Journal of Statistical Physics, Vol. 123, No. 3, May 2006 (© 2006)
DOI: 10.1007/s10955-005-9003-9

Towards a Theory of Transition Paths

Weinan E[1] and Eric Vanden-Eijnden[2]

Received July 19, 2005; accepted November 3, 2005
Published Online: May 16, 2006

We construct a statistical theory of reactive trajectories between two pre-specified sets A and B, i.e. the portions of the path of a Markov process during which the path makes a transition from A to B. This problem is relevant e.g. in the context of metastability, in which case the two sets A and B are metastable sets, though the formalism we propose is independent of any such assumptions on A and B. We show that various probability distributions on the reactive trajectories can be expressed in terms of the equilibrium distribution of the process and the so-called committor functions which give the probability that the process reaches first B before reaching A, either backward or forward in time. Using these objects, we obtain (i) the distribution of reactive trajectories, which gives the proportion of time reactive trajectories spend in sets outside of A and B; (ii) the hitting point distribution of the reactive trajectories on a surface, which measures where the reactive trajectories hit the surface when they cross it; (iii) the last hitting point distribution of the reactive trajectories on the surface; (iv) the probability current of reactive trajectories, the integral of which on a surface gives the net average flux of reactive trajectories across this surface; (v) the average frequency of reactive trajectories, which gives the average number of transitions between A and B per unit of time; and (vi) the traffic distribution of reactive trajectories, which gives some information about the regions the reactive trajectories visit regardless of the time they spend in these regions.

KEY WORDS: Transition path theory; transition state theory; transition path sampling; matastability; reactive trajectories; transition pathways.

The dynamical behavior of many systems arising in physics, chemistry, biology, etc. is dominated by rare but important transition events between long lived states. Understanding the mechanism and computing the rate of these transitions is a topic that has attracted a lot of attention for many years. Most of the theoretical and computational approaches to this problem (see e.g. Ref. 9 for a review) rely on the separation of time scales between the fast evolution of the system within

[1] Department of Mathematics and PACM, Princeton University, Princeton, NJ 08544.
[2] Courant Institute of Mathematical Sciences, New York University, New York, NY 10012.

the long lived states and the slow transition events between these states. In specific situations when this separation of time scales can be explicitly linked to the existence of a small parameter in the system, like e.g. a small noise amplitude in a system whose evolution is governed by a gradient flow with a smooth potential plus noise, a fairly complete mathematical theory can be developed based e.g. on Freidlin-Wentzell theory of large deviations for stochastic differential equations[8] (see also Refs. 1–3, 10, 13, 14, 16, 17 for some specific developments in the context of metastable systems). In practice, however, it may be difficult or impossible to identify such a small parameter explicitly and verify whether the system at hand satisfies the assumptions underlying this theory. Therefore, it is suitable to develop a framework for the description of transition events which does not explicitly rely on the separation of time scales even if the aim is to apply this framework to situations with time scale separation. The classical transition state theory (TST) (see e.g. Ref. 18 for a review and Refs. 20 and 21 for a modern account of the theory) can be viewed as the first attempt in this direction. TST gives the exact average frequency of transitions across an arbitrary dividing surface, which corresponds to the situation when the (two) long lived states are extended so as to exactly partition state-space into two sets with the dividing surface in between. The problem with TST is that its predictions rely heavily on the choice of dividing surface, and in any event it gives minimal information about the mechanism of transition—TST tells where the system crosses the dividing surface, but the theory cannot give a more global picture of the transition mechanism between two separated sets in state-space. A pioneering development in the direction of such a global picture is the theoretical background of the transition path sampling (TPS) technique introduced by Bolhuis, Chandler, Dellago, and Geissler.[4] The main idea behind TPS is to use a path integral formulation to assign probabilistic weights to transition paths (or reactive trajectories) between any two predefined sets in the state-space of the system. This approach is mostly useful when these sets are long lived states, but it does not rely on this assumption. TPS leads to a description of the mechanism of transition in path-space. One may, however, find it more interesting to have an understanding of this mechanism in the actual state-space of the system, and ask for instance about the regions that the reactive trajectories are likely to visit; what the dynamical bottlenecks, or transition state regions, are between the long lived states; etc. Extracting this information from the TPS path ensemble is not straightforward (for some developments in this direction see Refs. 11, 12). In this paper, we propose a framework complementary to that of TPS and construct a state-space based statistical theory of transition pathways. This framework underlies an efficient numerical technique, the so-called finite temperature string method (FTS), for the actual identification of the transition pathways,[6,7,15] but here we shall focus on the theory.

Specifically we shall consider a diffusion process with generator

$$L = b(x) \cdot \nabla + a(x) : \nabla\nabla \tag{1}$$

where $b : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is C^1 and $a : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ is C^2 and satisfies $\langle \xi, a(x)\xi \rangle > 0$ for all $\xi \in \mathbb{R}^n$, $x \in \Omega$. Equipped with Neumann boundary conditions on $\partial\Omega$, (1) defines a strong Markov process with continuous path taking value in Ω with reflection at the boundary. Assume that b and a are such that the process is ergodic with respect to the probability distribution (measure) $\mu(\cdot)$ which is absolutely continuous with respect to the Lebesgue measure:

$$d\mu(x) = m(x)\,dx, \tag{2}$$

for some C^2 probability density function $m(x) > 0$, and denote by $\{X_t\}_{t\in\mathbb{R}}$ an equilibrium path of the process in a given realization. Let $A \subset \Omega$ and $B \subset \Omega$ be two μ-measurable open subsets of Ω which we assume to be disjoint, with a smooth boundary, but not necessary connected (which eventually allows to consider situations with more than two connected sets, $\{A_j\}_{j=1,\ldots,n}$, by iteration: for $k = 1, \ldots, n$, take $A = A_k$, $B = \cup_{j\neq k} A_j$). Since the process is ergodic, the path $\{X_t\}_{t\in\mathbb{R}}$ makes transitions between A and B infinitely often. The question we are interested in is: How do these transitions occur? Clearly given any open μ-measurable set $C \subset \Omega \setminus (A \cup B)$, the proportion of time the process spends in C while it is not in $A \cup B$ is

$$\mu(C)/\mu(\Omega_{AB}) \qquad \text{where} \quad \Omega_{AB} = \Omega \setminus \overline{(A \cup B)}. \tag{3}$$

However, this ratio does not give the proportion of time the process spends in C *while making a transition from A to B* since it may happen, for example, that the process leaves A, enters C, then returns to A before visiting B.

To make things more precise, let us define the ordered family of times $\{t_j^-, t_j^+\}_{j\in\mathbb{Z}}$ such that

$$X_{t_j^-} \in \partial A, \quad X_{t_j^+} \in \partial B, \quad \forall t \in (t_j^-, t_j^+) \ : \ X_t \in \Omega_{AB}. \tag{4}$$

Then

Definition 1. (***AB*-reactive trajectories**) We call AB-reactive trajectory each portion of the trajectory $\{X_t\}_{t\in(t_j^-,t_j^+)}$, $j \in \mathbb{Z}$, during which the process makes a transition from A to B. The set

$$\bigcup_{j\in\mathbb{Z}} \{X_t\}_{t\in(t_j^-,t_j^+)} \equiv \{X_t\}_{t\in R} \quad \text{with} \ \ R := \bigcup_{j\in\mathbb{Z}} (t_j^-, t_j^+) \tag{5}$$

is called the set of *AB-reactive trajectories*.

