Published in final edited form as:
J Phys Chem B. 2013 October 24; 117(42): 13115–13119. doi:10.1021/jp403043a.

Diffusion along the Splitting/Commitment Probability Reaction Coordinate

Alexander M. Berezhkovskii[*] and
Mathematical and Statistical Computing Laboratory, Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892, United States

Attila Szabo
Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

Abstract

The splitting or commitment probabilities of states in the region of configuration space that separates reactants and products play an important role in the theory of chemical reactions. Assuming that the splitting probability changes more slowly than any other coordinate, we project multidimensional diffusive dynamics onto it. The resulting one-dimensional diffusion equation is not exact because the assumed separation of time scales does not hold in general. Nevertheless, this equation has the remarkable property that it always predicts the exact value of the number of transitions between reactants and products per unit time at equilibrium and hence the exact reaction rate. In the special case of two deep basins separated by a harmonic saddle, this equation is equivalent to the one that describes diffusion along a coordinate perpendicular to the transition state, defined as the surface starting from which reactants and products are reached with equal probability.

$$\nabla \cdot \mathbf{D}(\mathbf{x}) e^{-\beta U(\mathbf{x})} \nabla \phi(\mathbf{x}) = 0,$$
$$J = \int_I \nabla \phi(\mathbf{x}) \cdot \mathbf{D}(\mathbf{x}) p_{eq}(\mathbf{x}) \nabla \phi(\mathbf{x}) d\mathbf{x},$$
$$\frac{\partial p(\phi,t)}{\partial t} = J \frac{\partial^2}{\partial \phi^2} \frac{p(\phi,t)}{p_{eq}(\phi)}.$$

1. INTRODUCTION

In 1989, Bryngelson and Wolynes[1] proposed that the dynamics of protein folding can be described as diffusion along a collective reaction coordinate in the presence of the corresponding potential of mean force. Socci et al.[2] then showed, without using any adjustable parameters, that the rate of folding of a lattice model of a protein can be quantitatively reproduced by the diffusion model. Specifically, they found that the rate obtained from diffusion in the potential of mean force along a coordinate defined as the

---

[*] Corresponding Author: alexanderb@mail.nih.gov; telephone number: 301-451-7244.

Notes

The authors declare no competing financial interest.

fraction of native contacts, using a diffusion constant extracted from the relaxation time of the unfolded state, agreed with the simulated rate over a wide range of temperatures. More recently, Best and Hummer[3] showed that the folding rate of the off-lattice coarse-grained structure-based models can also be accurately described using the one-dimensional diffusion model. Finally, this model has provided a remarkably useful framework for the analysis and interpretation of experimental data.[4,5]

In general, multidimensional dynamics can be reduced to the dynamics of a single coordinate only when the dynamics along all other coordinates is sufficiently fast to establish local equilibrium. Motivated in part by successes mentioned above, some time ago we asked[6] if a one-dimensional reaction coordinate can exist under less restrictive conditions (i.e., in the absence of time-scale separation). For multidimensional diffusive dynamics, we showed that, when the reactants and products are separated by a single high harmonic saddle, the direction perpendicular to the stochastic separatrix is always a good reaction coordinate. Specifically, the rate obtained by using Kramers'theory[7] for one-dimensional diffusion in the potential of mean force along this coordinate was shown to be the same as that given by multidimensional Langer theory.[8]

The stochastic separatrix is the surface with the splitting or commitment probability equal to one-half. The splitting probability of a state is the probability that, starting from this state, the system reaches the product region of the configuration state before the reactant region. In the context of protein folding, it was introduced by Du et al.[9] under the name p_{fold}: the probability of folding before unfolding. For a variety of reasons, the splitting/commitment probability has been considered as the optimal reaction coordinate (see refs 10–13, and references therein). In this paper, we shall prove that, if multidimensional diffusive dynamics is projected onto this coordinate, assuming that all other coordinates are in local equilibrium, the resulting diffusion equation always leads to the exact reactive flux at equilibrium, even when it does not correctly describe the dynamics. This is similar to the statement recently made by Krivov[14] that "the equilibrium flux between two boundary states can be computed exactly as diffusion on a free energy profile associated with the" splitting probability coordinate. Vanden-Eijnden has pointed out to us that the present work is related to milestoning using isocommitor surfaces where the exact mean first-passage times and transition rates can be obtained.[15] In the special case of a high harmonic saddle, we will show that diffusion along the splitting probability coordinate is equivalent to diffusion along the direction perpendicular to the stochastic separatrix. In this sense, the present paper generalizes our previous work[6] which showed that this direction provides a one-dimensional reaction coordinate that leads to the exact reaction rate when the products and reactants are separated by a high multidimensional harmonic saddle.

