P. L. Geissler, “Transition Path Sampling,” *Advances in Chemical Physics* **123**, 1 (2002).

- [ERVE02a] W. E, W. Ren, and E. Vanden-Eijnden, “String method for the study of rare events,” *Phys. Rev. B* **66**, 052301 (2002).
- [ERVE02b] W. E, W. Ren, and E. Vanden-Eijnden, “Energy Landscapes and Rare Events,” ICM report (2002).
- [EGC02] R. Elber, A. Ghosh, and A. Cárdenas, “Long time dynamics of complex systems”, *Account of Chemical Research* **35**,396–403 (2002).
- [ECGS03] R. Elber, A. Ghosh, A. Cárdenas, and H. Stern, “Bridging the gap between reaction pathways, long time dynamics and calculation of rates”, *Advances in Chemical Physics* **126**, 93–129 (2003).
- [Eyr35] H. Eyring, *J. Chem. Phys.* , **3** 107 (1935).
- [FSDC02] A. Fischer, Ch. Schütte, P. Deuffhard, and F. Cordes, “Hierarchical Uncoupling-Coupling of Metastable Conformations Computational Methods for Macromolecules” in: *Challenges and Applications*, Proceedings of the 3rd International Workshop on Algorithms for Macromolecular Modeling. Edited by T. Schlick and H. H. Gan, LNCSE, Springer, 24 (2002)
- [FW98] M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, 2nd ed. (Springer, 1998).
- [FS01] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithm to Applications*, 2nd. ed. (Elsevier, 2001).
- [Gar89] G. W. Gardiner, *Handbook of stochastic methods*, 2nd ed. (Springer, 1997).
- [HJ00] G. Henkelman and H. Jónsson, “Improved Tangent Estimate in the NEB Method for Finding Minimum Energy Paths and Saddle Points,” *J. Chem. Phys.* **113**, 9978 (2000).
- [HTV94] T. Huber, A. E. Torda, and W.F. van Gunsteren, “Local elevation: A method for improving the searching properties of molecular dynamics simulation,” *J. Comput.-Aided Mol. Design* **8**, 695 (1994).
- [HMS02] W. Huisinga, S. Meyn, and Ch. Schütte, “Phase Transitions and Metastability in Markovian and Molecular Systems,” submitted (2002).
- [ILP03] M. Iannuzzi, A. Laio, and M. Parrinello, “Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics,” *Phys. Rev. Lett.* **90**, 238302 (2003)
- [JJ01] G. H. Jóhannesson and H. Jónsson, “Optimization of Hyperplanar Transition States,” *J. Chem. Phys.* **115**, 9644 (2001).
- [JMJ98] H. Jónsson, G. Mills, K. W. Jacobsen, “Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions,” p. 385 in *Classical and Quantum Dynamics in Condensed Phase Simulations*, ed. B. J. Berne, G. Ciccotti, and D. F. Coker (World Scientific, 1998)
- [LP02] A. Laio and M. Parrinello, “Escaping free-energy minima,” *Proc. Nat. Acad. Sc. USA* **99**, 12562 (2002).
- [Pap76] G. Papanicolaou, “Some probabilistic problems and methods in singular perturbations,” *Rocky Mountain Math. J.* **6**, 653 (1976).
- [Pra86] L. R. Pratt, “A statistical method for identifying transition states in high dimensional problems,” *J. Chem. Phys.* **9**, 5045 (1986).
- [Ren02] W. Ren, *Numerical Methods for the Study of Energy Landscapes and Rare Events*, PhD thesis, New York University, 2002.
- [Ren03] W. Ren, “Higher order string method for finding minimum energy paths,” *Comm. Math. Sci.* **1**, 377 (2003).
- [RCB77] J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, “Numerical integration of the Cartesian equation of motion of a system with constraints: molecular dynamics of N-alkanes,” *J. Comp. Phys.* **23**, 327 (1977).

- [SH02] Ch. Schütte and W. Huisinga, “Biomolecular Conformations can be Identified as Metastable Sets of Molecular Dynamics,” submitted (2002).
- [SHD02] Ch. Schütte, W. Huisinga, and P. Deuffhard, “Transfer Operator Approach to Conformational Dynamics in Biomolecular Systems,” pp.191–223 in *Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems* edited by B. Fiedler (2001)
- [SV00] M. R. Sorensen and A. Voter, “Temperature-accelerated dynamics for simulation of infrequent events,” *J. Chem. Phys.* **112**, 9599 (2000).
- [SC98] M. Sprik and G. Ciccotti, “Free energy from constrained molecular dynamics,” *J. Chem Phys.* **109**, 7737 (1998).
- [Str01] J. E. Straub, “Reaction rates and transition pathways”, p. 199 in *Computational biochemistry and biophysics*, ed. O. M Becker, A. D MacKerell, Jr., B. Roux, and M. Watanabe (Marcel Decker, Inc. 2001).
- [UE89] A. Ulitsky and R. Elber, “A new technique to calculate steepest descent paths in flexible polyatomic systems,” *J. Chem. Phys.* **92**, 1510 (1990).
- [Vot97] A. Voter, “Hyperdynamics: Accelerated molecular dynamics of infrequent events,” *Phys. Rev. Lett.* **78**, 3908 (1997).
- [WL01] F. Wang and P. D. Landau, “Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States,” *Phys. Rev. Lett.* **86**, 2050 (2001).
- [WCH01] T. O. White, G. Ciccotti, and J.-P. Hansen, “Brownian dynamics with constraints.” *Mol. Phys.* **99**, 2023, (2001)
- [Wig38] E. Wigner, *Trans. Faraday Soc.* **34**, 29 (1938).
- [Yam60] T. Yamamoto, *J. Chem Phys.* , **33**, 281 (1960).