BA-reactive trajectories can be defined similarly by interchanging the roles of A and B in the definition. We now ask: What is the probability distribution μ_{AB} supported on Ω_{AB} such that the AB-reactive trajectories are ergodic with respect to μ_{AB}? In other words, the probability distribution μ_{AB} must be such that for any

μ-measurable open set $C \subset \Omega_{AB}$, we have

$$\lim_{T\to\infty} \frac{\int_{R\cap[-T,T]} \mathbf{1}(X_t \in C)\, dt}{|R \cap [-T, T]|} = \mu_{AB}(C). \tag{6}$$

where $\mathbf{1}(X_t \in C)$ is the indicator function of the set $\{t : X_t \in C\}$, and $|R \cap [-T, T]|$ is the length of the set $R \cap [-T, T]$.

The distribution μ_{AB} is given in Proposition 2 below. To prepare for this Proposition, let

$$t_A^+(t) = \inf\{t' \geq t : X_{t'} \in \bar{A}\}, \qquad t_B^+(t) = \inf\{t' \geq t : X_{t'} \in \bar{B}\}, \tag{7}$$

be the first entrance times after time t in $\bar{A}$ or $\bar{B}$, respectively, and

$$t_A^-(t) = \sup\{t' \leq t : X_{t'} \in \bar{A}\}, \qquad t_B^-(t) = \sup\{t' \leq t : X_{t'} \in \bar{B}\}, \tag{8}$$

be the last exit times before time t from $\bar{A}$ or $\bar{B}$, respectively.[19] Define

$$q_+(x) = \mathbf{P}_x\{t_B^+(t) < t_A^+(t)\}, \qquad q_-(x) = \mathbf{P}_x\{t_B^-(t) < t_A^-(t)\} \tag{9}$$

where $\mathbf{P}_x$ denotes the probability conditional on $X_t = x$. $q_+(\cdot)$ and $q_-(\cdot)$ map Ω to $[0, 1]$: $q_+(x)$ is the probability conditional on $X_t = x$ that the process will reach first B before reaching A in the future of time t. Similarly, the function $q_-(x)$ is the probability conditional on $X_t = x$ that the process left last A rather than B in the past of time t. We call $q_+(\cdot)$ and $q_-(\cdot)$ the *forward and backward committor functions*, respectively (these functions are also called capacitors in the probability literature[19]). These functions satisfy the backward Kolmogorov equations

$$\begin{cases} 0 = b \cdot \nabla q_+ + a : \nabla\nabla q_+, \\ q_+|_{\partial A} = 0, \quad q_+|_{\partial B} = 1, \quad \partial_{\hat{n}} q_+|_{\partial\Omega} = 0, \end{cases} \tag{10}$$

and

$$\begin{cases} 0 = -b \cdot \nabla q_- + \dfrac{2}{m}\mathrm{div}(am) \cdot \nabla q_- + a : \nabla\nabla q_-, \\ q_-|_{\partial A} = 1, \quad q_-|_{\partial B} = 0, \quad \partial_{\hat{n}} q_-|_{\partial\Omega} = 0, \end{cases} \tag{11}$$

Here $\partial_{\hat{n}}$ denotes the normal derivative on $\partial\Omega$, and the operator at the right hand-side of (11) is the generator of the time-reversed process associated with (1):

$$L^R := -b \cdot \nabla + \frac{2}{m}\mathrm{div}(am) \cdot \nabla + a : \nabla\nabla \tag{12}$$

For time-reversible processes, $-bm + \mathrm{div}(am) = 0$ and $L^R \equiv L$.

We have

Proposition 2. (Distribution of *AB*-reactive trajectories) *The distribution μ_{AB} defined in (6) is given by*

$$d\mu_{AB}(x) = Z_{AB}^{-1}\, q_+(x) q_-(x) m(x)\, dx \tag{13}$$

where

$$Z_{AB} = \int_{\Omega_{AB}} q_+(x) q_-(x) m(x)\, dx$$

Proof: The ratio in the limit at the left-hand side of (6) can be written as

$$\frac{\int_{R\cap[-T,T]} \mathbf{1}(X_t \in C)\, dt}{|R \cap [-T, T]|} = \frac{\int_{-T}^{T} \mathbf{1}(X_t \in C)\mathbf{1}(t_B^+ < t_A^+)\mathbf{1}(t_B^- < t_A^-)\, dt}{\int_{-T}^{T} \mathbf{1}(X_t \in \Omega_{AB})\mathbf{1}(t_B^+ < t_A^+)\mathbf{1}(t_B^- < t_A^-)\, dt}$$

Taking the limit as $T \to \infty$ and using the ergodicity together with the strong Markov property, we deduce that

$$\mu_{AB}(C) = \frac{\int_C \mathbf{P}_x\{t_B^+ < t_A^+ \text{ and } t_B^- < t_A^-\} m(x)\, dx}{\int_{\Omega_{AB}} \mathbf{P}_x\{t_B^+ < t_A^+ \text{ and } t_B^- < t_A^-\} m(x)\, dx}$$

By Markovianity $\mathbf{P}_x\{t_B^+ < t_A^+ \text{ and } t_B^- < t_A^-\} = q_+(x) q_-(x)$ and this completes the proof. □

Example 1. Consider the process generated by the stochastic differential equation

$$dX_t = -\nabla V(X_t)\, dt + \sqrt{2}\, dW_t \tag{14}$$

on $\mathbb{R}^n$. Here W_t is a Wiener process, and $V(\cdot)$ is a C^2-function which grows at infinity sufficiently fast so that

$$Z := \int_{\mathbb{R}^n} e^{-V(x)} dx < \infty \tag{15}$$

(14) defines a process $\{X_t\}_{t\in\mathbb{R}}$ with generator $L = -\nabla V \cdot \nabla + \Delta$ which is ergodic with respect to the distribution

$$d\mu(x) = Z^{-1} e^{-V(x)}\, dx \tag{16}$$

Thus Proposition 2 applies to (14). Since the process $\{X_t\}_{t\in\mathbb{R}}$ is also time-reversible, i.e. $\{X_t\}_{t\in\mathbb{R}}$ and $\{X_{-t}\}_{t\in\mathbb{R}}$ are statistically equivalent, we have that $q_+(x) = 1 - q_-(x) \equiv q(x)$, where $q(\cdot)$ solves the backward Kolmogorov equation[5]

$$0 = -\nabla V \cdot \nabla q + \Delta q, \qquad q|_{x\in A} = 0, \quad q|_{x\in B} = 1, \tag{17}$$

The probability density function associated with μ_{AB} in the case when

$$\begin{aligned} V(x_1, x_2) &= \tfrac{5}{2}\left(1 - x_1^2\right)^2 + 5x_2^2 \\ A &= \{x_1 < -0.8\}, \quad B = \{x_1 > 0.8\} \end{aligned} \tag{18}$$

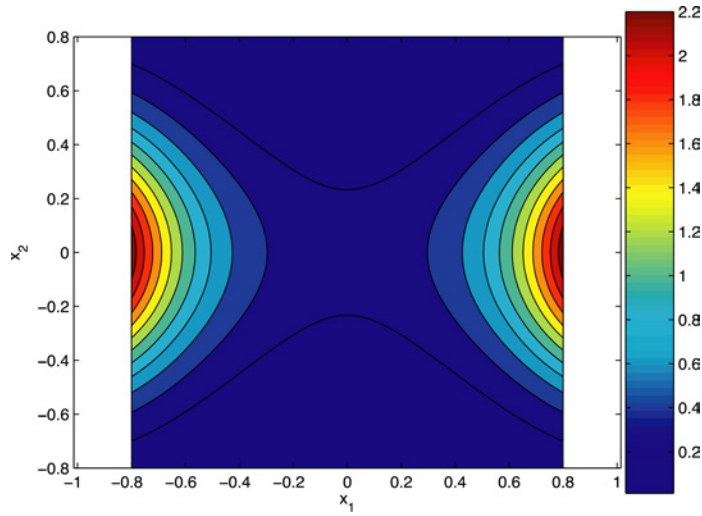

Fig. 1. The level sets of the equilibrium probability density function $\bar{Z}^{-1}e^{-V(x_1,x_2)}$ associated with (14) with the parameters as in (18) (this density is shown normalized in $\mathbb{R}^2/(A \cup B)$, i.e. $\bar{Z} = \bar{Z}_1\bar{Z}_2$, with $\bar{Z}_1 = \int_{-0.8}^{0.8} e^{-\frac{5}{2}(1-x_1^2)^2}\,dx_1$ and $Z_2 = \int_{\mathbb{R}} e^{-5x_2^2}\,dx_2$). The left white strip shows part of $A = \{x < -0.8\}$, the right white strip shows part of $B = \{x > 0.8\}$. The level sets of the density coincide with the level sets of $V(x_1, x_2)$. This graph shows that, when the trajectory leaves the sets A and B, it preferably visits regions near the two minima of $V(x_1, x_2)$ located at $(x_1, x_2) = (\pm 1, 0)$. Of course most of these excursions out of A and B do not lead to transitions between A and B.