2. UNIDIRECTIONAL REACTIVE FLUX

In this section we obtain some results of transition path theory[12,13] by adapting the point of view adopted by Kramers in his seminal work.[7] Consider a system whose dynamics is described as diffusion in an N-dimensional potential of mean force $U(\mathbf{x})$. The probability density $p(\mathbf{x},t)$ of finding the system at point $\mathbf{x}$ at time t satisfies the Smoluchowski equation

$$\frac{\partial p(\mathbf{x},t)}{\partial t}=\nabla\cdot\mathbf{D}(\mathbf{x})e^{-\beta U(\mathbf{x})}\nabla e^{-\beta U(\mathbf{x})}p(\mathbf{x},t) \qquad (2.1)$$

where $\beta = (k_B T)^{-1}$ and $\mathbf{D}(\mathbf{x})$ is the position-dependent matrix of diffusion coefficients. Let A and B be domains in configuration space separated by an intermediate region I, so that any trajectory from A to B (and vice versa) must cross the I-region. We are interested in the flux or the number of transitions per unit time between A and B when the system is at equilibrium. This is the same as the flux between B and A, and therefore we shall denote both fluxes by J. In the special case that A and B are deep basins and I is the barrier region, J is of particular interest because it describes the transitions between the two wells. In this case J is equal to the product of the rate constant from A to B (B to A) and the normalized equilibrium population of A (B).

Following Kramers, we establish a steady state in the I-region by maintaining the distribution on the surface separating A and I (denoted by ∂A) at equilibrium while forcing the distribution to vanish on the surface separating I and B (denoted by ∂B). Then J is equal to the flux through any dividing surface in the I-region. The resulting steady-state distribution has a simple physical interpretation: it is the probability density obtained from only those fragments of the equilibrium trajectory that enter the I-region from the A-domain. This probability density $p_{ss}(\mathbf{x})$ satisfies

$$\nabla\cdot\mathbf{D}(\mathbf{x})e^{\beta U(\mathbf{x})}\nabla e^{-\beta U(\mathbf{x})}p_{\mathrm{ss}}(\mathbf{x})=0 \qquad (2.2)$$

for $\mathbf{x} \in I$, with boundary conditions $p_{ss}(\mathbf{x} \in \partial A) = p_{eq}(\mathbf{x})$ and $p_{ss}(\mathbf{x} \in \partial B) = 0$, where $p_{eq}(\mathbf{x})$ is the equilibrium distribution, $p_{eq}(\mathbf{x}) = \exp(-\beta U(\mathbf{x}))/Z$, $Z = \int \exp(-\beta U(\mathbf{x}))d\mathbf{x}$, with the integral being over the regions A, I, and B. The flux J is given by the surface integral

$$J=-\int_{\sum}\hat{\mathbf{n}}(\mathbf{x})\mathbf{D}\cdot(\mathbf{x})e^{-\beta U(\mathbf{x})}\nabla e^{\beta U(\mathbf{x})}p_{\mathrm{ss}}(\mathbf{x})\,\mathrm{d}S \qquad (2.3)$$

where Σ is an arbitrary dividing surface in the I-region and $\mathbf{n}(\hat{\mathbf{x}})$ is the unit vector normal to this surface at point $\mathbf{x}$ directed toward the B-domain.

The steady-state distribution, $p_{ss}(\mathbf{x})$, and hence the flux, J, can be expressed in terms of the splitting/commitment probability of the states in the I-region. To see this, let us write $p_{ss}(\mathbf{x})$ as

$$p_{\mathrm{ss}}(\mathbf{x})=(1-\phi(\mathbf{x}))p_{\mathrm{eq}}(\mathbf{x}) \qquad (2.4)$$

where, at the moment, $\varphi(\mathbf{x})$ is an arbitrary function, and substitute it into eq 2.2. In this way we find that $\varphi(\mathbf{x})$ satisfies

$$\nabla\cdot\mathbf{D}(\mathbf{x})e^{-\beta U(x)}\nabla\phi(\mathbf{x})=0 \quad (2.5)$$

for $\mathbf{x}\in I$, with boundary conditions $\varphi(\mathbf{x}\in\partial A)=0$ and $\varphi(\mathbf{x}\in\partial B)=1$. Since eq 2.5 is Onsager's equation[16] for the splitting probability, it follows from the boundary conditions that $\varphi(\mathbf{x})$ is the probability of reaching the B-domain before the A-domain starting from $\mathbf{x}$, and hence it is the splitting probability.

