Stochastic actions for diffusive dynamics: Reweighting, sampling, and minimization

Artur B. Adib

Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-0520, USA

(Dated: February 3, 2008)

In numerical studies of diffusive dynamics, two different action functionals are often used to specify the probability distribution of trajectories, one of which requiring the evaluation of the second derivative of the potential in addition to the force. Here it is argued that both actions are equivalent prescriptions for the purposes of reweighting and sampling trajectories, whereas the most probable path is more generally given by the global minimum of the action involving the second derivative term. The answer to this apparent paradox lies in the non-differentiable character of Brownian paths, as well as in the “entropy” associated with a given trajectory.

Keywords: Onsager-Machlup functional, Jacobian term, Instanton

I. INTRODUCTION

The present paper is concerned with the probability distribution of diffusive trajectories, particularly in discrete form as suited for numerical computations. In its simplest version, the problem is that of specifying the probability that a Brownian particle moving according to the overdamped Langevin equation

$$\dot{x}(t) = \frac{F(x(t))}{\zeta} + \sqrt{2D} \mathcal{R}(t)$$

will follow a particular path from an initial to a final position in a given time. Here $F(x) = -U'(x)$ is the force acting on the particle, $\zeta$ is the friction coefficient, $D$ is the diffusion constant, and $\mathcal{R}(t)$ is a random force term satisfying the white noise relation $\langle \mathcal{R}(t) \mathcal{R}(t') \rangle = \delta(t-t')$. The temperature in energy units, $kT$, is related to the above parameters via the Einstein relation, $D = kT/\zeta$.

Unlike their deterministic counterparts, stochastic differential equations such as Eq. (1) remain ambiguous until one specifies the discretization rule for evaluating the various time-dependent quantities. Two typical rules are the Ito and Stratonovich discretizations, in which $F(x(t))$ is evaluated at the beginning and at the mid-point of the time-interval, respectively (see Eq. (15), for example). In the continuum limit, it is well known that both choices lead to (a) the same Fokker-Planck equation (provided $D$ is independent of $x$; cf. Eq. (15)), and (b) the same path integral formula for the propagator, expressed in terms of the so-called action functional (cf. Eqs. (2)-(3)).

The above results are consistent with the intuitive expectation that the measurable statistical properties of a Brownian particle should be insensitive to the particular stochastic calculus chosen to describe its motion. Nonetheless, an ambiguity of practical character arises when deriving a discrete form for the probability distribution of trajectories, i.e. the discretized action: Depending on the type and the stage at which one carries out the discretization, the ensuing action can take on different functional forms (cf. Sec. III). In particular, in previous numerical studies concerned with sampling and optimizing diffusive trajectories, one typically finds two different expressions for the stochastic action, and no consensus seems to exist as to which form should be used in a given application.

It is the purpose of this paper to shed some light on this issue. Of course, the scope of this problem goes well beyond the prototypical case of a Brownian particle moving in a field of force; for example, chemical reactions in solution are typically described by effective equations of motion that are straightforward multi-dimensional extensions of Eq. (1). A specific example of increasing interest in recent years is the microscopic description of the mechanisms underlying protein folding, which often requires the characterization of diffusive paths connecting unfolded and folded states. In these and similar cases, it is crucial to understand and resolve the aforementioned ambiguity. The next section offers a brief introduction to this problem, followed by specific applications to reweighting, sampling, and minimization in Sec. III.

II. TWO ACTIONS

The central quantity of this work is the relative probability of observing a given diffusive trajectory connecting an initial $(x_i)$ to a final $(x_f)$ configuration. In the continuum limit, this probability is directly available from the path integral expression for the propagator, $P(x_f|x_i; t)$.

For the dynamics in Eq. (1), regardless of the discretization adopted, one obtains

$$P(x_f|x_i; t) = \int_{x(0)=x_i}^{x(t)=x_f} Dx(s) e^{S[x(s)]/2D},$$

where

$$S[x(s)] = \frac{\Delta U}{\zeta} + \frac{1}{2} \int_0^t ds \left[ \dot{x}^2 + \left( \frac{F}{\zeta} \right)^2 + \frac{2D}{\zeta} F^2 \right]$$

is the so-called action functional. Here $\Delta U \equiv U(x_f) - U(x_i)$, and the term involving the formal time derivative of $x$, viz. $\int ds \dot{x}^2$, is a shorthand for the free diffusion contribution to the path-integration measure, cf. Eq. (7) (recall that Brownian trajectories are non-differentiable).

Equation (2) shows that each trajectory contributes to...
the propagator with a weight ultimately dictated by its action. This will be the object of interest in the remainder of the paper.