is shown in Fig. 2. Here $q_+(x_1, x_2) = 1 - q_-(x_1, x_2) \equiv q(x_1)$ with

$$q(x_1) = \frac{\int_{-0.8}^{x_1} e^{\frac{5}{2}(1-z^2)^2}\,dz}{\int_{-0.8}^{0.8} e^{\frac{5}{2}(1-z^2)^2}\,dz} \tag{19}$$

The density associated with μ_{AB} is to be compared with $\bar{Z}^{-1}e^{-V(x_1,x_2)}$ where $\bar{Z} = \bar{Z}_1\bar{Z}_2$ with $\bar{Z}_1 = \int_{-0.8}^{0.8} e^{-\frac{5}{2}(1-x_1^2)^2}\,dx_1$ and $Z_2 = \int_{\mathbb{R}} e^{-5x_2^2}\,dx_2$ shown in Fig. 1.

Example 2. Consider the process generated by

$$\begin{cases} dX_t = U_t\,dt, \\ dU_t = -\nabla_x V(X_t)\,dt - U_t\,dt + \sqrt{2}\,dW_t \end{cases} \tag{20}$$

on $(x, u) \in \mathbb{R}^n \times \mathbb{R}^n$ (x is referred to as the configuration, u as the velocity). Assuming that $V(\cdot)$ is C^2 and grows sufficiently fast at infinity so that (15) is satisfied, the process generated by (20) is ergodic with respect to

$$d\mu(x, u) = Z_H^{-1}e^{-H(x,u)}\,dx\,du \tag{21}$$

where $H(x,u) = \frac{1}{2}|u|^2 + V(x)$ and $Z_H = \int_{\mathbb{R}^n \times \mathbb{R}^n} e^{-H(x,u)}\,dx\,du$. Since the generator associated with (20), $L = u \cdot \nabla_x - \nabla_x V \cdot \nabla_u - u \cdot \nabla_u + \Delta_u$, is degenerate, we need to take some care in applying Proposition 2. To avoid difficulties, we will assume that the sets A and B are defined as follows. First, we define two open connected C^1-sets $A' \subset \mathbb{R}^n$ and $B' \subset \mathbb{R}^n$ in configuration space. Then we extend them in the full state-space as

$$\begin{aligned} A &= \{(x,u) : x \in A',\, H(x,u) \leq \min_{x \in \partial A'} V(x)\}, \\ B &= \{(x,u) : x \in B',\, H(x,u) \leq \min_{x \in \partial B'} V(x)\}. \end{aligned} \tag{22}$$

It can be checked that Proposition 2 applies in this case. Since $\{X_t, U_t\}_{t\in\mathbb{R}}$ is statistically equivalent to $\{X_{-t}, -U_{-t}\}_{t\in\mathbb{R}}$, it follows that $q_+(x,u) = 1 - q_-(x,-u) \equiv q(x,u)$ and (13) reduces to

$$d\mu_{AB}(x,u) = \bar{Z}_{AB}^{-1} q(x,u)(1 - q(x,-u))e^{-H(x,u)}\,dx\,du, \tag{23}$$

where $\bar{Z}_{AB} = \int_{\Omega_{AB}\times\mathbb{R}^n} q(x,u)(1-q(x,-u))e^{-H(x,u)}\,dx\,du$ and $q(\cdot)$ satisfies[5]

$$\begin{cases} 0 = u \cdot \nabla_x q - \nabla_x V \cdot \nabla_u q - u \cdot \nabla_u q + \Delta_u q, \\ q|_{(x,u)\in\partial A} = 0, \quad q|_{(x,u)\in\partial B} = 1. \end{cases} \tag{24}$$

The probability distribution μ_{AB} gives the information about the proportion of time AB-reactive trajectories spend in any given subset of Ω_{AB}. If A and B are metastable sets, i.e. sets with small volumes that concentrate most of the probability, there must be some dynamical bottlenecks between these sets where the AB-reactive trajectories are likely to spend most of their time—in the example shown in Figs. 1 and 2 the dynamical bottleneck is the region near the saddle point $(x_1, x_2) = (0,0)$ where the distribution μ_{AB} does indeed concentrate. Thus Proposition 2 allows one to identify these dynamical bottlenecks or *transition state regions*. On the other hand, it may be useful to identify the regions the reactive trajectories are likely to visit regardless of the amount of time these trajectories spend in these regions. To help characterize these regions, we introduce several objects. First in Proposition 3, we introduce the *hitting point distribution of AB-reactive trajectories* which tells where the AB-reactive trajectories hit a surface S when they cross that surface. Next in Proposition 4 we show that, provided that S is a level set of $q_+(\cdot)$, the hitting point distribution of AB-reactive trajectories is also the *last hitting point distributions of AB-reactive trajectories*, which tells where the AB-reactive trajectories last leave this surface. In Proposition 5, we identify the *probability current of AB-reactive trajectories*, the integral of which on any surface S gives the net average flux of AB-reactive trajectories across this surface. In Proposition 6, we derive the *average frequency of AB-reactive trajectories*, which gives the average number of transitions from A to B per unit

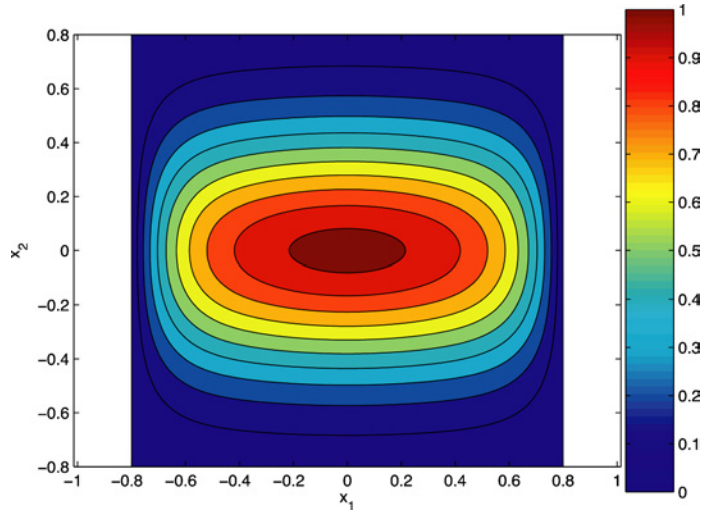

Fig. 2. The level sets of the equilibrium probability density function $Z_{AB}^{-1} q(x_1)(1 - q(x_1))e^{-V(x_1,x_2)}$ associated with distribution of AB-reactive trajectories μ_{AB} in the example of (14) with the parameters as in (18). Here $\mu_{AB} = \mu_{BA}$ by time reversibility and $q(x_1)$ is given in (19). Notice that the density concentrates around the saddle point $(x_1, x_2) = (0, 0)$ of $V(x_1, x_2)$. This graph shows that during the transitions, the reactive trajectories spend most of their time in the region around the saddle point. This region is the dynamical bottleneck of the reaction and therefore qualifies as the *transition state region*.

of time. Finally in Proposition 8 we derive the *AB-traffic distribution* which gives a measure of the regions that the AB-reactive trajectories visit irrespective of the time they spend in these regions.