If the dividing surface Σ is chosen to be an isocommittor surface $\varphi(\mathbf{x})=\varphi$, $0<\varphi<1$, then the integral in eq 2.3 can be rewritten as

$$J=\int_I\delta(\phi(\mathbf{x})-\phi)\nabla\phi(\mathbf{x})\cdot\mathbf{D}(\mathbf{x})p_{\mathrm{eq}}(\mathbf{x})\nabla\phi(\mathbf{x})\,\mathrm{d}\mathbf{x} \quad (2.6)$$

where we have used the fact that $p_{\mathrm{ss}}(\mathbf{x})$ is given by eq 2.4 and that $p_{\mathrm{eq}}(\mathbf{x})=\exp(-\beta U(\mathbf{x}))/Z$. Since there are neither sinks nor sources in the I-region, the flux J is a constant independent of φ. To show this formally consider the following sequence of identities

$$\begin{aligned}\frac{\partial J}{\partial\phi}&=\int_I\frac{\partial}{\partial\phi}\partial(\phi(\mathbf{x})-\phi)\nabla\phi(\mathbf{x})\cdot\mathbf{D}(\mathbf{x})p_{\mathrm{eq}}(\mathbf{x})\nabla\phi(\mathbf{x})\,\mathrm{d}\mathbf{x}\\&=-\int_I\nabla\delta(\phi(\mathbf{x})-\phi)\cdot\mathbf{D}(\mathbf{x})p_{\mathrm{eq}}(\mathbf{x})\nabla\phi(\mathbf{x})\,\mathrm{d}\mathbf{x}\\&=\int_I\delta(\phi(\mathbf{x})-\phi)\nabla\cdot\mathbf{D}(\mathbf{x})p_{\mathrm{eq}}(\mathbf{x})\nabla\phi(\mathbf{x})\,\mathrm{d}\mathbf{x}=0\end{aligned} \quad (2.7)$$

where in succession we used the identity $\nabla\delta(\varphi(\mathbf{x})-\varphi)=-\partial\delta(\varphi(\mathbf{x})-\varphi)/\partial\varphi\nabla\varphi(\mathbf{x})$, integration by parts, and the Onsager equation for $\varphi(\mathbf{x})$, eq 2.5. Since J is a constant, we can integrate both sides of eq 2.6 over φ from zero to unity. In this way we find

$$J=\int_I\nabla\phi(\mathbf{x})\cdot\mathbf{D}(\mathbf{x})p_{\mathrm{eq}}(\mathbf{x})\nabla\phi(\mathbf{x})\,\mathrm{d}\mathbf{x} \quad (2.8)$$

This result was first obtained by E and Vanden-Eijnden.[12] Expressions for the flux J in eq 2.6 will play a key role in the next section.

3. DIFFUSION EQUATION ALONG THE SPLITTING PROBABILITY COORDINATE

Let $p(\varphi,t)$ be the probability density of finding the system at time t on the isocommittor surface $\varphi(\mathbf{x})=\varphi$ in the I-region, $0<\varphi<1$. It can be expressed in terms of the probability density $p(\mathbf{x},t)$ as

$$p(\phi,t)=\int_I \delta(\phi(\mathbf{x})-\phi)p(\mathbf{x},t)\,\mathrm{d}\mathbf{x} \quad (3.1)$$

Multiplying both sides of eq 2.1 by $\delta(\varphi(\mathbf{x}) - \varphi)$, integrating the resulting equation over all $\mathbf{x}$ in the I-region, integrating by parts twice, and taking advantage of the identity $\nabla\delta(\varphi(\mathbf{x}) - \varphi) = -\partial\delta(\varphi(\mathbf{x}) - \varphi)/\partial\varphi\nabla\varphi(\mathbf{x})$ and the Onsager equation, eq 2.5, we find that

$$\frac{\partial p(\phi,t)}{\partial t}=\frac{\partial^2}{\partial\phi^2}\int_I \delta(\phi(\mathbf{x})-\phi)\,p(\mathbf{x},t)\nabla\phi(\mathbf{x})\cdot\mathbf{D}(\mathbf{x})\nabla\phi(\mathbf{x})\mathrm{d}\mathbf{x} \quad (3.2)$$

In general, this equation cannot be closed, because the dynamics along φ is non-Markovian, and hence no simple diffusion equation for $p(\varphi,t)$ exists.

