An Introduction to Langer’s Theory

Lu Hong
Department of Bioengineering & Therapeutic Sciences
University of California, San Francisco
San Francisco, CA 94158
lu.hong@ucsf.edu

Abstract

This note provides a pedagogical introduction to Langer’s theory for activated rate processes in multiple dimensions at the high friction limit, with an emphasis on the connection between the theory and the property of the backward committor/splitting probability near the saddle point. The intended audience is assumed to have some familiarity with linear algebra and statistical mechanics while knowledge of stochastic processes is not strictly necessary.

Keywords: Kramers’ theory, KLBS theory

Contents

1 A brief review of stochastic process 2
1.1 Overdamped Langevin equation 2
1.2 Smoluchowski equation 4

2 Kramers’ theory at the high friction limit 5
2.1 Kramers’ theory via the flux-over-population method 6
2.2 Kramers’ theory via MFPT 8

3 Langer’s extension to higher dimensions 9
3.1 A relation between the steady-state probability and the backward committor 10
3.2 An ansatz for the backward committor 11
3.3 Properties of the vector normal to the isocommittor planes 13
3.3.1 $v_+$ as a generalized eigenvector 13
3.3.2 The value of the free constant $a$ 14
3.3.3 The magnitude of $v_+$ 15
3.4 Properties of the steady-state flux vector field 15
3.4.1 The flux vector field has a constant direction 15
3.4.2 The flux vector field is constant along its streamlines 16
3.4.3 Flux surface integrals are constant over dividing surfaces 16
3.5 Langer’s rate constant 19
3.6 Limitation of Langer’s theory 19

4 Simplification of Langer’s theory by Berezhkovskii and Szabo 20
4.1 Gaussian surface integral over a plane 20
4.2 Projection of Langer’s result to one dimension 22

5 Acknowledgement 23
This note is intended as an introduction to Langer’s theory [1], the multidimensional extension of Kramers’ theory for activated rate processes in the high friction limit [2]. The note is organized as follows: Section 1 briefly reviews some elements of the theory of stochastic process, which prepares for the introduction of Kramers’ theory in Section 2, where the flux-over-population method is demonstrated in the relatively simple context of a one-dimensional double-well potential. The majority of this note is devoted to the discussion of Langer’s theory in Section 3. The derivation begins by relating the steady-state probability density in the flux-over-population method to the backward committor, followed by a detailed analysis of the behavior of the committor near the saddle point. This allows us to derive an expression for the probability flux vector field, culminating in the presentation of the multidimensional rate constant in equation (106). Lastly, Section 4 discusses some more recent development by Berezhkovskii and Szabo [3] that allows one to project Langer’s result back to one dimension. Together, the results presented in this note are sometimes referred to as the KLBS theory.

Given the venerable age of Langer’s theory and the existence of several review articles [4–6] and textbook [7] partly devoted to this subject, one might wonder why this note is necessary. This introduction certainly draws heavily on the aforementioned works, but differs in two regards. First, almost all current accounts of Langer’s theory are quite concise. While this feature is excellent for experts who are in need of a quick refresher, it poses significant challenges for beginners, because the derivation of the theory can be quite tedious and involve some mathematical tricks that are not standard knowledge for a typical reader. One goal of this note, therefore, is to derive Langer’s theory while erring on the side of presenting too many, rather than too few, algebraic details.

Second, an interesting feature of Langer’s theory is that the expression for the rate constant critically depends on several basic geometric features of the committor/splitting probability near the saddle point. This connection was not explicitly stated in Langer’s original paper, and is usually not fully developed in recent accounts of the theory. This is, again, unfortunate, since this connection provides a simple setting for gaining an intuitive understanding of the behavior of the committor near an idealized transition state. Given the prominent role the committor plays in modern rate theories, such as the transition path theory [8], such intuitions can be valuable in understanding the more recent development. Therefore, as described above, a second goal of this note is to make explicit the connection of Langer’s theory to the committor function and provide a detailed analysis thereof.

The primary intended audience for this note are students and researchers in chemistry and biophysics interested in condensed-phase simulations of macromolecules. The reader should be familiar with linear algebra and have some exposure to classical statistical mechanics and some rate theory (e.g., Arrhenius equation); knowledge of stochastic processes is not strictly necessary although results from the theory will be invoked in some derivations.

1 A brief review of stochastic process

In this section we briefly review some concepts from the theory of stochastic process relevant for our discussion. Readers who have never studied this subject should either consult standard references on stochastic processes and statistical mechanics, or simply take the results presented below as given and fill in their missing background knowledge later. The presentation in this section partially follows that in Chapter 15 of [9].