**A. Derivation**

Although in the continuum limit the action is given unambiguously by the above expression, different discrete forms can be obtained depending on how and when one discretizes the problem. Here I will follow the (implicit or explicit) argument that often appears in the literature to justify the use of two particular discrete actions. The first arises by performing the discretization at the level of Eq. (1) using the Ito rule (which incidentally leads to the algorithm traditionally adopted for the numerical generation of Brownian trajectories), namely

$$x_{n+1} = x_n + \frac{\Delta t}{\zeta} F_n + \sqrt{2D\Delta t} R_n,$$

(4)

where $\Delta t$ is the time step, $F_n \equiv F(x_n)$, and $R_n$ is a Gaussian random variable of unit variance. Because the distribution $P(R_n) = e^{-R_n^2/2}/\sqrt{2\pi}$ is known and the force only depends on the previous position $x_n$, the short-time conditional probability $P(x_{n+1}|x_n; \Delta t)$ can be straightforwardly obtained by changing variables from $R_n$ to $x_{n+1}$, which in this particular case comes with unit Jacobian. The corresponding action, which comes from the definition

$$P(x_N|x_{N-1}; \Delta t) \cdots P(x_1|x_0; \Delta t) \equiv e^{-S_1[x_0 \cdots x_N]/2D}$$

(5)

up to a path-independent factor, is thus given by

$$S_1[x_0 \cdots x_N] = \frac{\Delta t}{2} \sum_{n=0}^{N-1} \left( \frac{x_{n+1} - x_n}{\Delta t} - \frac{F_n}{\zeta} \right)^2,$$

(6)

where $x_0 = x_i$ and $x_N = x_f$. Note that this dictates the exact probability of a path generated by Eq. (4) for any finite $\Delta t$. On the other hand, a direct discretization at the level of Eq. (3) gives the action

$$S_2[x_0 \cdots x_N] = \frac{\Delta U}{\zeta} + \frac{\Delta t}{2} \sum_{n=0}^{N-1} \left( \frac{x_{n+1} - x_n}{\Delta t} \right)^2 + \left( \frac{F_n}{\zeta} \right)^2 + \frac{2D}{\zeta} F_n^2.$$

(7)

Aside from the formal $x^2$ term correctly expressed above, Eq. (6) involves ordinary integrals only, and hence is insensitive to the particular discretization rule adopted in Eq. (7).

**B. When are the actions equivalent?**

The two actions derived above are manifestly different; in particular, the presence of the derivative of the force makes $S_2$ especially inconvenient for numerical implementations, which presumably explains why $S_1$ is more widely adopted in the literature (see also Ref. [18] for a new action functional without this term). Nonetheless, the very validity of $S_1$ as the action for diffusive problems has been questioned by some authors.[21,22] The goal of this section is to clarify the (non-)equivalence between these two actions.

To begin the analysis, it is instructive to first consider an obvious situation where the actions are not equivalent, namely, when they are evaluated in the continuum limit over differentiable trajectories. In this case, Eq. (6) becomes

$$S_1[x(s)] = \frac{\Delta U}{\zeta} + \frac{1}{2} \int_0^t ds \left[ \dot{x}^2 + \left( \frac{F}{\zeta} \right)^2 \right],$$

(8)

where the cross-term $\dot{x}F/\zeta$ was integrated by parts to yield the surface term $\Delta U/\zeta$. Comparing this result with the analogous limit of $S_2$, which is repeated here for the convenience of the reader (cf. Eq. (3)),

$$S_2[x(s)] = \frac{\Delta U}{\zeta} + \frac{1}{2} \int_0^t ds \left[ \dot{x}^2 + \left( \frac{F}{\zeta} \right)^2 + \frac{2D}{\zeta} F' \right],$$

(9)

we see that these two expressions differ by a non-trivial term whose magnitude is dictated by the ratio $D/\zeta$; i.e., the actions are generally not equivalent when evaluated over differentiable functions (the case where $F'$ is constant is an exception, e.g. the harmonic potential case considered by Onsager and Machlup[22].)