However, this equation can be closed, if we assume that φ is a slow variable, in the sense that local equilibrium is rapidly established on surfaces with fixed φ, $p(\mathbf{x},t)|_{\varphi(\mathbf{x})=\varphi} \propto e^{-\beta U(\mathbf{x})}p(\varphi,t)$. In this case, the time evolution of the probability density of every state with the same φ is the same, and we can write $p(\mathbf{x},t)$ as

$$p(\mathbf{x},t)=\frac{p_{\mathrm{eq}}(\mathbf{x})}{\int_I \delta(\phi(\mathbf{y})-\phi(\mathbf{x}))p_{\mathrm{eq}}(\mathbf{y})\,\mathrm{d}\mathbf{y}}\times p(\phi(\mathbf{x}),t) \quad (3.3)$$

$\mathbf{x} \in I$, where the denominator ensures that eq 3.1 holds. A similar approximation has been used by Peters et al.[17] as the basis of their dynamical self-consistency test for identifying the optimal reaction coordinate. Substituting $p(\mathbf{x}, t)$ in eq 3.3 into eq 3.2 we obtain

$$\frac{\partial p(\phi,t)}{\partial t}=\frac{\partial^2}{\partial\phi^2}\left[\int_I \delta(\phi(\mathbf{x})-\phi)\nabla\phi(\mathbf{x})\cdot\mathbf{D}(\mathbf{x})p_{\mathrm{eq}}(\mathbf{x})\nabla\phi(\mathbf{x})\,\mathrm{d}\mathbf{x}\right]\frac{p(\phi,t)}{\int_I \delta(\phi(\mathbf{y})-\phi)p_{\mathrm{eq}}(\mathbf{y})\,\mathrm{d}\mathbf{y}} \quad (3.4)$$

Remarkably, the integral in the square brackets is the exact flux J given in eq 2.6. As J is independent of φ, it can be pulled outside the derivatives. Since the equilibrium distribution of φ is

$$p_{\mathrm{eq}}(\phi)=\int_I \delta(\phi(\mathbf{x})-\phi)p_{\mathrm{eq}}(\mathbf{x})\,\mathrm{d}\mathbf{x} \quad (3.5)$$

Equation 3.4 can be written as

$$\frac{\partial p(\phi,t)}{\partial t}=J\frac{\partial^2}{\partial\phi^2}\frac{p(\phi,t)}{p_{\mathrm{eq}}(\phi)} \qquad (3.6)$$

This is a key result of this paper. A remarkable feature of eq 3.6 is that it contains the exact flux J. However, in general, it does not provide an exact description of the projected dynamics since the local equilibrium ansatz in eq 3.3 is an approximation. Nevertheless, if one calculates the number of unidirectional transitions between the boundaries of the I-region per unit time at equilibrium from this approximate one-dimensional evolution equation, it will be exact, irrespective of the complexity of the underlying multidimensional free energy surface.

To see this, we note that eq 3.6 can be rewritten in the standard form of the one-dimensional Smoluchowski equation (i.e., eq 2.1 in one dimension), if one defines a position-dependent diffusion coefficient $D(\varphi)$ as

$$D(\phi)p_{\mathrm{eq}}(\phi)=J \qquad (3.7)$$

Then eq 2.8 for the flux, specialized to one-dimensional diffusion along the φ-coordinate, takes the form $J=\int_0^1(\mathrm{d}\phi/\mathrm{d}\phi)D(\phi)\,p_{\mathrm{eq}}(\phi)\,(\mathrm{d}\phi/\mathrm{d}\phi)\mathrm{d}\phi=\int_0^1 J\,\mathrm{d}\phi=J$, where we have used eq 3.7. It should not be surprising that the projection of the multidimensional diffusive dynamics onto the φ-coordinate yields a one-dimensional diffusion equation that leads to the correct flux J, because to do the projection, one must know $\varphi(\mathbf{x})$ and hence, from eq 2.8, J. Using eq 3.7, we find that the one-dimensional version of the Onsager equation, eq 2.5, has the form $Jd^2\varphi/\mathrm{d}\varphi^2 = 0$, which yields the exact splitting probability. Finally, we mention that using the transformation[18] $z \propto \int \mathrm{d}\varphi/D(\varphi)^{1/2} = \int (p_{\mathrm{eq}}(\varphi)/J)^{1/2}\mathrm{d}\varphi$, one can convert eq 3.6 into a Smoluchowski equation with a constant diffusion coefficient.