1.1 Overdamped Langevin equation

Consider a (closed) system that is in contact with a large heat bath. In a typical (classical) simulation, the system consists of the macromolecule(s) of interest as well as some water molecules and counterions. The time evolution of the full system is determined by the Hamiltonian equations, and the equilibrium phase space distribution function is microcanonical. To avoid explicit representation of the bath, we “abstract away” the bath degrees of freedom, and the remaining system dynamics can be described via the generalized
Langevin equation:

\[
\begin{align*}
    r'_i(t) &= \frac{p_i(t)}{m_i} \\
    p'_i(t) &= -\frac{\partial W(r)}{\partial r_i} - \int_0^t \sum_j \sqrt{m_i m_j} \gamma_{ij}(t - \tau)r'_j(\tau) d\tau + \xi_i(t)
\end{align*}
\]  

(1)  

Here, \(r_i, p_i,\) and \(m_i\) are the position, momentum, and (renormalized) mass of the \(i\)th degree of freedom in the remaining system, and \(W(r)\) is the potential of mean force obtained from the full potential energy function by averaging over the bath degrees of freedom. The generalized Langevin equation can be derived via either the harmonic bath model or, more rigorously, via the Mori–Zwanzig theory.

The generalized Langevin equations is reminiscent of Newton’s equation of motion, but the price we pay for doing away with the bath degrees of freedom is the difficulty of dealing with a set of stochastic integro-differential equations. Specifically, compared to Newton’s equations, there are two additional terms in (2) that provide a coarse-grained model of the effect of the bath on the system. Here, we have “decomposed” the effect of the bath into two parts. First, the fluctuation term \(\xi_i(t)\) represents a random force (i.e., noise) acting on the system. Although the motion of the bath is fully deterministic, by ignoring the molecular details we model \(\xi_i(t)\) as a random, or stochastic, process. For a system solvated by a dense bath, such as liquid water, that affects the system dynamics through soft collisions (i.e., weak noise), a common model for \(\xi_i(t)\) is a Gaussian random process with zero mean.

Second, the dissipation term \(\int_0^t \gamma_{ij}(t - \tau)r'_j(\tau) d\tau\) is a convolution integral that acts as a friction force slowing down the system. This integral is called a memory integral; the term \(\gamma_{ij}(t)\) is known as the memory kernel, or dynamic friction kernel, and it encodes the “memory” of the motion of the system by the bath. Physically, the memory integral represents the fact that the bath requires a finite amount of time to respond to fluctuations in the system and this lag affects the motion of the system. At equilibrium, the fluctuation-dissipation theorem now reads

\[
\mathbb{E}[\xi_i(t)\xi_j(t')] = 2k_B T \sqrt{m_i m_j} \gamma_{ij}(|t - t'|)
\]  

(3)  

Now, we make some further simplifying assumptions to make (2) more analytically tractable. First, we assume that the bath responds instantaneously to the motion of the system; i.e., we assume that the memory integral decays instantaneously and the bath has no memory of the system history. This is a good model when the (renormalized) mass of the system is much larger than that of the bath. In this case, the memory kernel becomes \(\gamma_{ij}(t) = 2\gamma_{ij}\delta(t)\), where we have defined

\[
\gamma_{ij} = \int_0^\infty \gamma_{ij}(t) dt
\]  

(4)  

as the static friction kernel, or simply the friction coefficient, and \(\delta(t)\) is a Dirac delta function. The fluctuation-dissipation theorem now reads

\[
\mathbb{E}[\xi_i(t)\xi_j(t')] = 2k_B T \sqrt{m_i m_j} \gamma_{ij} \delta(|t - t'|)
\]  

(5)  

Stochastic processes with an autocorrelation function of this form are called white noises, and the terminology reflects the fact that the power spectral density of the process is a constant over all frequencies. With these assumptions, the generalized Langevin equation becomes

\[
\begin{align*}
    r'_i(t) &= \frac{p_i(t)}{m_i} \\
    p'_i(t) &= -\frac{\partial W(r)}{\partial r_i} - \sum_j \sqrt{m_i m_j} \gamma_{ij} r'_j(t) + \xi_i(t)
\end{align*}
\]  

(6)  

which is known as the Langevin equation. Compared to the generalized Langevin equation, the Langevin equation describes stochastic processes that are Markovian (i.e., memoryless).
Second, for dense solvent, such as water, the high friction and frequent collisions with the system leads to, on a short timescale (i.e., \( t < \gamma^{-1} \)), rapid fluctuations in the acceleration \( p_i'(t) \). However, on a longer timescale the change in the time-averaged velocity will be small as the effect of collisions cancel out each other. This allows us to set the acceleration in (7) to zero, which leads to

\[
\sum_j \sqrt{m_i m_j} \gamma_{ij} r_j'(t) = -\frac{\partial W(r)}{\partial r_i} + \xi_i(t)
\]  