To begin, let S be a piecewise C^1 surface of co-dimension 1 contained in Ω_{AB}, and let $d\sigma_S(x)$ be the surface element (Lebesgue measure) on S. Denote by

$$d\nu_S(x) = C_S^{-1} m(x)\, d\sigma_S(x) \quad \text{with} \quad C_S = \int_S m(x)\, d\sigma_S(x), \tag{25}$$

the distribution supported on S induced by μ. The distribution ν_S tells where the trajecory (not just the reactive portions of it) hit S when it crosses this surface. The following proposition gives the corresponding distribution $\nu_{S,AB}$ which tells where the AB-reactive trajectories hit S when they cross this surface

Proposition 3. (Hitting point distribution of *AB*-reactive trajectories) *The distribution of hitting points on S by the AB-reactive trajectories is*

$$d\nu_{S,AB}(x) = C_{S,AB}^{-1} q_+(x) q_-(x) m(x)\, d\sigma_S(x) \tag{26}$$

where

$$C_{S,AB} = \int_S q_+(x) q_-(x) m(x)\, d\sigma_S(x) \tag{27}$$

Proof: Use the identity

$$\nu_{S,AB}(C \cap S) = \lim_{d\to 0+} \lim_{T\to\infty} \frac{\int_{-T}^{T} \mathbf{1}(X_t \in C \cap S_d)\mathbf{1}(t_B^+ < t_A^+)\mathbf{1}(t_B^- < t_A^-)\, dt}{\int_{-T}^{T} \mathbf{1}(X_t \in S_d)\mathbf{1}(t_B^+ < t_A^+)\mathbf{1}(t_B^- < t_A^-)\, dt}$$

where $C \in \Omega_{AB}$ is any μ-measurable subset and S_d is the slab of thickness d around S: $S_d := \{x : \text{dist}(x, S) < d\}$. Proceeding as in the proof of Proposition 2, we deduce that

$$\nu_{S,AB}(C \cap S) = \lim_{d\to 0+} \frac{\int_{C\cap S_d} q_+(x) q_-(x) m(x)\, dx}{\int_{S_d} q_+(x) q_-(x) m(x)\, dx} = \frac{\int_{C\cap S} q_+(x) q_-(x) m(x)\, d\sigma_S(x)}{\int_S q_+(x) q_-(x) m(x)\, d\sigma_S(x)},$$

consistent with (26). □

The distribution $\nu_{S,AB}$ gives more than the hitting point distributions if one uses a special class of surfaces S, namely the level sets of $q_+(x)$ which we denote as:

$$S_+(z) = \{x : q_+(x) = z\}, \qquad (z \in (0, 1)). \tag{28}$$

Since $q_+(\cdot)$ is C^1 by assumption, $S_+(z)$ is a C^1 surface of co-dimension 1 which partitions Ω into two sets, one containing A and the other containing B. The family $\{S_+(z)\}_{z\in(0,1)}$ defines a co-dimension 1 foliation of Ω_{AB}, i.e.

$$S_+(z) \cap S_+(z') = \emptyset \quad \text{if } z \neq z', \qquad \bigcup_{z\in(0,1)} S_+(z) = \Omega_{AB}.$$

The foliation $\{S_+(z)\}_{z\in[0,1]}$ is also called the *reaction coordinate* associated with the function $q_+(\cdot)$ (and, by extension, we sometimes refer to $q_+(\cdot)$ itself as the reaction coordinate).

Given an AB-reactive trajectory $\{X_t\}_{t\in(t_j^-,t_j^+)}$, define the last hitting point on $S_+(z)$ as the point $x \in S_+(z)$ such that

$$X_{t_j^\star(z)} = x \quad (t_j^\star(z) \in (t_j^-, t_j^+)), \quad \forall t \in (t_j^\star(z), t_j^+) \ : \ X_t \in \{x : q_+(x) > z\} \tag{29}$$

Define also the distribution of last hitting point of the AB-reactive trajectories as the distribution $\nu_{z,AB}$ supported on $S_+(z)$ such that for any μ-measurable subset

$C \in \Omega_{AB}$, we have

$$\nu_{z,AB}(C \cap S_+(z)) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{j=-N}^{N} \mathbf{1}(X_{t_j^\star(z)} \in C \cap S_+(z)) \tag{30}$$

Then

Proposition 4. (Last hitting point distribution of *AB*-reactive trajectories) *For each $z \in (0, 1)$, the last hitting point distribution $\nu_{z,AB}$ defined in (30) coincides with the distribution $\nu_{S,AB}$ defined in (26) evaluated on $S \equiv S_+(z)$, i.e*

$$\nu_{z,AB} \equiv \nu_{S_+(z),AB}$$

Since by definition $q_+(x) = z$ is constant on $S_+(z)$, on this surface the distribution $\nu_{S_+(z),AB} \equiv \nu_{z,AB}$ reduces to

$$d\nu_{z,AB}(x) = C_{AB}^{-1}(z) q_-(x) m(x)\, d\sigma_z(x) \tag{31}$$

where $d\sigma_z(x)$ is the surface element on $S_+(z)$, $d\sigma_z \equiv d\sigma_{S_+(z)}$, and

$$C_{AB}(z) = \int_{S_+(z)} q_-(x) m(x)\, d\sigma_z(x). \tag{32}$$

Proof: Let

$$t^+_{\Omega_+(z)} = \inf\{t' \geq t : X_{t'} \in \Omega_+(z)\}, \qquad \Omega_+(z) = \{x : q(x) \leq z\}.$$

and for $x \in \Omega_{AB} \setminus \Omega_+(z)$, define

$$\hat{q}_+(x) = \mathbf{P}_x\{t_B^+ < t^+_{\Omega_+(z)}\}.$$

$\hat{q}_+(x)$ is the probability conditional on $X_t = x$ that the process will reach first B before reaching $\Omega_+(z)$ in the future of time t. By strong Markovianity the probability to reach first A before reaching B starting from x (which is $1 - q_+(x)$) is the probability to reach first $\Omega_+(z)$ before reaching B starting from x (which is $1 - \hat{q}_+(x)$) times the probability to reach first A before reaching B starting form $S_+(z)$ (which is $1 - z$). Therefore

$$1 - q_+(x) = (1 - \hat{q}_+(x))(1 - z) \quad \Leftrightarrow \quad \hat{q}_+(x) = \frac{q_+(x) - z}{1 - z} \tag{33}$$

(This identity can also be derived by noting that $\hat{q}_+(\cdot)$ satisfies (10) on $\Omega_{AB} \setminus \Omega_+(z)$ with boundary conditions $\hat{q}_+|_{S_+(z)} = 0$, $\hat{q}_+|_{\partial B} = 1$, $\partial_{\hat{n}}\hat{q}_+|_{\partial\Omega} = 0$.) For any surface $S \subset \Omega_{AB} \setminus \Omega_+(z)$, by a straightforward generalization of Proposition 3, we can define the distribution of hitting points on S of the tails of AB-reactive

trajectories on $(t_j^\star(z), t_j)$, $\{X_t\}_{t\in\cup_j(t_j^\star(z),t_j^+)}$, during which they never revisit $\Omega_+(z)$. This distribution is (26) with $q_+(x)$ replaced by $\hat{q}_+(x)$, i.e. using (33)

$$d\hat{\nu}_{S,AB}(x) = \hat{C}_{S,AB}^{-1}(q_+(x) - z)q_-(x)m(x)\,d\sigma_S(x),$$

where

$$\hat{C}_{S,AB} = \int_S (q_+(x) - z)q_-(x)m(x)\,d\sigma_S(x)$$

In particular, if $S \equiv S_+(z')$ with $z' > z$, we have $\hat{\nu}_{S_+(z'),AB} \equiv \nu_{z',AB}$ with the distribution $\nu_{z,AB}$ given in (31). The proposition then follows by letting $z' \to z$, since in this limit the only hitting points on $S_+(z')$ by $\{X_t\}_{t\in(t_j^\star(z),t_j^+)}$ are the last hitting points. □