When states of the system are discrete and the dynamics is described by a master equation, the analogue of the above results is derived in the Appendix.

To make contact with our previous work,[6] we shall show that, when the intermediate region is a harmonic saddle and the matrix of diffusion coefficients is position-independent, $\mathbf{D}(\mathbf{x}) = \mathbf{D}$, eq 3.6 is equivalent to the diffusion equation along a coordinate q which is perpendicular to the stochastic separatrix. To do this we need to use some results from our previous paper. Placing the origin at the saddle point, we can write $\beta U(\mathbf{x}) = \mathbf{x}\cdot\mathbf{K}\mathbf{x}/2$, $\mathbf{x} \in I$, where $\mathbf{K}$ is the matrix of the force constants at the saddle point measured in units of $k_B T$, det $\mathbf{K} < 0$. Let $\mathbf{e}$ be a unit vector perpendicular to the stochastic separatrix which is the hyperplane formed by all points with $\varphi(\mathbf{x}) = 0.5$. In the special case of a harmonic saddle, the splitting probability is

$$\phi(\mathbf{x})=\sqrt{\kappa/2\pi}\int_{-\infty}^{\mathbf{e}\cdot\mathbf{x}}\mathrm{e}^{-\kappa z^2/2}\,\mathrm{d}z \qquad (3.8)$$

where $\kappa = |\mathbf{e}\cdot\mathbf{K}^{-1}\mathbf{e}|^{-1}$. All points $\mathbf{x}$ for which $\mathbf{e}\cdot\mathbf{x}$ is the same, form isocommittor surfaces which are parallel hyperplanes perpendicular to $\mathbf{e}$.

The flux J can be obtained by multiplying the Langer rate constant by the equilibrium population of the well. This leads to

$$J=\sqrt{\frac{(2\pi)^{N-2}}{|\det\mathbf{K}|}}\frac{\mathbf{e}\cdot\mathbf{D}\mathbf{e}}{Z|\mathbf{e}\cdot\mathbf{K}^{-1}\mathbf{e}|} \qquad (3.9)$$

The equilibrium distribution along the q-coordinate, which is perpendicular to the stochastic separatrix, is

$$\begin{aligned}p_{\mathrm{eq}}(q)&=\frac{1}{Z}\int\delta(q-\mathbf{e}\cdot\mathbf{x})\mathrm{e}^{-\mathbf{x}\cdot\mathbf{K}\mathbf{x}/2}\,\mathrm{d}\mathbf{x}\\&=\frac{1}{Z}\sqrt{\frac{(2\pi)^{N-1}\kappa}{|\det\mathbf{K}|}}\mathrm{e}^{\kappa q^2/2}\end{aligned} \qquad (3.10)$$

Using eqs 3.8–3.10 and the relationship $p(\varphi,t)d\varphi = p(q,t)\mathrm{d}q$, we can transform eq 3.6 into

$$\frac{\partial p(q,t)}{\partial t}=\mathbf{e}\cdot\mathbf{D}\mathbf{e}\frac{\partial}{\partial q}p_{\mathrm{eq}}(q)\frac{\partial}{\partial q}\frac{p(q,t)}{p_{\mathrm{eq}}(q)} \qquad (3.11)$$

which is the Smoluchowski equation along the q-coordinate. In this special case, because vector $\mathbf{e}$ points along the unstable diffusive mode of the saddle point dynamics, eq 3.11 is actually exact within the framework of the multidimensional model.

To obtain a one-dimensional description of the entire system, one must treat the dynamics not only in the I-region, but also in the A and B domains. The simplest way of doing this is to impose the appropriate boundary conditions on eq 3.6, which describes the dynamics only in the intermediate region. Specifically, we will consider the end-points of the I-region, $\varphi = 0$ and 1, as reversible binding sites representing the A and B domains. Let $P_A(t) = \int_A p(\mathbf{x},t)\mathrm{d}\mathbf{x}$ ($P_B(t) = \int_B p(\mathbf{x},t)\mathrm{d}\mathbf{x}$) be the probability of finding the system at time t in the A (B) domain. Because of probability conservation, we have

$$p_A(t)+\int_0^1 p(\phi,t)\,\mathrm{d}\phi+p_B(t)=1 \qquad (3.12)$$

Since the flux must be continuous at the A–I boundary

$$\frac{\mathrm{d}P_A(t)}{\mathrm{d}t}=J\frac{\partial}{\partial\phi}\frac{p(\phi,t)}{p_{\mathrm{eq}}(\phi)}\Big|_{\phi=0} \quad (3.13)$$