(8)

The motion described by (8) goes by many names, such as diffusion, Brownian motion, or overdamped Langevin dynamics; the equation itself is known as the overdamped Langevin equation. In some applications, the cross-correlation terms in the memory kernel are ignored, in which case (8) takes on a simpler form

\[
m_i \gamma_{ii} r_i'(t) = -\frac{\partial W(r)}{\partial r_i} + \xi_i(t)
\]  

(9)

### 1.2 Smoluchowski equation

For the purpose of describing the kinetics of activated rate processes at equilibrium, working directly with (8) is inconvenient, since trajectories consistent with (8) are individual realizations of the system dynamics, while we are more interested in the statistics of an ensemble of such realizations. In other words, we are more interested in \( p(r,t) \), the probability density of the system (strictly speaking, \( p(r,t) \) is a conditional probability density function more appropriately denoted as \( p(r,t|r_0,t_0) \)). The time evolution of \( p(r,t) \) for processes governed by (8) is given by

\[
\frac{\partial p(r,t)}{\partial t} = -\nabla \cdot J(r,t) \quad \text{with} \quad J(r,t) = -D\pi(r)\nabla p(r,t) / \pi(r)
\]  

(10)

where \( \pi(r) \) is the Boltzmann distribution in the configuration space and the stationary solution to (10), \( J(r,t) \) is the (probability) flux, and \( D \) is the (position-independent) diffusion matrix, which we assume to be symmetric positive definite. Elements of the diffusion matrix are related to the static friction kernel defined in (4) via the relation

\[
D_{ij} = \frac{k_B T}{\sqrt{m_i m_j} \gamma_{ij}}
\]  

(11)

Equations such as (10) are known as the Smoluchowski equation, a special case of a class of partial differential equations known as the Fokker-Planck equations, or Kolmogorov’s forward equations. It is also common to refer to the first part of (10) as the continuity equation for the probability density, and the second part of (10) as a “constitutive” equation.

For the following discussion it is also convenient to rewrite (10) in two forms. First, we can rewrite (10) component-wise:

\[
\partial_t p(r,t) = \sum_{ij} \partial_i D_{ij} [\beta \partial_j W(r) + \partial_j] p(r,t)
\]  

(12)

where we have used a short-hand notation \( \partial_i = \partial / \partial t \) and \( \partial_i = \partial / \partial r_i \). For readers not familiar with the component-wise notation, they should convince themselves that, e.g., the \( i \)th component of a matrix-vector product \( A x \) can be expressed as \( \sum_j A_{ij} x_j \).

Second, we can rewrite (10) as an equation involving the Fokker-Planck operator \( L^\dagger \),

\[
\frac{\partial p(r,t)}{\partial t} = L^\dagger p(r,t) \quad \text{where} \quad L^\dagger = \nabla \cdot D\pi(r)\nabla \pi^{-1}(r)
\]  

(13)

Here, the symbol \( \dagger \) indicates that the Fokker-Planck operator \( L^\dagger \) is the adjoint of another operator \( L \) known as the generator,

\[
L = \pi^{-1}(r) \nabla \cdot D\pi(r)\nabla
\]  

(14)
The concept of an adjoint operator is a generalization of the Hermitian transpose of a matrix. As (13) shows, the Fokker-Planck operator dictates the time evolution of the probability density \( p(r, t) \). There is a similar interpretation for the generator: the generator dictates the time evolution of (conditional) ensemble averages, or observables, of the form \( u(r, t) = \mathbb{E}[f(r(t)) | r(0) = r_0] \), for some suitable scalar function \( f \) defined on the configuration space; i.e.,

\[
\frac{\partial u(r, t)}{\partial t} = Lu(r, t) \tag{15}
\]

2 Kramers’ theory at the high friction limit

In this section we consider Kramers’ theory in the context of a one-dimensional Brownian particle whose motion is described by the one-dimensional version of the overdamped Langevin equation (8),

\[
m \gamma \dot{r}(t) = -\frac{dW}{dr} + \xi(t) \quad \text{where} \quad \mathbb{E}[\xi(t)] = 0, \mathbb{E}[\xi(t)\xi(t')] = 2k_B T m \gamma \delta(|t - t'|) \tag{16}
\]

From (10), it follows that the probability density \( p(r, t) \) satisfies the following Fokker-Planck equation

\[
\frac{\partial p(r, t)}{\partial t} = -\frac{\partial J(r, t)}{\partial r} \quad \text{with} \quad J(r, t) = -D \left[ \beta \frac{dW}{dr} + \frac{\partial}{\partial r} \right] p(r, t) \tag{17}
\]

where \( D = k_B T / m \gamma \) is the position-independent diffusion constant. The stationary solution to (17) is the Boltzmann distribution \( \pi(r) = e^{-\beta W(r)} / Z \), where \( Z \) is a configurational partition function.