Consider now the case where the actions are evaluated over Brownian paths, which are non-differentiable and probabilistic in nature. The continuum limit of Eq. (9) then leads to a stochastic integral, which in general does not satisfy the usual relations from differentiable calculus. Indeed, in the limit of infinitely many intervals, the cross term that gave rise to the surface term above becomes an Ito integral, i.e.

$$\lim_{\Delta t \to 0} \sum_{n=0}^{N-1} (x_{n+1} - x_n) F_n \equiv \int_{x_i}^{x_f} dx F(x),$$

(10)

which satisfies the (probabilistic) equality (see also Appendix)

$$\int_{x_i}^{x_f} dx f'(x) = \Delta f - D \int_0^t ds f''(x(s)).$$

(11)

Here $D$ is the diffusion constant associated with the trajectory $x(t)$, $f(x)$ is an arbitrary differentiable function, and $\Delta f = f(x_f) - f(x_i)$. Note that in the zero-noise limit ($D \to 0$) one recovers the fundamental theorem of differentiable calculus; for any finite $D$, however, this result allows one to replace Ito integrals by ordinary time-integrals. Consequently, when expressed in terms of ordinary time-integrals only, the limiting form of $S_1$ acquires.
a term proportional to the second-derivative of the potential in addition to the terms shown in Eq. (5), correctly recovering Eq. (9). This shows the (probabilistic) equivalence between $S_1$ and $S_2$ in the continuum limit when evaluated over Brownian trajectories with same diffusion constant as that appearing in the definition of $S_2$.

These observations are summarized in Figures 1 and 2.

### III. APPLICATIONS

The fact that the relative probability of observing a given diffusive path is given by a Boltzmann-like factor (cf. Eq. 2) shows that the action and the diffusion constant are for trajectories what the energy and the temperature are for conformations in the canonical ensemble. This analogy has led many authors to translate equilibrium methods (e.g., sampling, reweighting and minimization techniques), to the context of diffusive trajectories. In this section, I will argue that the ensemble of trajectories thus generated (or reweighted) is insensitive to the choice between Eq. 6 or Eq. 7, while most probable trajectories are more accurately obtained by minimizing $S_2$ instead of $S_1$.

#### A. Reweighting

If one is given an ensemble of diffusive trajectories generated under some conditions (e.g. for some potential $U(x)$ and diffusion constant $D$), it is in principle possible to “reweight” them so as to obtain a new ensemble corresponding to different conditions (e.g. a different potential $U^\ast(x)$, with same diffusion constant). In this example, the reweighting factor is given by

$$P^\ast\{\{x_i\}|x_0\} \over P\{\{x_i\}|x_0\} = e^{-\Delta S/2D},$$

where $\Delta S = S^\ast[x_0 \cdots x_N] - S[x_0 \cdots x_N]$ is the action difference due to the new potential $U^\ast(x)$, and $D$ is the diffusion constant associated with the given ensemble of trajectories. The notation $P\{\{x_i\}|x_0\}$ stands for the probability of observing a trajectory $\{x_i\} \equiv x_1 \cdots x_N$ given the initial point $x_0$. The results of the previous section immediately indicate that the choice between $S_1$ or $S_2$ for $S$ is immaterial for this expression: Since the actions are equal (in a probabilistic sense) when evaluated over this ensemble, the reweighting factor is effectively the same whether one uses the former or the latter action.

Figure 3 presents a simple example of this observation, where free trajectories ($U(x) = 0$) were reweighted so as to give an ensemble corresponding to a particle subject to a quartic potential $U^\ast(x) = x^4/4$ and the parameters $D = 1$, $\zeta = 1$.

\[\langle x^4(t)\rangle_U = \langle e^{-\Delta S/2D} x^4(t)\rangle_U,\]
FIG. 3: The time-dependent moment \( \langle x^4(t) \rangle \) for a Brownian particle starting at the origin and moving on the quartic potential \( U^*(x) = x^4/4 \). The shaded gray curve is the result of a direct simulation using Eq. (1). The curves with smaller fluctuations about \( \langle x^4(t) \rangle = 1 \) are the results obtained by reweighting \( 10^7 \) free \( (U(x) = 0) \) Brownian trajectories according to Eq. (13), using both \( S_1 \) (black solid curve) and \( S_2 \) (red dashed curve). The remaining curves correspond to a one-ended boundary condition version of the “transition path sampling” (TPS) Monte Carlo algorithm with relative probability given by Eq. (14) and action given by \( S_1 \) (blue solid curve) and \( S_2 \) (green dashed curve). The total number of trajectories used for the TPS algorithm was \( 5 \times 10^6 \). (The noise level for the TPS results is high even for such a large number of samples because the present Monte Carlo algorithm has not been optimized for efficient trajectory sampling). The parameters adopted throughout are \( \Delta t = 0.01, \zeta = 1 \), and \( D = 1 \). For these values of \( D \) and \( \zeta \), the exact asymptotic result (e.g. via the virial theorem) is unity.

where the specific functional form of \( \Delta S \) depends on the action adopted, and the particle always starts at the origin for both averages.