Next, we consider the probability current of AB-reactive trajectories. To motivate this concept, consider the evolution equation

$$\begin{cases} \dfrac{\partial \hat{m}}{\partial t} = -\nabla\cdot(b(x)\hat{m}(x,t) - \mathrm{div}(a(x)\hat{m}(x,t))) \\ \hat{m}(\cdot,t)|_{\partial A} = m(\cdot), \quad \hat{m}(\cdot,t)|_{\partial B} = 0 \quad \partial_{\hat{n}}\hat{m}(\cdot,t)|_{\partial\Omega} = 0, \end{cases} \tag{34}$$

for some initial condition at $t = -T$. This equation described the evolution of the probability density $\hat{m}(\cdot,\cdot) : \Omega_{AB}\times[-T,\infty) \to \mathbb{R}_+$, of a process kept at equilibrium on ∂A and absorbed on ∂B. Such a process allows one to focus on the transitions from A to B of the original process and ignore the transitions from B to A. (34) can be written in conservation form as

$$\frac{\partial \hat{m}}{\partial t} = -\mathrm{div}\, J(x,t) \quad \text{where } J(x,t) = b(x)\hat{m}(x,t) - \mathrm{div}(a(x)\hat{m}(x,t)). \tag{35}$$

We claim that $\hat{m}(\cdot,t)$ assumes the following stationary value as the initial condition is pushed back towards the infinite past:

$$m(\cdot,t) \to m(\cdot)q_-(\cdot) \qquad \text{as } T\to\infty. \tag{36}$$

This equation follows from the strong Markov property and expresses the fact that at equilibrium, the probability density that the process associated with (35) be at position x is proportional to the probability density that the original process be at x (which is $m(x)$) times the probability that it came from A rather than B (which is $q_-(x)$). Notice that $q_+(\cdot)$ does not appear in (36) because the process whose density obeys (37) is not constrained to go to A after leaving B and instead may return to B before reaching A. When $\hat{m} \to mq_-$, $J \to J_{AB}$, the probability current of AB-reactive trajectories. Therefore

Proposition 5. (Probability current of *AB*-reactive trajectories) *The vector field $J_{AB} : \Omega_{AB} \to \mathbb{R}^n$ given by*

$$J_{AB}(x) = b(x)m(x)q_-(x) - div(a(x)m(x)q_-(x)) \tag{37}$$

is the probability current of AB-reactive trajectories.

In other words, given any piecewise C^1 surface $S \subset \Omega_{AB}$ of co-dimension 1,

$$\Gamma_{S,AB} = \int_S \hat{n}_S(x) \cdot J_{AB}(x)\, d\sigma_S(x) \tag{38}$$

gives the average net flux of AB-reactive trajectories across S. Here $\hat{n}_S(x)$ is the unit outward pointing normal to S and $d\sigma_S(x)$ is the surface element on S.

We shall give an alternative proof of Proposition 5 based on a direct definition of the probability current of AB-reactive trajectories (see (50) below). Since this proof is somewhat tedious and rather technical, we defer it till the end of the paper. Notice that for time-reversible processes for which $q_+ = 1 - q_- \equiv q$ and $bm = \mathrm{div}(am)$ the probability current reduces to

$$J_{AB}(x) = m(x)a(x)\nabla q(x). \tag{39}$$

Notice also that Proposition 5 implies that the net average flux across any surface defined as the boundary of a region in Ω_{AB} (i.e. a region which does not contain A nor B) is zero, $\Gamma_{\partial C,AB} = 0$ if $C \subset \Omega_{AB}$. This follows from the fact that probability current in (37) is divergence free:

$$\mathrm{div}\, J_{AB}(x) = q_-(x)(L^* m)(x) + m(x)(L^R q_-)(x) = 0. \tag{40}$$

and that (38) can by expressed as (using Gauss' theorem):

$$\Gamma_{\partial C,AB} = \int_{\partial C} \hat{n}_{\partial C}(x) \cdot J_{AB}(x)\, d\sigma_{\partial C}(x) = \int_C \mathrm{div}\, J_{AB}(x)\, dx = 0, \tag{41}$$

Therefore (38) is especially useful if S is any dividing surface between A and B, i.e. a surface such that A is on one side of it and B on the other (in this case we must take $\hat{n}_S(x)$ pointing toward B). In this case indeed $\Gamma_{S,AB}$ is independent of the particular dividing surface we pick and is the *average frequency of AB-reactive trajectories*, which gives the averaged number of transitions from A to B per unit of time. We shall show that it can also be expressed as:

Proposition 6. (Average frequency of *AB*-reactive trajectories) *The average number of transitions from A to B per unit of time is given by:*

$$k_{AB} = \int_{\Omega_{AB}} \langle \nabla q_+, a \nabla q_+ \rangle m dx = \int_{\Omega_{AB}} \langle \nabla q_-, a \nabla q_- \rangle m dx \tag{42}$$

Notice that the average number of transitions from A to B is the same as the average number of transitions from B to A since we are at equilibrium, i.e. $k_{AB} = k_{BA}$, and it is also half the average number of transitions between A and B.

Proof: If S is any piecewise C^1 dividing surface in Ω_{AB}, we have that $k_{AB} \equiv \Gamma_{S,AB}$. Therefore, using $S \equiv S_+(z)$ in (38) and the co-area formula,

$$d\sigma_z(x) = |\nabla q_+(x)|\delta(q_+(x) - z)\,dx,$$

the average frequency can be expressed as

$$\begin{aligned} k_{AB} &= \int_{S_+(z)} \hat{n}_{S_+(z)} \cdot (mq_-b - \operatorname{div}(amq_-))\,d\sigma_z(x) \\ &= \int_{\Omega_{AB}} \nabla q_+ \cdot (mq_-b - \operatorname{div}(amq_-))\delta(q_+(x) - z)\,dx. \end{aligned}$$

Integrating by part the second term in the integral and using (10) we deduce that

$$\begin{aligned} k_{AB} &= \int_{\Omega_{AB}} mq_-(b \cdot \nabla q_+ + a : \nabla\nabla q_+)\delta(q_+(x) - z)\,dx \\ &\quad + \int_{\Omega_{AB}} mq_-a : \nabla q_+ \nabla q_+ \delta'(q_+(x) - z)\,dx \\ &= -\frac{d}{dz}\int_{\Omega_{AB}} mq_-a : \nabla q_+ \nabla q_+ \delta(q_+(x) - z)\,dx. \end{aligned}$$

Integrating both term of this equality on z from z to 1 using the fact that k_{AB} is independent on z and $q_- = 0$ on ∂B (i.e. when $z = 1$) this gives

$$(1 - z)k_{AB} = \int_{\Omega_{AB}} mq_-a : \nabla q_+ \nabla q_+ \delta(q_+(x) - z)\,dx.$$

Integrating again on z from 0 to 1, we arrive at

$$\frac{1}{2}k_{AB} = \int_{\Omega_{AB}} mq_-a : \nabla q_+ \nabla q_+ dx.$$

Since $k_{AB} = k_{BA}$, and k_{BA} is given by a similar expression with $q_\pm$ replaced by $1 - q\pm$, we also have

$$\frac{1}{2}k_{AB} = \int_{\Omega_{AB}} m(1 - q_-)a : \nabla q_+ \nabla q_+ dx.$$

which can be added to the previous expression to give the first equality in (42). The second equality follows by repeating the proof using the time-reversed process. □