An analogous relationship holds at the I–B boundary. If we describe the exit from the binding site corresponding to the A-domain to the I-region by the rate constant k_A (k_A^{-1} is the mean lifetime of state A) and the entrance to this site by the so-called radiation boundary condition with the intrinsic rate κ_A, then the flux across the A–I boundary is

$$J\frac{\partial}{\partial\phi}\frac{p(\phi,t)}{p_{\mathrm{eq}}(\phi)}\Big|_{\phi=0}=\kappa_A p(0,t)-k_A P_A(t) \quad (3.14)$$

Because of detailed balance, $\kappa_A p_{\mathrm{eq}}(0)=k_A P_A^{\mathrm{eq}}$, where P_A^{eq} is the normalized equilibrium population of the A-domain. Thus we can write this flux as

$$J\frac{\partial}{\partial\phi}\frac{p(\phi,t)}{p_{\mathrm{eq}}(\phi)}\Big|_{\phi=0}=k_A P_A^{\mathrm{eq}}\left(\frac{p(0,t)}{p_{\mathrm{eq}}(0)}-\frac{P_A(t)}{P_A^{\mathrm{eq}}}\right) \quad (3.15)$$

The analogous equation holds at the I–B boundary ($\varphi = 1$).

Equation 3.6 supplemented by the above boundary conditions at the end-points completely determine the dynamics of the entire system given J, $p_{\mathrm{eq}}(\varphi)$, $P_{A,B}^{\mathrm{eq}}$, and $k_{A,B}$. It can be shown that, if one calculates the average number of transitions between the A and B domains per unit time, in the framework of this model, it will be correct (i.e., equal to J) irrespective of the values of $k_{A,B}$. However, appearance of the trajectories predicted by this simple model will be sensitive to the values of these parameters, because they determine the residence or dwell times of the system in the A and B states. A sensible choice for $k_{A,B}$ is to equate them to the reciprocal of the mean first-passage times from the A- and B-domains to the I-region, averaged over the equilibrium distribution of the starting positions. Clearly, the accuracy of this simple model depends on the nature of the underlying multidimensional free energy surface. It will be only a crude approximation for the dynamics when the free energy surface $U(\mathbf{x})$ is essentially flat, even though it correctly predicts the flux. However, it works quite well when the A and B domains are deep basins.

4. CONCLUDING REMARKS

If one assumes that the splitting/commitment probability is the slowest coordinate and projects the multidimensional Smoluchowski equation onto it, the resulting one-dimensional diffusion equation always leads to the exact result for the flux (number of transitions per unit time) between two domains at equilibrium. This is true even when the splitting probability does not change sufficiently slowly for all other coordinates to be effectively equilibrated, so that the projected dynamics is non-Markovian and hence cannot be described by a one-dimensional Smoluchowski equation. An analogous result holds when the configuration space is discrete and the dynamics is described by a master equation (see Appendix). In this case one can reduce the dimensionality of the rate matrix by grouping states with the same splitting probabilities, assuming that the occupancies of these states have the same dependence on time. It can be shown that the value of the equilibrium reactive flux, as given by the expression derived independently by Metzner et. al.[19] and Berezhkovskii et al.,[20] remains unchanged under such transformation.

For the formalism developed above to be useful in practice, one must be able to find a good approximation for the splitting/commitment probability of states in the region separating reactants from products. This is an active area of research (see refs 13–15, 21, 22, and references therein), which is beyond the scope of this paper. Suffice it to say that variational principles play an important role in this regard. Finally, we like to take this opportunity to point out that the variational principle we proposed[6] for a double-well system with a single harmonic saddle point is a special case of the variational principle of E and Vanden-Eijnden,[11–13] which in turn is an extension of Dirichlet's principle for Laplace's equation. When the reactants and products are separated by a single harmonic saddle, we showed that, by minimizing the flux given in eq 3.9 with respect to the direction of the unit vector **e**, one recovers the corresponding multidimensional Langer result. E and Vanden-Eijnden showed that the flux is the minimum of the functional $\int_I \nabla f(\mathbf{x})\cdot\mathbf{D}p_{\mathrm{eq}}(\mathbf{x})\nabla f(\mathbf{x})\,\mathrm{d}\mathbf{x}$ with respect to functions $f(\mathbf{x})$, that satisfy the same boundary conditions as the splitting probability. In the case of a harmonic saddle where $p_{\mathrm{eq}}(\mathbf{x}) = \exp(-\mathbf{x}\cdot\mathbf{K}\mathbf{x}/2)/Z$, if one chooses a trial function of the form $f(\mathbf{x})=(\kappa/2\pi)^{1/2}\int_{-\infty}^{\mathbf{e}\cdot\mathbf{x}}\exp(-\kappa z^2/2)\,\mathrm{d}z$ with $\kappa = |\mathbf{e}\cdot\mathbf{K}^{-1}\mathbf{e}|^{-1}$, then the above functional turns out to be identical to the flux in eq 3.9. Thus, in this case, minimization of the functional with respect to the direction of vector **e** is equivalent to minimization of the flux obtained from the potential of mean force along the direction **e**, using the one-dimensional Kramers theory.[6]