B. Sampling

The case of sampling is subtler. Since now the actions generate the ensemble of trajectories itself, one can no longer straightforwardly rely on the above argument, which assumed the ensemble of trajectories was given. Here one is sampling the entire space of trajectories with weight dictated by the action, i.e.

\[
P^i[x_i|x_0] \propto e^{-\mathcal{S}[x_0\cdots x_N]/2D},
\]

where \( D \) is the desired diffusion constant. The question remains as to whether the ensemble thus generated is sensitive to the choice between \( S_1 \) and \( S_2 \).

A short (but not particularly illuminating) answer can be given as follows. Since the variable \( x \) is Markovian (cf. Eq. (1)), the statistical properties of its trajectories are fully specified by its propagator, \( P(x|x_0;t) \). The problem of equivalence of trajectory ensembles thus reduces to that of equivalence of propagators. When one adopts \( S_1 \) as the action, one is sampling exactly the same ensemble of trajectories as that generated by Eq. (1). In the continuum limit, it is known that the propagator corresponding to this Ito-like equation of motion solves the Fokker-Planck equation

\[
\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x} \left( \frac{F}{\zeta} P \right) + D \frac{\partial^2 P}{\partial x^2}.
\]

But this is precisely the same Fokker-Planck equation solved by the path integral in Eqs. (2)-(3), which defines the propagator of the ensemble of trajectories generated by \( S_2 \) in the continuum limit; i.e. in this limit, the propagators corresponding to the ensembles of trajectories generated by \( S_1 \) and \( S_2 \) are identical. Thus, by the Markovian argument above, the two ensembles are statistically equivalent.

A more constructive proof of equivalence can also be given by recalling that, effectively, \( S_1 \) and \( S_2 \) only differ when the trajectories over which they are evaluated have a diffusion constant other than that entering the definition of \( S_2 \). Thus, what one needs to show is that the probability of sampling a path with diffusion constant different from the one specified in Eq. (14) is vanishingly small, whether one uses \( S_1 \) or \( S_2 \); for the remaining trajectories whose sampling probability is in principle non-zero, the actions are effectively equal and hence lead to the same relative weight between the trajectories.

To proceed, notice that both actions contain a Wiener (free diffusion) term, and thus – irrespective of the action chosen – Eq. (14) can be factored as

\[
P^i[x_i|x_0] \propto P_{\text{free}}^{(D)}[x_i|x_0] \cdot P_{\text{int}}[x_i|x_0],
\]

where

\[
P_{\text{free}}^{(D)}[x_i|x_0] \propto e^{-\sum_{n=0}^{N-1} \frac{(x_{n+1} - x_n)^2}{2D}},
\]

and \( P_{\text{int}} \) is an action-specific distribution accounting for the interaction of the particle with the potential \( U(x) \). The proof is finished by showing that the presence of the modulating factor \( P_{\text{free}}^{(D)} \) leads to a sharply peaked (i.e. zero variance) distribution of diffusion constants about \( D \) (cf. Appendix), which effectively rules out the sampling of trajectories that violate the equivalence between \( S_1 \) and \( S_2 \).

The above observations are illustrated in Figure 3 again for a particle moving in the quartic potential \( U(x) = x^4/4 \). The sampling method adopted is a simple extension of the transition path sampling algorithm, where only the initial point \( x(0) = 0 \) is fixed. The method uses a straightforward Monte Carlo procedure, where new trajectories are attempted by displacing one of the various \( x_i \) by a given amount, and accepting or rejecting
the new trajectory according to the standard Metropolis rule\cite{8} (recall the analogy between action-diffusivity and energy-temperature above).

C. Minimization and most probable trajectories

What has been shown so far is that in trajectory space, the surfaces defined by $S_1$ and $S_2$ are nearly coincident over the overwhelming majority of trajectories representative of the desired ensemble. Nonetheless, appreciable discrepancies between these two surfaces arise when one considers trajectories with “wrong” diffusivity; in particular, the actions generally lead to entirely different values when evaluated over differentiable functions (Sec. II B). As we shall see below, this bears important consequences to the study of most probable trajectories.

Since the probability of observing a given trajectory decreases exponentially fast with its action (cf. Eq. (14)), the trajectories with greatest statistical weight are those for which the action is minimal. For the moment, this will be our operational definition of “most probable trajectory.” Consider therefore the problem of minimizing $S_1$ and $S_2$ in the space of all possible trajectories $x_0 \cdots x_N$. It is easy to see that the minima of both actions correspond to differentiable functions in the continuum limit. Indeed, because of the free diffusion contribution common to both actions, namely

$$S_{\text{free}}[x_0 \cdots x_N] = \frac{1}{2} \sum_{n=0}^{N-1} \frac{\Delta x_n^2}{\Delta t},$$

where $\Delta x_n = x_{n+1} - x_n$, the actions diverge in the continuum limit when evaluated over diffusive trajectories, for which $\Delta x_n^2 \sim O(\Delta t)$. Differentiable functions, on the other hand, yield a finite $S_{\text{free}}$ contribution, as in this case $\Delta x_n^2 = O(\Delta t^2)$ (i.e. $\dot{x}$ is finite). Thus, we arrive at the following situation of conflict concerning most probable trajectories: Not only are they highly atypical of diffusive processes (insofar as they are differentiable), but also their very identity is sensitive to the choice between $S_1$ and $S_2$ (as these actions are generally different for differentiable paths). The story of minimization is thus qualitatively different from the case of reweighting and sampling.