The probability flux amplitude, $|J_{AB}(\cdot)|$, can in principle be taken as a measure of the regions that the AB-reactive trajectories are likely to visit regardless of

the amount of time they spend in these regions. For convenience, we shall however characterize these *transition regions* (or *transition tubes*) by introducing instead yet another object, called the *AB-traffic distribution*. This distribution has the advantage over $|J_{AB}(\cdot)|$ that it explicitly depends only on the equilibrium density, $m(\cdot)$, and the committor functions, $q_{\pm}(\cdot)$, and not on the actual drift $b(\cdot)$ and diffusion tensor $a(\cdot)$ of the process. Therefore, it may be more convenient to use in situation when $m(\cdot)$ and $q_{\pm}(\cdot)$ are extracted from a time-series whose underlying $b(\cdot)$ and $a(\cdot)$ may not be explicitly known (as is often the case in applications). We will introduce the AB-traffic distribution by appropriately extending the family of last hitting point distributions $\{\nu_{z,AB}\}_{z\in(0,1)}$ to a distribution on Ω_{AB}. To this end, let us define

Definition 7. (**Pruned *AB*-reactive trajectories**) For each $j \in \mathbb{Z}$, the subset $\{X_{t_j^\star(z)}\}_{z\in(0,1)}$ consisting of the sequence of last hitting points from $S_+(z), z \in (0, 1)$ of the jth AB-reactive trajectory is called the pruned AB-reactive trajectory. The set

$$\{X_{t_j^\star(z)}\}_{j\in\mathbb{Z},z\in(0,1)} \tag{43}$$

is called the set of pruned AB-reactive trajectories.

For each $j \in \mathbb{Z}$, $X_{t_j^\star(z)}$ viewed as a function of z, $X_{t_j^\star(\cdot)} : (0, 1) \to \Omega_{AB}$, gives the succession of last hitting points of the jth AB-reactive trajectory on $S_+(z)$ when z varies from 0 to 1.

The natural way to associate a distribution to the pruned AB-reactive trajectories is by looking for the distribution μ_{AB}^F such that for any μ-measurable subset $C \in \Omega_{AB}$, we have (compare (30))

$$\mu_{AB}^F(C) = \tilde{\mu}_{AB}^F(C)/\tilde{\mu}_{AB}^F(\Omega_{AB}) \tag{44}$$

where

$$\tilde{\mu}_{AB}^F(C) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{j=-N}^{N} \int_0^1 \mathbf{1}\Big(X_{t_j^\star(z)} \in C\Big) \frac{dz}{\Big|\nabla q_+\Big(X_{t_j^\star(z)}\Big)\Big|}. \tag{45}$$

To justify (45), note that for each $j \in \mathbb{Z}$, $\{X_{t_j^\star(z)}\}_{z\in(0,1)}$ consists of the subset of the jth AB-reactive trajectory pruned of the loops where this trajectory revisit regions it already visited (which by Proposition 4 introduces no bias). The factor $1/|\nabla q_+(X_{t_j^\star(z)})|$ is included to make (45) less sensitive to the actual parametrization of the pruned trajectories in terms of z. In fact, it is easy to see that this factor makes (45) is gauge invariant, in the sense that it is left invariant by any transformation $z \to z' = f(z), q(\cdot) \to q'(\cdot) = f(q(\cdot))$, where $f : (0, 1) \to (0, 1)$ is a strictly monotone function. Thus μ_{AB}^F indeed measures which regions the AB-trajectories visit regardless of the time they spend in these regions. We call this distribution the *AB-traffic distribution*. We have

Proposition 8. (*AB*-traffic distribution) *The AB-traffic distribution defined in (45) is*

$$\mu_{AB}^F(dx) = C^{-1} C_{AB}^{-1}(q_+(x)) q_-(x) m(x)\, dx. \tag{46}$$

where

$$C = \int_{\Omega_{AB}} C_{AB}^{-1}(q_+(x)) q_-(x) m(x)\, dx \tag{47}$$

Proof: Notice first that $\mathbf{1}(X_{t_j^\star(z)} \in C) = \mathbf{1}(X_{t_j^\star(z)} \in C \cap S_+(z))$ since $X_{t_j^\star(z)} \in S_+(z)$. Therefore, using (30), (45) can be written as

$$\tilde{\mu}_{AB}^F(C) = \int_0^1 C_{AB}^{-1}(z) \int_{C \cap S_+(z)} q_-(x) m(x) \frac{d\sigma_z(x)}{|\nabla q_+(x)|} dz.$$

Using the co-area formula, we deduce that

$$\begin{aligned}
\tilde{\mu}_{AB}^F(C) &= \int_0^1 C_{AB}^{-1}(z) \int_C q_-(x) \delta(q_+(x) - z) m(x)\, dx dz \\
&= \int_C C_{AB}^{-1}(q_+(x)) q_-(x) \Big(\int_0^1 \delta(q_+(x) - z)\, dz \Big) m(x)\, dx \\
&= \int_C C_{AB}^{-1}(q_+(x)) q_-(x) m(x)\, dx,
\end{aligned}$$

which, using (44), is consistent with (46). □

For time-reversible processes such that $\{X_t\}_{t\in\mathbb{R}}$ and $\{X_{-t}\}_{t\in\mathbb{R}}$ are statistically equivalent and $q_+(x) = 1 - q_-(x) \equiv q(x)$, from (32) we have

$$\begin{aligned}
C_{AB}(z) &= \int_{q(x)=z} (1 - q(x)) m(x)\, d\sigma_z(x) \\
&= (1 - z) \int_{q(x)=z} m(x)\, d\sigma_z(x)
\end{aligned}$$

As a result μ_{AB}^F reduces to

$$d\mu_{AB}^F(x) = C_F^{-1}(q(x)) m(x)\, dx \qquad (\{X_t\}_{t\in\mathbb{R}} \stackrel{d}{=} \{X_{-t}\}_{t\in\mathbb{R}}) \tag{48}$$

where

$$C_F(z) = \int_{q(x)=z} m(x)\, d\sigma_z(x). \tag{49}$$

The distribution μ_{AB}^F corresponding to (14) with the parameters as in (19) is shown in Fig. 3. Note that contrast to the distribution μ_{AB} shown in Fig. 2 which

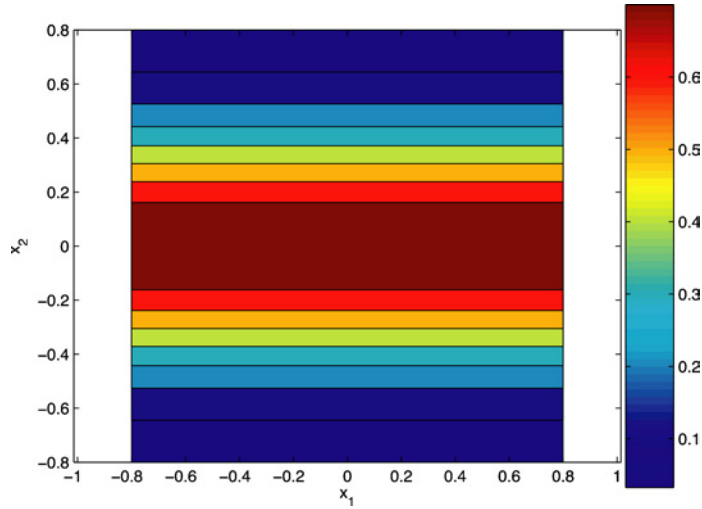

Fig. 3. The level sets of the probability density function associated with the traffic distribution μ^F_{AB} in the example of (14) with the parameters as in (18). In this example, this density is simply $\frac{5}{8}\bar{Z}_2^{-1}e^{-5x_2^2}$. Notice that the density is concentrated on the channel around the x_1-axis connecting A to B. This region qualifies as the *transition path region*. The probability current in this example is $J_{AB}(x) = Z^{-1}\bar{Z}_1^{-1}e^{-5x_2^2}\hat{e}_1$, where e_1 is the unit vector in the x_1-direction. The current is concentrated in the transition path region, and the flow lines of the current are parallel to the level set of the probability density function associated with μ^F_{AB}. The average frequency is given by $k_{AB} = \bar{Z}_1 Z_1^{-1} = \int_{-0.8}^{0.8} e^{-\frac{5}{2}(1-x_1^2)^2}\,dx_1 / \int_{\mathbb{R}} e^{-\frac{5}{2}(1-x_1^2)^2}\,dx_1 \approx 0.32$.