Here, we take \( W(r) \) to be an (asymmetric) double-well potential, where the minimum of the reactant well \( A \) is located at \( r_A \), the minimum of the product well \( B \) is located at \( r_B \), and the transition state \( r^\ddagger \) is identified as the position of the peak of \( W(r) \) between the two wells; without loss of generality, we assume that \( r_A < r^\ddagger < r_B \). An example of such a potential is shown in Figure 1. Further, we assume that the height of the barrier \( W(r^\ddagger) \) is much larger than \( k_B T \), so that there is a separation of timescales between barrier crossing and within-well equilibration. A related assumption here is that barrier crossing is much slower than the correlation time of the dynamic friction kernel so that the static friction kernel approximation can be justified. Together, these assumptions amount to the situation where a single slow degree of freedom in the system \( r \) is sufficient for describing the reaction, while the rest of the degrees of freedom relax much faster than the timescale of barrier crossing along \( r \); we should thus interpret “bath” as also including other degrees of freedom of the macromolecule that are not explicitly treated, and interpret \( r \) not necessarily as a position variable but more generally as a reaction coordinate (i.e., some function of the position variables \( r_i \)).

We mention here in passing that Kramers also derived expressions for the rate constant in the weak and moderate-to-high friction regimes. We will not discuss these results since they are not as relevant for condensed-phase simulations.

![Figure 1: An example one-dimensional double-well potential of mean force. The potential is given by \( W(r) = -\ln \left( e^{-2(x+1)^2} + e^{-2(x-1)^2+1} \right) \).](image)
2.1 Kramers’ theory via the flux-over-population method

Kramers derived an expression for the rate constant of the \( A \rightarrow B \) reaction, \( k_{AB} \), using what is now known as the flux-over-population method. Consider a hypothetical procedure in which an ensemble of \( n \) particles are prepared in the reactant well \( A \) (usually we set \( n = 1 \)), where they rapidly reach thermal equilibrium on a timescale much faster than that of escaping to the product well \( B \). Whenever a particle escapes from the reactant well and reaches the product well, it is immediately removed and a new one is added to the reactant well, such that the reactant population is always maintained at \( n \). As we show in the analysis below, the exact positions at which a particle is considered escaped and at which a new one is inserted are not very important. As the system reaches a non-equilibrium steady-state, there is a non-zero probability current \( J \), or flux, into the product well, which is the reaction rate (i.e., the number of transitions from \( A \) to \( B \) per unit time). We normalize this reaction rate by the reactant population, which gives the rate constant,

\[
k_{AB} = \frac{J}{n}
\]

hence the name “flux-over-population”.

Before flushing out the algebraic details, let us pause for a moment and consider why the flux-over-population procedure is needed. In other words, why does the calculation of an equilibrium rate constant invoke the flux of a seemingly contrived non-equilibrium process? A simple answer is that the (net) flux between any two states (micro- or macro-) in equilibrium is zero, due to detailed balance, and thus gives no information about kinetics. To see why the flux-over-population procedure circumvent this problem, let us consider the behavior of particles at some point near the transition state and see how they contribute to the flux across that point at equilibrium. At any given moment, such a particle can be categorized into one of four groups, depending on its past and future behavior:

1. The particle came from \( A \) and will move to \( B \) before going back to \( A \).
2. The particle came from \( A \) and will go back to \( A \) before moving to \( B \).
3. The particle came from \( B \) and will move to \( A \) before going back to \( B \).
4. The particle came from \( B \) and will go back to \( B \) before moving to \( A \).

Behavior described in group 2 and 4 are known as barrier recrossing and does not contribute to the flux, since each pair of crossing-recrossing cancel out each other (this is to be contrasted with the procedure in transition-state theory, where all (re)crossing events are counted towards the total flux across some dividing surface [10]). Because the net equilibrium flux is zero, this implies that the flux contributed by group 1 and 3 cancels as well. Note that, for ergodic dynamics, this categorization is exhaustive and it is not possible for a particle to cross the barrier once and stay in one of the two wells forever. By removing particles reaching the product well, the flux-over-population procedure removes the flux contribution from group 3 (and group 4), while the only remaining nonzero flux contribution from group 1 stays close to its equilibrium value at the steady state due to the separation of timescales. As such, the nonequilibrium steady-state flux \( J \) in this procedure gives the true \( A \) to \( B \) reaction rate.