The answer to the first conflict is relatively easily settled, and as we shall see hints at the solution of the second. The phenomenon associated with it is not unlike that observed in equilibrium sampling studies: Although the most likely conformations of a given system (e.g. atomic clusters, proteins, liquids, etc.) in principle correspond to minima in the energy landscape, in practice thermal fluctuations will keep the system near to, but hardly in the precise location of the minima themselves. Thus, although differentiable paths are virtually never observed in diffusive problems, they can provide the motif upon which fluctuations – i.e. diffusive trajectories – emerge. This in turn reveals the importance of the neighborhood of the minima in such analyses, which was neglected in the above discussion of most probable trajectories.

An illustration of how the nature of such neighborhoods can affect the definition of most probable trajectories is given in Fig. 4 for a simple one-dimensional potential. By analogy, we are thus led to look at relative probabilities of the type

$$P([x_i]_0) \frac{\delta x'_1 \cdots \delta x'_N}{P([x_i]_0) \delta x_1 \cdots \delta x_N} = e^{-|\Delta S - 2D\Delta \Sigma|/2D},$$

where $\Delta S = S[x_0' \cdots x_N'] - S[x_0 \cdots x_N]$ and similarly for the “trajectory entropy” difference $\Delta \Sigma$, where $\Sigma[x_1 \cdots x_N] = \ln \delta x_1 \cdots \delta x_N$. Of course, the definition of $\Sigma$ depends on the definition of the measure $\delta x_1 \cdots \delta x_N$, which is not unique; essentially, it depends on how far and in what ways one is willing to accept points away from a given motif trajectory as mere fluctuations about it. Perhaps the simplest definition is obtained by looking at the fluctuations of each individual $x_n$ for a given typical action fluctuation $\delta S$, i.e. by inverting the expansion

$$\delta S = \frac{\partial S}{\partial x_n} \delta x_n + \frac{1}{2} \frac{\partial^2 S}{\partial x_n^2} \delta x_n^2 + \ldots.$$  

Assuming one is dealing with local minima or points of very low gradients in comparison to curvatures, a second-order truncation yields the simple measure

$$\delta x_1 \cdots \delta x_N = \sqrt{2\delta S / S''(x_1)} \cdots \sqrt{2\delta S / S''(x_N)}.$$

FIG. 4: Simple illustration of the entropic role played by the curvature near conformational minima. For the same amount of thermal energy, say $\delta U$, the typical amplitude $\delta x$ of conformational fluctuations near $x$ is much smaller than that near $x'$, $\delta x'$ (use e.g. $\delta U \approx \frac{1}{2} U''(x)\delta x^2$). Thus, although conformations in the neighborhood of $x$ are energetically favored over those near $x'$, the opposite is true when one considers the entropy due to the different curvatures, and the final relative population of $x$- and $x'$-“like” conformations should embody both contributions, e.g. $P(x')/P(x) = |P(x')/P(x)| \times [U''(x')/U''(x)]^{-1/2}$ rather than simply $P(x')/P(x)$. 


where $S''(x_n)$ is a shorthand for $\partial^2 S[x_0 \cdots x_N]/\partial x_n^2$. I will now proceed to show that with this simple choice of measure in Eq. (19), the second conflict above is eliminated.

According to the new relative probability in Eq. (19), in order to find most likely trajectories the quantity to be minimized is not the action alone, but rather the “trajectory free energy”

$$F[x_0 \cdots x_N] \equiv S[x_0 \cdots x_N] - 2D \Sigma[x_0 \cdots x_N].$$

Because the actions diverge for diffusive trajectories, any minimization process will quickly be restricted to what would be differentiable trajectories in the continuum limit. Let us therefore evaluate the action derivatives required by $\Sigma$ (cf. Eqs. (20) and (22)) over such trajectories. From the expressions in Eqs. (6) and (7), up to terms explicitly of first order in $\Delta t$ one gets

$$\frac{\Delta t}{2} \frac{\partial S_1}{\partial x_n} = -(x_{n+1} - 2x_n + x_{n-1})$$