is concentrated in the region near the saddle point $(x_1, x_2) = (0, 0)$ (which is here the transition state region), μ^F_{AB} is concentrated on a channel centered around the x_1-axis and connecting A to B: this channel is the transition paths region in this example.

We conclude this paper with a proof of Proposition 5.

Proof of Proposition 5. The average net flux across any surface in Ω_{AB} (not necessarily a dividing one) can be expressed as the difference between the fluxes across dividing surfaces which have common portions. Therefore, by Gauss' theorem and since the candidate probability current J_{AB} is divergence free by (40), it suffices to verify that (38) holds with the probability current in (37) for any dividing surface. Let S be a piecewise C^1 dividing surface of co-dimension 1 in Ω_{AB} and denote by Ω_A the closed set which contains S and all the portion of Ω on the side of S which contains A. The average net flux of AB-reactive trajectories

across any such S can be expressed as the limit

$$\Gamma_{S,AB} = \lim_{\tau\to 0+} \frac{1}{\tau} \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathbf{1}(X_t \in \Omega_A)\mathbf{1}(t_B^-(t) < t_A^-(t)) \times \mathbf{1}(X_{t+\tau} \in \Omega_A^c)\mathbf{1}(t_B^+(t+\tau) < t_{\Omega_A}^+(t+\tau))\, dt \tag{50}$$

where $\Omega_A^c = \Omega \setminus \Omega_A$ is the complement of Ω_A in Ω and

$$t_{\Omega_A}^+(t) = \inf\{t' \geq t : X_{t'} \in \bar{\Omega}_A\}.$$

By ergodicity and strong Markovianity, (50) is

$$\Gamma_{S,AB} = \lim_{\tau\to 0+} \frac{1}{\tau} \int_{\Omega_A} m(x) q_-(x) \int_{\Omega_A^c} p_\tau^x(y) \hat{q}_+(y)\, dy dx. \tag{51}$$

Here

$$\hat{q}_+(x) = \mathbf{P}_x\{t_B^+(t) < t_{\Omega_A}^+(t)\} \tag{52}$$

and $p_\tau^x(y)$ is the transition probability density function of the process: for all μ-measurable set $C \subset \Omega$, $\mathbf{P}_x\{X_t \in C\} = \int_C p_t^x(y)\, dy$. To compute the limit as $\tau \to 0+$ in (51) use:

Lemma 9. *As $\tau \to 0+$, we have*

$$\frac{1}{\tau} \int_{\Omega_A^c} p_\tau^x(y) \hat{q}_+(y)\, dy dx \to \langle \hat{n}_S(x), a(x)\hat{q}_+(x)\rangle d\sigma_S(x) \tag{53}$$

in the sense of distributions, where $\hat{n}_S(x)$ is the unit normal to S pointing outward of Ω_A and $d\sigma_S(x)$ is the surface element (Lebesgue measure) on S.

The distribution on S at the right hand-side of (53) is related to the so-called equilibrium distribution introduced in potential theory.[19]

Proof: Use

$$\int_\Omega p_\tau^x(y) f(y)\, dy = f(x) + (Lf)(x)\tau + o(\tau)$$

to deduce that

$$\int_{\Omega_A^c} p_\tau^x(y) \hat{q}_+(y)\, dy = q_+(x)\mathbf{1}(x \in \Omega_A^c) + \tau L\big(\hat{q}_+(x)\mathbf{1}\big(x \in \Omega_A^c\big)\big) + o(\tau).$$

For $x \in \Omega_A$, the first term is zero which implies that

$$\frac{1}{\tau} \int_{\Omega_A^c} p_\tau^x(y) \hat{q}_+(y)\, dy dx \to L\big(\hat{q}_+(x)\mathbf{1}\big(x \in \Omega_A^c\big)\big)\, dx \qquad (x \in \Omega_A)$$

as $\tau \to 0+$. Since for all test function $f : \Omega \to \mathbb{R}$ (using integration by parts),

$$\int_\Omega f(x)\nabla\big(\mathbf{1}\big(x \in \Omega_A^c\big)\big)dx = -\int_{\Omega_A^c} \nabla f(x)\,dx$$
$$= \int_S f(x)\hat{n}_S(x)\,d\sigma_S(x)$$

we have the distribution identity $\nabla(\mathbf{1}(x \in \Omega_A^c))\,dx = \hat{n}_S(x)\,d\sigma_S(x)$. Similarly $a(x) : \nabla(\hat{q}_+(x)\nabla\mathbf{1}(x \in \Omega_A^c))\,dx = 0$ since (using integration by parts and $\hat{q}_+ = 0$ on S)

$$\int_\Omega f(x)a(x) : \nabla\big(\hat{q}_+(x)\nabla\mathbf{1}\big(x \in \Omega_A^c\big)\big)\,dx$$
$$= -\int_\Omega \hat{q}_+(x)\big\langle \mathrm{div}(a(x)\nabla f(x)),\, \nabla\mathbf{1}\big(x \in \Omega_A^c\big)\big\rangle dx$$
$$= -\int_S \hat{q}_+(x)\langle \mathrm{div}(a(x)\nabla f(x)),\, \hat{n}_S(x)\rangle d\sigma_S(x) = 0.$$

Therefore by expansion we arrive at

$$\begin{aligned} L\big(\hat{q}_+(x)\mathbf{1}\big(x \in \Omega_A^c\big)\big)\,dx &= (L\hat{q}_+)(x)\mathbf{1}\big(x \in \Omega_A^c\big)\,dx + \hat{q}_+(x)\langle \hat{n}_S(x), b(x)\rangle d\sigma_S(x) \\ &\quad + \langle \hat{n}_S(x), a(x)\nabla\hat{q}_+(x)\rangle d\sigma_S(x) \\ &\quad + a(x) : \nabla\big(\hat{q}_+(x)\nabla\mathbf{1}\big(x \in \Omega_A^c\big)\big)\,dx \\ &= \langle \hat{n}_S(x), a(x)\nabla\hat{q}_+(x)\rangle d\sigma_S(x). \end{aligned}$$

since $L\hat{q}_+ = 0$ in Ω_A^c and $\hat{q}_+ = 0$ on S, and this concludes the proof of the Lemma.