At steady state, the population distribution \( p_{ss}(r) \) is time-independent, and thus by (17) the flux \( J \) is now both time- and position-independent. The two quantities are related via

\[
J = -D \beta \frac{dW}{dr} \frac{\partial}{\partial r} p_{ss}(r) = -D e^{-\beta W(r)} \frac{d}{dr} \left[ e^{\beta W(r)} p_{ss}(r) \right]
\]

(19)

Now, divide both sides by \( De^{-\beta W(r)} \) and integrate both sides over the interval \( [r_A, r_B] \), which gives

\[
J \int_{r_A}^{r_B} D^{-1} e^{\beta W(r)} dr = -e^{\beta W(r)} p_{ss}(r) \bigg|_{r=r_B}^{r=r_A}
\]

(20)

We now consider the behavior of \( p_{ss}(r) \) near the two boundaries: \( r_A \) and \( r_B \). At \( r=r_B \), we require \( p_{ss}(r_B) = 0 \) to satisfy the absorbing boundary condition caused by particle removal. At positions in the reactant well
away from the absorbing boundary condition at \( r_B \), we assume that the system is close to equilibrium, i.e., \( p_{ss}(r) \approx \pi(r) \), because the particle insertion procedure maintains the reactant population \( n \) in the reactant well, and inserted particles thermalize much faster than the reaction timescale. This observation has two consequences. First, it implies that \( p_{ss}(r_A) = \pi(r_A) \). Second, it implies that the population in the reactant well is

\[
n = \int_{-\infty}^{r^\dagger} p_{ss}(r) \, dr \approx \int_{-\infty}^{r^\dagger} \pi(r) \, dr \tag{21}
\]

Now, apply the boundary conditions at \( p_{ss}(r_A) \) and \( p_{ss}(r_B) \) to (20), then apply the flux-over-population formula (18), and we arrive at

\[
k_{AB} = \frac{e^{\beta W(r_A)} p_{ss}(r_A) - e^{\beta W(r_B)} p_{ss}(r_B)}{\int_{r_A}^{r_B} D^{-1} e^{\beta W(r)} \, dr \int_{-\infty}^{r^\dagger} p_{ss}(r) \, dr} = \frac{\int_{r_A}^{r_B} D^{-1} e^{\beta W(r)} \, dr \int_{-\infty}^{r^\dagger} \pi(r) \, dr}{\int_{r_A}^{r_B} D^{-1} e^{\beta W(r)} \, dr \int_{-\infty}^{r^\dagger} e^{-\beta W(r)} \, dr} \tag{22}
\]

If we assume that \( W(r) \) is harmonic around \( r^\dagger \) and \( r_A \), we write, for \( W(r) \) near \( r^\dagger \) and \( r_A \),

\[
W(r) \approx \Delta W^\dagger - \frac{1}{2} \kappa^\dagger (r - r^\dagger)^2 \quad \text{and} \quad W(r) \approx \frac{1}{2} \kappa_A (r - r_A)^2 \tag{25}
\]

where we have defined \( \Delta W^\dagger = W(r^\dagger) - W(r_A) \) and set \( W(r_A) = 0 \) (resetting the zero of the potential of mean force has the effect of changing the partition function, which has no effect on our derivation). Here, \( \kappa_A \) and \( \kappa^\dagger \) are the force constants of the harmonic potentials; equivalently, \( \kappa_A \) and \( \kappa^\dagger \) can also be interpreted as the curvatures at \( r_A \) and \( r^\dagger \), respectively. Applying (25) to (24) gives the following approximate results

\[
\int_{r_A}^{r_B} D^{-1} e^{\beta W(r)} \, dr \approx \int_{-\infty}^{\infty} D^{-1} e^{\beta (\Delta W^\dagger - \kappa^\dagger r^2/2)} \, dr = D^{-1} \sqrt{\frac{2\pi}{\beta \kappa^\dagger}} e^{\beta \Delta W^\dagger} \tag{26}
\]

\[
\int_{-\infty}^{r^\dagger} e^{-\beta W(r)} \, dr \approx \int_{-\infty}^{\infty} e^{-\beta \kappa_A r^2/2} \, dr = \sqrt{\frac{2\pi}{\beta \kappa_A}} \tag{27}
\]

These results provide a good approximation whenever the region of validity for the harmonic approximation is large enough that the added probability mass by taking the integration limit to infinity is negligible.