$$\frac{\Delta t}{4} \frac{\partial^2 S_1}{\partial x_n^2} = 1 - \frac{\Delta t}{\zeta} [(x_{n+1} - x_n)F'_n - F_n + F_{n-1}]$$

and

$$\frac{\Delta t}{2} \frac{\partial S_2}{\partial x_n} = -(x_{n+1} - 2x_n + x_{n-1})$$

$$\frac{\Delta t}{4} \frac{\partial^2 S_2}{\partial x_n^2} = 1.$$  

A couple of observations are immediately apparent from these expressions. First, the gradients of both $S_1$ and $S_2$ are of order $O(\Delta t)$ for differentiable trajectories [use $x_{n+1} - 2x_n + x_{n-1} = O(\Delta t^2)$ and $(x_{n+1} - x_n)F'_n - F_n + F_{n-1} = O(\Delta t^2)$], whereas their curvatures diverge as $O(1/\Delta t)$. Second, to order $\Delta t$, the relative curvatures of $S_1$ are trajectory-dependent, whereas in the case of $S_2$ such curvatures are constant.

Thanks to the first observation, the definition in Eq. (22) is valid, and we obtain the trajectory entropy corresponding to $S_1$ by invoking the result in Eq. (25)

$$\Sigma_1[x_0 \cdots x_N] = -\frac{1}{2} \ln \prod_{n=1}^{N} \left[ 1 + \frac{\Delta t}{\zeta} F'_n + O(\Delta t^2) \right]$$

$$= -\frac{\Delta t}{2\zeta} \sum_{n=1}^{N} F'_n + O(N \Delta t^2)$$

where use was made of $(x_{n+1} - x_n)F'_n = O(\Delta t)$, and trajectory-independent constants are omitted. An analogous computation of the trajectory-dependent components of $S_2$ reveals that

$$\Sigma_2[x_1 \cdots x_N] = O(N \Delta t^2),$$

which vanishes in the continuum limit; i.e., only $S_1$ leads to a trajectory-dependent entropy. Therefore, when evaluated over differentiable trajectories in the continuum limit, the quantity to be minimized in the case of $S_1$ is the “free energy”

$$F_1[x(s)] = S_1[x(s)] - \frac{D}{\zeta} \int_0^t ds F',$$

whereas in the case of $S_2$ the free energy is simply the action itself, i.e.

$$F_2[x(s)] = S_2[x(s)].$$

Comparing these results with Eqs. (8) and (9), we see that when fluctuations are taken into account the quantities to be minimized coincide (i.e. $F_1 = F_2$), and consequently the aforementioned ambiguity disappears. The present treatment of most probable trajectories then leads to a functional that agrees with more elaborate mathematical approaches using induced measures in function space.

To illustrate the above observations, consider the one-dimensional problem of a Brownian particle moving according to Eq. (1) in the double-well potential

$$U(x) = 3(x^2 - 1)^2,$$

and let us ask what is the most probable trajectory motif that takes the particle from the minimum at $x = -1$ to the minimum at $x = +1$, in one unit of time for $D = \zeta = 1$. Figure 5 shows the predictions obtained by direct minimization of the actions alone. In agreement with our expectation, the actions predict smooth – i.e. “differentiable” – paths connecting the end-points, with
FIG. 6: Distribution of barrier-crossing times in the barrier region \(-0.3 < x < 0.3\) of the double-well potential of Fig. 5 according to the transition path sampling algorithm (TPS) using both \(S_1\) (black solid curve) and \(S_2\) (red dashed curve) for a total trajectory length of \(t = 1\) with \(\Delta t = 0.01\) (longer trajectory lengths do not change the shape of the instanton, and hence yield essentially the same barrier-crossing times per instanton). The results from a direct Langevin simulation using Eq. (4) is also shown (shaded gray curve). In accordance with the discussion of Sec. III B, the statistics is insensitive to the choice of action. Nonetheless, the predictions based on the minima of the actions alone (vertical blue bars) are quantitatively different (cf. Fig. 5). The result based on \(S_2\), whose entropic correction is trajectory-independent, is in good agreement with the observed maximum in the distribution. For comparison, the Kramers time \(\tau_K \approx 4\) (obtained from the exponential decay of the correlation function \(\langle x(0)x(t) \rangle \sim e^{-t/\tau_K}\)) was also calculated, and falls well outside the range of the plot.

the specific shape depending on whether one uses \(S_1\) or \(S_2\). In particular, the predicted paths spend rather different times in the barrier region, herein called the “barrier-crossing time” (note that this quantity is generally different from the transition, or Kramers’ time, which is largely dominated by the residence time in each well). In agreement with our discussion above, however, \(S_2\) already embodies the trajectory entropies, and hence its minimum predicts more accurately the overall shape of the most probable trajectories, in particular their barrier-crossing times (cf. Fig. 6).