Going back to the proof of Proposition 6, using (53) in (51), we have

$$\Gamma_{S,AB} = \int_S m(x)q_-(x)\langle \hat{n}_S(x), a(x)\nabla\hat{q}_+(x)\rangle d\sigma_S(x).$$

This is true for any piecewise C^1-surface S. Now take $S \equiv S_+(z)$. Then by (33), $\hat{q}_+(x) = (q_+(x) - z)/(1 - z)$, and

$$\Gamma_{S_+(z),AB} = \int_{S_+(z)} \frac{m(x)q_-(x)}{1 - z}\langle \hat{n}_{S_+(z)}(x), a(x)\nabla q_+(x)\rangle d\sigma_z(x). \tag{54}$$

By construction $\Gamma_{S_+(z),AB}$ must be the same for all $\{S_+(z)\}_{z\in(0,1)}$. This can be checked explicitly upon noting that

$$\Gamma_{S_+(z),AB} = B(z)/(1 - z) \tag{55}$$

where (using the co-area formula)

$$B(z) = \int_{S_+(z)} m(x)q_-(x)\langle \hat{n}_{S_+(z)}(x), a(x)\nabla q_+(x)\rangle d\sigma_z(x)$$
$$= \int_{\Omega_{AB}} m(x)q_-(x)\langle \nabla q_+(x), a(x)\nabla q_+(x)\rangle \delta(q_+(x) - z)\, dx \qquad (56)$$

Therefore (using the identity $\nabla q_+(x)\delta'(q_+(x) - z) = \nabla\delta(q_+(x) - z)$, integration by parts and $a : \nabla\nabla q_+ = -b \cdot \nabla q_+$)

$$\begin{aligned} B'(z) &= -\int_{\Omega_{AB}} m(x)q_-(x)\langle \nabla q_+(x), a(x)\nabla q_+(x)\rangle \delta'(q_+(x) - z)\, dx \\ &= -\int_{\Omega_{AB}} m(x)q_-(x)a(x) : \nabla q_+(x)\nabla\delta(q_+(x) - z)\, dx \\ &= \int_{\Omega_{AB}} \nabla \cdot (m(x)q_-(x)a(x)\nabla q_+(x))\delta(q_+(x) - z)\, dx \\ &= \int_{\Omega_{AB}} (\mathrm{div}(a(x)m(x)q_-(x)) \cdot \nabla q_+(x) \\ &\quad + m(x)q_-(x)a(x) : \nabla\nabla q_+(x))\delta(q_+(x) - z)\, dx \qquad (57) \\ &= \int_{\Omega_{AB}} (\mathrm{div}(a(x)m(x)q_-(x)) \cdot \nabla q_+(x) \\ &\quad - m(x)q_-(x)b(x) \cdot \nabla q_+(x))\delta(q_+(x) - z)\, dx \\ &= \int_{S_+(z)} \hat{n}_{S_+(z)} \cdot (\mathrm{div}(a(x)m(x)q_-(x)) - m(x)q_-(x)b(x))\, d\sigma_z(x). \end{aligned}$$

But this means that (denoting by $\Omega_{z,z'}$ the region between $S_+(z)$ and $S_+(z')$ and using Gauss' theorem)

$$\begin{aligned} B'(z) &= \int_{S_+(z')} \hat{n}_{S_+(z)} \cdot (\mathrm{div}(a(x)m(x)q_-(x)) - m(x)q_-(x)b(x))\, d\sigma_{z'}(x) \\ &\quad + \int_{\Omega_{z,z'}} \nabla \cdot (\mathrm{div}(a(x)m(x)q_-(x)) - m(x)q_-(x)b(x))\, dx \\ &= B'(z'), \end{aligned}$$

where we used the divergence free property in (40). It follows that $B'(z) = cst$, and since $B(1) = 0$ from (56), we must have $B(z) = C(1 - z)$ where the constant C must be $\Gamma_{S_+(z),AB}$ by (55), i.e. we have the identity

$$\Gamma_{S_+(z),AB} = -B'(z).$$

From the last equality in (57) it follows that

$$\Gamma_{S_+(z),AB} = \int_{S_+(z)} \hat{n}_{S_+(z)} \cdot (m(x)q_-(x)b(x) - \mathrm{div}(a(x)m(x)q_-(x)))\, d\sigma_z(x)$$
$$\equiv \int_{S_+(z)} \hat{n}_{S_+(z)} \cdot J_{AB}(x) d\sigma_z(x).$$

Since J_{AB} is divergence free, this expression is in fact valid for any dividing piecewise C^1 surface S and not only $S_+(z)$ by Gauss' theorem, which terminates the proof. □

ACKNOWLEDGMENTS

This work is part of a joint project with Weiqing Ren: we thank him for stimulating discussions. We also thank G. Ben Arous, G. Ciccotti, P. Constantin, A. Fischer, P. Metzner, and C. Schütte for helpful suggestions. The work of E is partially supported by ONR grant N00014-01-1-0674. The work of E. V.-E. is partially supported by NSF grants DMS02-09959 and DMS02-39625, and by ONR grant N00014-04-1-0565.

REFERENCES

1. A. Bovier, M. Eckhoff, V. Gayrard and M. Klein, Metastability in reversible diffusion processes I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc.* **6**:1–26 (2004).
2. A. Bovier, V. Gayrard and M. Klein, Metastability in reversible diffusion processes. II. Precise estimates for small eigenvalues. *J. Eur. Math. Soc.* **7**:69–99 (2005).
3. M. V. Day, Mathematical approaches to the problem of noise-induced exit, pp. 269–287. In: *Stochastic analysis, control, optimization and application*, Systems Control Found. Appl., Birkhõuser Boston, Boston, MA, 1999.
4. C. Dellago, P. G. Bolhuis and P. L. Geissler, Transition path sampling. *Adv. Chem. Phys.* **123** (2002)
5. R. Durrett. *Stochastic Calculus.* CRC Press, 1996.
6. W. E, W. Ren and E. Vanden-Eijnden, Finite temperature string method for the study of rare events. *J. Phys. Chem. B* **109**, 6688–6693 (2005).
7. W. E, W. Ren and E. Vanden-Eijnden, Transition pathways in complex systems: Reaction coordinates, isocommittor surfaces, and transition tubes. *Chem. Phys. Lett.* **413**: 242–247 (2005).
8. M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, 2nd ed. Springer, 1998.
9. P. Hänggi, P. Talkner and M. Borkovec, Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **62**: 251–342 (1990).
10. W. Huisinga, S. Meyn and Ch. Schütte, Phase Transitions and Metastability in Markovian and Molecular Systems. *Ann. Appl. Probab.* **14**: 419–458 (2004).
11. G. Hummer, From transition paths to transition states and rate coefficients. *J. Chem. Phys.* **120**: 516–523 (2004).
12. A. Ma and A. R. Dinner, Automatic Method for Identifying Reaction Coordinates in Complex Systems. *J. Phys. Chem. B* **109**: 6769–6779 (2005).

13. R. S. Maier and D. L. Stein, Limiting exit location distributions in the stochastic exit problem. *SIAM J. Appl. Math.* **57**: 752–790 (1997).
14. B. J. Matkowsky and Z. Schuss, The Exit Problem for Randomly Perturbed Dynamical Systems. *SIAM J. App. Math.* **33**: 365–382 (1977).
15. W. Ren, E. Vanden-Eijnden, P. Maragakis and W. E, Transition Pathways in Complex Systems: Application of the Finite-Temperature String Method to the Alanine Dipeptide. *J. Chem. Phys.* **123**: 134109 (2005)
16. Z. Schuss, Singular Perturbation Methods on Stochastic Differential Equations of Mathematical Physics. *SIAM Review* **22**: 119–155 (1980).
17. Z. Schuss and B. J. Matkowsky, The Exit Problem: A New Approach to Diffusion Across Potential Barriers. *SIAM J. App. Math.* **36**: 604–623 (1979).
18. 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).
19. A.-S. Sznitman, *Brownian Motion, Obstacles and Random Media.* Springer, 1998.
20. F. Tal, E. Vanden-Eijnden, Transition state theory and dynamical corrections in ergodic systems. *Nonlinearity* **19**: 501–509 (2006).
21. E. Vanden-Eijnden, F. Tal, Transition state theory: Variational formulation, dynamical corrections, and error estimates. *J. Chem. Phys.* **123**: 184103 (2005).