Substituting (26) and (27) back to (24) gives

\[
k_{AB} = \frac{\beta D}{2\pi} \sqrt{\frac{\kappa_A \kappa^\dagger}{\beta}} e^{-\beta \Delta W^\dagger} \tag{28}
\]

This is the rate constant predicted by Kramers’ theory at the high friction limit. Since the dynamics at high friction is diffusive, it is also known as the spatial-diffusion-limited rate. The rate constant for the reverse reaction can be obtained analogously by replacing the reactant well \( A \) with the product well \( B \) in the preceding derivation. If the diffusion constant is position-dependent, the derivation up to (and include) (24) is still valid. The diffusion constant can be “folded” into the potential of mean force \( W(r) \) as

\[
e^{\beta W(r)} / D(r) = e^{\beta W(r) - \ln D(r)} = e^{\beta [W(r) - k_B T \ln D(r)]} \tag{29}
\]

This defines a new potential surface \( W(r) - k_B T \ln D(r) \), which may reach its maximum at a position other than \( r^\dagger \).

In the literature it is also common to parameterize the harmonic potential near \( r^\dagger \) and \( r_A \) as

\[
W(r) \approx \Delta W^\dagger - \frac{1}{2} m \omega^2 \left( r - r^\dagger \right)^2 \quad \text{and} \quad W(r) \approx \frac{1}{2} m \omega_A^2 (r - r_A)^2 \tag{30}
\]
Here, $\omega_A$ (and analogously, $\omega^\dagger$) is the angular frequency of a harmonic oscillator of mass $m$ on the potential surface $W(r) = m\omega^2_A r^2/2$; $\omega_A$ is related to the force constant/curvature $\kappa_A$ by $\kappa_A = m\omega_A^2$. Together with the definition of the diffusion constant $D = 1/\beta m \gamma$, (28) can be written alternatively as

$$k_{AB} = \frac{\omega_A \omega^\dagger}{2\pi \gamma} e^{-\beta \Delta W^\dagger} \tag{31}$$

### 2.2 Kramers’ theory via MFPT

There is an alternative approach to deriving (28) using the mean first-passage time (MFPT). Let $\Omega$ be the reaction coordinate space of the system. For a subset $B \subseteq \Omega$, the MFPT $m_B(r)$ is the average time for the system initiated at $r$ to reach $B$. It can be shown, using potential theory, that the MFPT satisfies the following (Dirichlet) boundary value problem,

$$\begin{cases}
Lm_B(r) = -1, & r \in \Omega \setminus B \\
m_B(r) = 0, & r \in B
\end{cases} \tag{32}$$

where the notation $\Omega \setminus B$ denotes the complement of $B$ in $\Omega$. For one-dimensional Brownian dynamics, the generator $L$ defined in (14) takes on the form

$$L = \frac{d}{dr} D \left[ \frac{d}{dr} - \beta W(r) \right] = D \frac{d^2}{dr^2} - D \beta \frac{dW}{dr} \frac{d}{dr} \tag{33}$$

Unfortunately, a full derivation of (32) from scratch would take too long; interested readers should consult standard references on Markov processes for more details. Nevertheless, an “intuitive” justification of (32) goes as follows: consider a hypothetical procedure where an ensemble of system trajectories are prepared, all with the same initial condition at position $r(0) = r_0$ outside $B$; we can calculate $m_B(r_0)$ by taking the average first hitting time to the set $B$ among the trajectories. For each given trajectory $r(t)$, moving forward in time by a small amount $\delta t$ reduces the first hitting time at $r(t + \delta t)$ by the same amount of $\delta t$; as such, the time derivative of $m_B(r_0)$ should be $-1$. Since the generator acting on an observable $m_B(r_0)$ gives its time derivative, it follows that $Lm_B(r_0) = -1$. On the other hand, if $r_0$ is already in $B$, then the first hitting time to $B$ is zero by construction, hence the boundary condition in (32).

To determine the rate constant for the transition from the reactant well to the product well, let us consider the MFPT to the product well from some point $r$ near the reactant well (as we will see, the particular choice of $r$ is not very important). More specifically, let $\Omega = \mathbb{R}$ and consider the MFPT for the set $B = [r_B, \infty)$. The differential equation $Lm_B(r) = -1$ can be solved using an integrating factor,

$$D\frac{d^2m_B}{dr^2} - D\beta \frac{dW}{dr} \frac{dm_B}{dr} = -1 \tag{34}$$

$$\int_t^{-\infty} \frac{d}{ds} \left[ e^{-\beta W(s)} \frac{dm_B}{ds} \right] ds = -D^{-1} \int_t^{-\infty} e^{-\beta W(s)} ds \tag{35}$$

$$\int_r^{r_B} \frac{dm_B}{dt} dt = -\int_r^{r_B} D^{-1} e^{\beta W(t)} \int_t^{-\infty} e^{-\beta W(s)} ds dt \tag{37}$$

$$m_B(r) = \int_r^{r_B} D^{-1} e^{\beta W(t)} \int_t^{-\infty} e^{-\beta W(s)} ds dt \tag{38}$$

It is easy to check that, under the harmonic approximations of (25), the function $e^{\beta W(t)} e^{-\beta W(s)}$ has a unique global maximum at $(s = r_A, t = r^\dagger)$ within the domain of integration in (38) (see Figure 2). Since $e^{\beta W(t)} e^{-\beta W(s)}$ decays exponentially away from the global minimum, we can simply extend the domain of integration to $\mathbb{R}^2$, in which case the double integral in (38) simplifies to the integrals considered in (26) and
Figure 2: The region of integration for the double integral in (38). The region of integration is shown in green. The maxima of $e^{\beta W(t)}e^{-\beta W(s)}$ are shown as intersections of the orange lines. Only one of the two maxima, $(s = r_A, t = r^\dagger)$, is located inside the integration limit.