IV. CONCLUSIONS

The present paper was primarily motivated by a conspicuous ambiguity in previous numerical studies of diffusive processes using path integral techniques. This ambiguity was introduced in Sec. III where it was suggested that at least two different actions, namely Eqs. (6) and (7), are in principle legitimate prescriptions for the discrete representation of the propagator, Eq. (2). It was then argued that the actions are indeed equivalent when evaluated over diffusive, non-differentiable trajectories with the same diffusion constant as that appearing in Eq. (7), while generally leading to different values when evaluated over differentiable paths (Sec. III B).

The consequences of these observations to numerical simulations were then explored in Sec. III where the equivalence of the actions for the purposes of reweighting and sampling diffusive paths was shown and numerically illustrated, while for problems concerned with the most likely “motif” of diffusive trajectories it was argued that \(S_2\) already contains the relevant entropic corrections connected with fluctuations about the trajectory motif, and hence should be favored over \(S_1\) in direct minimization studies.

A comforting message in this work is that, contrary to what has been suggested previously \(^{13}\), the second-derivative of the potential energy function (cf. Eq. (7)) is not always required for the correct description of diffusive problems (see also Ref. \(^{13}\)); it is only if one inquires about most probable trajectories that this numerically inconvenient term plays a role. It is expected that this and the other observations above will aid future authors in making judicious choices for the discrete action in diffusive problems.

Acknowledgments

The author would like to thank Attila Szabo and Gerhard Hummer for numerous discussions, and Gavin Crooks for pertinent suggestions. This research was supported by the Intramural Research Program of the NIH, NIDDK.

APPENDIX

This Appendix contains a derivation that illustrates two results, namely the validity of the Itô formula (Eq. (11)) for a particular choice of \(f(x)\), and the sharply peaked structure of the distribution of diffusivities about \(D\) when sampling from \(F_{\text{free}}\).

For the particular choice \(f(x) = x^2/2\) in Eq. (11), we need to show that, in a mean-square sense,

\[
\sum_{n=0}^{N-1} (x_{n+1} - x_n) (x_{n+1} - x_n) - \frac{x_N^2 - x_0^2}{2} + Dt = 0 \quad (A.33)
\]

as \(N \to \infty, \Delta t \to 0\), with \(t = N\Delta t\) constant. The proof offered below follows closely that of Gardiner\(^2\) and proceeds by showing that both the first and the second moments of the left-hand-side of Eq. (A.33) are zero for a free Brownian trajectory \(\{x_t\}\) (an extension to more
general cases, e.g. trajectories satisfying Eq. (4), is easily performed by keeping track of the leading order terms in $\Delta t$ only.

The trick is to re-express Eq. (A.33) in terms of increments $\Delta x_n = x_{n+1} - x_n$, which for free diffusion are exactly Gaussian-distributed. A simple algebraic manipulation of the sum in Eq. (A.33) gives

$$
\sum_{n=0}^{N-1} (x_{n+1} - x_n)x_n = \frac{1}{2} \sum_{n=0}^{N-1} [x_n^2 - x_n^2 - \Delta x_n^2]
$$

$$
= \frac{x_N^2 - x_0^2}{2} - \frac{1}{2} \sum_{n=0}^{N-1} \Delta x_n^2,
$$

(A.34)

and therefore we need to show that, in the aforementioned limit,

$$
M_p = \langle (\sum_n \Delta x_n^2 - 2Dt)^p \rangle_{x_0} \to 0
$$

(A.35)

for $p = 1, 2$, where the average subscript means that the initial point $x_0$ is fixed. The first moment $M_1$ is immediately verified, as $\langle \Delta x_n^2 \rangle_{x_0} = 2D\Delta t$. The second moment gives

$$
M_2 = \langle \sum_n \Delta x_n^4 \rangle_{x_0} + \langle \sum_{n\neq m} \Delta x_n^2 \Delta x_m^2 \rangle_{x_0} - (2D)^2,
$$

(A.36)

where the cross-term in the quadratic expansion effectively changed the sign of the last term. Now, for Gaussian variables, $\langle \Delta x_n^4 \rangle_{x_0} = 3 \langle \Delta x_n^2 \rangle_{x_0}^2$, and since $\Delta x_n$ is uncorrelated with $\Delta x_m$ for $n \neq m$, we get

$$
M_2 = 3N(2D\Delta t)^2 + N(N-1)(2D\Delta t)^2 - (2D)^2
$$

$$
= 2(2D)^2 t \Delta t,
$$

which vanishes in the continuum limit, thus proving Eq. (A.35).