Acknowledgments

We thank Sergei Krivov for many discussions and Margaret Johnson, Alexander Grosberg, and Eric Vanden-Eijnden for their comments on our manuscript. In addition, we thank Eric Vanden-Eijnden for sending us a draft of an unpublished manuscript which arrives at similar conclusions. This study was supported by the Intramural Research Program of the NIH, Center for Information Technology and National Institute of Diabetes and Digestive and Kidney Diseases.

References

1. Bryngelson JD, Wolynes PG. Intermediates and Barrier Crossing in a Random Energy Model (with Applications to Protein Folding). J Phys Chem. 1989; 93:6902–6915.
2. Socci ND, Onuchic JN, Wolynes PG. Diffusive Dynamics of the Reaction Coordinate for Protein Folding Funnels. J Chem Phys. 1996; 104:5860–5868.
3. Best RB, Hummer G. Diffusive Model of Protein Folding Dynamics with Kramers Turnover Rate. Phys Rev Lett. 2006; 96:228104. [PubMed: 16803349]

4. Munos V, Eaton WA. A Simple Model for Calculating the Kinetics of Protein Folding from Three-Dimensional Structures. Proc Natl Acad Sci US A. 1999; 96:11311–11316.

5. Kubelka J, Henry ER, Hofrichter J, Eaton WA. Chemical, Physical, and Theoretical Kinetics of an Ultrafast Folding Protein. Proc Natl Acad Sci US A. 2008; 105:18655–18662.

6. Berezhkovskii A, Szabo A. One-Dimensional Reaction Coordinates for Diffusive Activated Rate Processes in Many Dimensions. J Chem Phys. 2005; 122:014503.

7. Kramers HA. Brownian Motion in a Field of Force and the Diffusion Model of Chemical Reactions. Physica. 1940; 7:284–304.

8. Langer JS. Statistical Theory of Decay of Metastable States. Ann Phys. 1969; 54:258–275.

9. Du R, Pande VS, Grosberg AY, Tanaka T, Shakhnovich ES. On the Transition Coordinate for Protein Folding. J Chem Phys. 1998; 108:334–350.

10. Hummer G. From Transition Path to Transition States and Rate Ceofficients. J Chem Phys. 2004; 120:516–523. [PubMed: 15267886]

11. EW, Ren W, Vanden-Eijnden E. Transition Pathways in Complex Systems: Reaction Coordinates, Isocommitor Surfaces, and Transition Tubes. Chem Phys Lett. 2005; 413:242–247.

12. EW, Vanden-Eijnden E. Towards a Theory of Transition Path. J Stat Phys. 2006; 123:503–523.

13. EW, Vanden-Eijnden E. Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events. Annu Rev Phys Chem. 2010; 61:391–420. [PubMed: 18999998]

14. Krivov SV. On Reaction Coordinate Optimality. J Chem Theory Comput. 2013; 9:135–146. [PubMed: 26589017]

15. Vanden-Eijnden E, Venturoli M, Ciccotti G, Elber R. On the Assumptions Underlying Milestoning. J Chem Phys. 2008; 129:174102. [PubMed: 19045328]

16. Onsager L. Initial Recombination of Ions. Phys Rev. 1938; 54:554–557.

17. Peters B, Bolhuis PG, Mullen RG, Shea JE. Reaction Coordinates, One-Dimensional Smoluchowski Equations, and a Test for Dynamical Self-Consistency. J Chem Phys. 2013; 138:054106. [PubMed: 23406097]

18. Risken, H. The Fokker-Planck Equation. Springer; Berlin: 1984. p. 97

19. Metzner P, Schutte C, Vanden-Eijnden E. Transition Path Theory for Markov Jump Processes. Multiscale Model Simul. 2009; 7:1192–1219.