(27). Note that the validity of this approximation implies that $m_B(r)$ is a constant near and to the left of the reactant well, and that the particular choice for the boundaries of $B$ is not important as long as its lower bound is sufficiently to the right of $r^\dagger$.

Taken together, we have shown that the rate constant (28) derived from the flux-over-population method is equivalent to the inverse of the mean first-passage time; in fact, this equivalence is exact for a much broader range of stochastic processes than those considered in this note [11]. Unfortunately, despite its conceptual simplicity, the MFPT cannot in general be calculated analytically for problems in more than one dimension and thus we will not consider this method in the following sections; however, see Section VII. D. of [4] for an approximate approach to solving for the MFPT in higher dimensions using asymptotic methods.

3 Langer’s extension to higher dimensions

Now let us consider the extension of Kramers’ theory to an $N$-dimensional potential (of mean force) with the reactant and product wells separated by a saddle point; recall that the saddle point is a point on the potential energy surface $W(r)$ where the gradient is zero and the Hessian has exactly one unstable mode. An example two-dimensional double-well potential is shown in Figure 3A. Conceptually, the flux-over-population method employed in the one-dimensional case carries over with few changes. However, as we will see, the higher-dimensional setting of (10) does pose some algebraic challenges. The derivation presented here loosely follows the presentation in [1, 4, 5].

Similar to our treatment in one dimension, we assume a harmonic approximation at $r_A$, the bottom of the reactant well $A$,

$$W(r) \approx \frac{1}{2}(r - r_A)^T H_A (r - r_A)$$

where $H_A$ is the Hessian matrix of second-order derivatives evaluated at the minimum $r_A$, which we assume to be symmetric positive definite; note that the Hessian is also known as the force constant matrix (to see why, note that for a harmonic potential $\kappa r^2/2$ with force constant $\kappa$, the second-order derivative at the center $r = 0$ is simply $\kappa$). In terms of notations, in this note we will exclusively use notations of the form $x^T y$ to denote inner products in place of other common notations such as $x \cdot y$ or $\langle x, y \rangle$; readers more familiar with the other notations should convince themselves that, e.g., the quadratic form in (39) can be expressed in the angle bracket notation as $\langle r - r_A, H_A (r - r_A) \rangle/2$.

In addition, we assume that there is a saddle point $r^\dagger$ in between the reactant and product wells, which is
Figure 3: An example two-dimensional double-well potential. A) Contour plot of the double-well potential; the potential is given by \( W(x, y) = \frac{5}{2}(x^2 - 1)^2 + 5y^2 \) (see also [12]). The reactant well \( A \) and the product well \( B \) are defined, somewhat arbitrarily, as disks of radius 1/4 centered at the global minima \( r_A \) and \( r_B \). The white dashed vectors show the eigenvectors of \( H^\dagger \), while the white solid vectors show the eigenvectors of \( H^\dagger D \), with \( D = \{ \{4, 1\}, \{1, 2\} \} \). B) The exact backward committor \( q^-(x, y) \) and steady-state flux \( J(x, y) \). The committor is computed by solving (44) numerically using the \texttt{NDSolveValue} function with the appropriate \texttt{DirichletCondition} in Mathematica 12.1 and is shown as a contour plot (with 0.1 contour increment). The steady-state flux is calculated using (76) and represented as the orange vector field; the lengths of the streamline segments are proportional to the vector field magnitude. C) The approximate backward committor and steady-state flux from Langer’s theory. The committor is given by (74), and the steady-state flux is given by (78) with \( \pi(s) \) replaced with its harmonic approximation at the saddle point. For all calculations \( \beta \) is set to one.

the point with the minimum energy along the barrier ridge. A harmonic approximation at the saddle point gives

\[
W(r) \approx \Delta W^\dagger + \frac{1}{2}(r - r^\dagger)^T H^\dagger (r - r^\dagger)
\]  

(40)