As to the second assertion concerning the role of $P_{\text{free}}^{(D)}$ in fixing the diffusion constant, if one takes as an effective measure of diffusivity of a trajectory the quantity

$$
D_{\text{eff}}[x_0 \cdots x_N] \equiv \frac{\sum_{n=0}^{N-1} \Delta x_n^2}{2t},
$$

(A.37)

one can write Eq. (A.35) as

$$
M_p = \langle (D_{\text{eff}}[x_0 \cdots x_N] - D)^p \rangle_{x_0} \to 0
$$

(A.38)

for $p = 1, 2$, where the continuum limit was defined above. By itself, this result already shows the sharply peaked structure of the distribution of effective diffusivities about $D$. To see where this structure comes from, consider the case of free diffusion. Since in this case the angle-brackets are nothing but averages with respect to the density $P_{\text{free}}^{(D)}$, and since $P_{\text{free}}^{(D)}$ depends on the trajectory through $D_{\text{eff}}$, namely

$$
P_{\text{free}}^{(D)}[\{x_i\}|x_0] \propto e^{-N D_{\text{eff}}[x_0 \cdots x_N]/2D},
$$

(A.39)

in analogy with canonical averages one can write

$$
\langle f \rangle_{x_0} = \int dD' \Omega_x(D') P_{\text{free}}^{(D')}(D') f(D'),
$$

(A.40)

where $f$ is an arbitrary function of $D_{\text{eff}}[x_0 \cdots x_N]$, and

$$
\Omega_x(D') = \int d\{x_i\} \delta(D' - D_{\text{eff}}[x_0 \cdots x_N])
$$

(A.41)

is the “density of trajectories” with effective diffusion constant $D'$. With this relation in mind, Eq. (A.38) shows that the probability density in effective diffusivity space, viz. $\Omega(D') \times P_{\text{free}}^{(D')}(D')$, is sharply peaked about $D$, as desired. A corollary of this result is that $\Omega$ is a fast-growing function of the effective diffusion constant (indeed, an explicit computation shows that the result is analogous to the microcanonical density of states of an ideal gas, i.e. $\Omega(D) \sim D^{N/2}$ for large $N$).

From a practical standpoint, the above observation is more general than the well known result that the set of differentiable functions ($D_{\text{eff}} = 0$) has measure zero with respect to $P_{\text{free}}^{(D)}$, as it establishes that the set of all functions with “wrong” diffusion constant (i.e. $D_{\text{eff}} \neq D$) is also practically precluded from any sampling modulated by $P_{\text{free}}^{(D)}$.  

* Electronic address: adiba@mail.nih.gov

1. R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford Univ. Press, New York, 2001).

2. C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences* (Springer, Berlin, 2004), 3rd ed.

3. K. L. C. Hunt and J. Ross, J. Chem. Phys. **75**, 976 (1981).

4. L. R. Pratt, J. Chem. Phys. **85**, 5045 (1986).

5. C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).

6. P. Eastman, N. Gronbech-Jensen, and S. Doniach, J. Chem. Phys. **113**, 3823 (2001).

7. O. Mazonka, C. Jarzynski, and J. Blocki, Nucl. Phys. A **611**, 335 (1999).

8. D. M. Zuckerman and T. B. Woolf, J. Chem. Phys. **111**, 9475 (1999).

9. D. M. Zuckerman and T. B. Woolf, Phys. Rev. E **63**, 016702 (2000).

10. C. Xing and I. Andricioaei, J. Chem. Phys. **124**, 034110 (2006).

11. C. Dellago, P. G. Bolhuis, and D. Chandler, J. Chem. Phys. **108**, 9236 (1998).

12. R. Elber and D. Shalloway, J. Chem. Phys. **112**, 5539 (2000).
13 P. Faccioli, M. Sega, F. Pederiva, and H. Orland, Phys. Rev. Lett. 97, 108101 (2006).
14 A. Nitzan, Chemical Dynamics in Condensed Phases: Relaxation, Transfer, and Reactions in Condensed Molecular Systems (Oxford Univ. Press, New York, 2006).
15 M. Chaichian and A. Demichev, Path Integrals in Physics, Volume I: Stochastic Process and Quantum Mechanics (Institute of Physics, London, 2001).
16 M. P. Allen and D. J. Tildesley, Computer Simulation of Liquids (Clarendon Press, Oxford, 1987).
17 L. Onsager and S. Machlup, Phys. Rev. 91, 1505 (1953).
18 T. F. Miller III and C. Predescu, J. Chem. Phys. 126, 144102 (2007).
19 D. Durr and A. Bach, Comm. Math. Phys. 60, 153 (1978).