20. Berezhkovskii A, Hummer G, Szabo A. Reactive Flux and Folding Pathways in Network Models of Coarse-Grained Protein Dynamics. J Chem Phys. 2009; 130:205102. [PubMed: 19485483]

21. Johnson ME, Hummer G. Characterization of a Dynamic String Method for the Construction of Transition Pathways in Molecular Reactions. J Phys Chem. 2012; 116:8573–8583.

22. Krivov SV. The Free Energy Landscape Analysis of Protein (FIP35) Folding Dynamics. J Phys Chem. 2011; 115:12315–12324.

APPENDIX: DISCRETE STATE SPACE

We now briefly derive the analogue of some of the above results when the states are discrete (labeled by $i = 1, 2$, etc.) and the dynamics is described by a set of rate or master equations

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = \sum_j K_{ij} p_j(t) \qquad \text{(A.1)}$$

where $p_i(t)$ is the probability of finding the system in state i at time t, K_{ij} is the rate constant for the $j \to i$ transition, and $K_{ii} = -\Sigma_{j \neq i} K_{ij}$. This is the discrete analogue of the multidimensional Smoluchowski equation, eq 2.1. The equilibrium populations p_j^{eq} are solutions of $\sum_j K_{ij} p_j^{\mathrm{eq}} = 0$ and are normalized to unity.

As in the continuous case, we divide the states in the three groups. Let us label states in the intermediate region by m and n. We denote the splitting or commitment probability of an intermediate state m by φ_m. Consider a dividing surface that separates the intermediate states in such a way that all φ's on one side of the surface are greater than those on the other side. Let us adopt the convention that states in one group are labeled by "m" and those in the other group by "n" such that $\varphi_n \geq \varphi_m$ for all m and n. Metzner et al.[19] and Berezhkovskii et al.[20] have independently and in different ways shown that unidirectional reactive flux J through the intermediate region at equilibrium can be expressed in terms of the splitting probabilities as

$$J=\sum_{m,n} K_{mn} p_n^{\mathrm{eq}} (\phi_n - \phi_m) \qquad \text{(A.2)}$$

where the above labeling convention is understood.

Let us now try to reduce the dimensionality of the problem by grouping together states with the same splitting probability. We shall label each group by capital indices M and N so that $\varphi_M = \varphi_m$ for all $m \in M$. In this notation, the sum over all m is equal to the sum over groups M and all $m \in M$, $\Sigma_m = \Sigma_M \Sigma_{m \in M}$.

If we sum eq A.1 over all $m \in$ M and define $P_M(t) = \Sigma_{m \in M} p_m(t)$, we find that

$$\frac{\mathrm{d}P_M(t)}{\mathrm{d}t} = \sum_N \sum_{\substack{m \in M \\ n \in N}} K_{mn} p_n(t) \qquad \text{(A.3)}$$

To close this equation, we invoke the local equilibrium approximation and assume that for all $n \in N$

$$p_n(t) = \frac{p_n^{\mathrm{eq}}}{\sum_{n \in N} p_n^{\mathrm{eq}}} P_N(t) = \frac{p_n^{\mathrm{eq}}}{p_N^{\mathrm{eq}}} P_N(t) \qquad \text{(A.4)}$$

This means that all states with the same φ are assumed to have the same time dependence. This is true only in the limit that interconversion of states with the same φ is much faster than that among states with different φ's. Substituting eq A.4 into eq A.3, we note that eq A.3 can be written in the form of a master equation,

$$\frac{\mathrm{d}P_M(t)}{\mathrm{d}t} = \sum_N K_{MN} P_N(t) \qquad \text{(A.5)}$$

if we define the rate constants for transitions among groups with the same splitting probability by

$$K_{MN}P_N^{\mathrm{eq}}=\sum_{\substack{m\in M\\ n\in N}} K_{mn}p_n^{\mathrm{eq}} \tag{A.6}$$

Although the reduced rate equation, eq A.5, in general gives only an approximate description of the system dynamics, it does give the exact flux *J*. That is the unidirectional reactive flux calculated at equilibrium using the reduced rate matrix,

$$J=\sum_{M,N} K_{MN}P_N^{\mathrm{eq}}(\phi_N-\phi_M) \tag{A.7}$$

is actually exact (i.e., it is identical to the flux in eq A.2). To show this, substitute eq A.6 into eq A.7, use the fact that $\varphi_M = \varphi_m$ for all $m \in$ M, and recall that $\Sigma_m = \Sigma_M \Sigma_{m\in M}$.