Here, we assume that \( H^\dagger \) has exactly one negative eigenvalue associated with the unstable mode at \( r^\dagger \), while the rest of the eigenvalues are assumed to be strictly positive. In general, it is possible for \( H^\dagger \) to possess one or more eigenvalues of zero, which correspond to some underlying symmetries of the system. For example, the Hessian matrix of a system of two one-dimensional particles with positions \( x_1 \) and \( x_2 \) whose distance is subject to the harmonic restraint \( W(x_1, x_2) = \kappa(r_0 - (x_1 - x_2))^2/2 \) will have a zero eigenvalue corresponding to the translational symmetry of the system. Such symmetries can oftentimes be eliminated (e.g., here by working with the reaction coordinate \( r = x_1 - x_2 \)) and we shall not consider such cases for the following analysis. Lastly, we record here for future reference,

\[
\partial_t W(r) = \sum_j H_{ij}^\dagger (r_j - r_j^\dagger)
\]  

(41)

This result comes from the fact that \( H^\dagger \) is symmetric, and, for any quadratic form \( x^T A x \), \( \nabla_x x^T A x = (A + A^T)x \).

3.1 A relation between the steady-state probability and the backward committor

Instead of seeking a steady-state solution \( p_{ss}(r) \) to (10) directly, we take an indirect approach that relates \( p_{ss}(r) \) to a function called the backward committor \( q^-(r) \), which is the probability that, going back in time, a trajectory starting at \( r \) will reach the reactant well before the product well. As we will see, \( p_{ss}(r), \) a nonequilibrium quantity, is related to the behavior of \( q^-(r) \) at equilibrium by a surprisingly simple equation. Since Brownian dynamics under detailed balance is time-reversible (in the sense that a time-reversed process is statistically indistinguishable from the original one), it is more convenient in this context to think of \( q^-(r) \) as representing the probability that, going forward in time, a trajectory starting at \( r \) will reach the reactant well before the product well. The choice of the superscript notation is intended to distinguish \( q^- \) from \( q^+ \), the forward committor, which is the probability that, going forward in time, a trajectory starting at \( r \) will
reach the product well before the reactant well; for time-reversible processes at equilibrium the two quantities are related by \( q^+(r) + q^-(r) = 1 \). The committor is also known as splitting probability, or, in the protein folding literature, \( p_{\text{fold}} \). One should be careful not to confuse the definition of time reversibility in the theory of stochastic process with the concept of reversibility in Hamiltonian dynamics (i.e., flipping both time and momentum leaves the process invariant), the concept of reversibility in thermodynamics, or the concept of reversible reactions in kinetics.

In this section we show that

\[
q^-(r) = p_{\text{ss}}(r)/\pi(r) \tag{42}
\]

First, as in the one-dimensional case, the steady-state distribution \( p_{\text{ss}}(r) \) satisfies the following boundary value problem,

\[
\begin{align*}
L^\dagger p_{\text{ss}}(r) &= 0, \quad r \in \Omega \setminus (A \cup B) \\
p_{\text{ss}}(r) &= \pi(r), \quad r \in A \\
p_{\text{ss}}(r) &= 0, \quad r \in B
\end{align*} \tag{43}
\]

Second, similar to the MFPT, the backward committor also satisfies a boundary value problem,

\[
\begin{align*}
Lq^-(r) &= 0, \quad r \in \Omega \setminus (A \cup B) \\
q^-(r) &= 1, \quad r \in A \\
q^-(r) &= 0, \quad r \in B
\end{align*} \tag{44}
\]

In general, the time evolution of the backward committor is actually dictated by \( L^R \); the generator for the time-reversed process, but \( L^R = L \) for time-reversible processes. Again, we will not attempt to derive (44) rigorously. An “intuitive” justification is as follows: consider the same hypothetical procedure as before, where an ensemble of system trajectories are prepared, all with the same initial condition at position \( r(0) = r_0 \) outside of \( A \) or \( B \); we can calculate \( q^-(r_0) \) by determining the fraction of these trajectories that reach \( A \) first before \( B \). The fate of each individual realization of the system dynamics is time-independent: a trajectory either eventually reaches \( A \) first before it reaches \( B \), or it doesn’t; therefore, the fraction \( q^-(r_0) \) is time-independent, and thus \( Lq^-(r_0) \), which gives the time derivative, is zero. If \( r_0 \) is already in \( A \), then \( q^- = 1 \) by construction since it is not possible for the trajectory to reach \( B \) before \( A \); similarly, if \( r_0 \) is already in \( B \), then \( q^- = 0 \) since the trajectory is already in \( B \) before it can reach \( A \).

Comparing (13) with (14), we see that the generator is related to the Fokker-Planck operator by

\[
L = \pi(r)^{-1}L^\dagger\pi(r) \tag{45}
\]

Substituting (45) in (44) gives \( \pi(r)^{-1}L^\dagger\pi(r)q^-(r) = 0 \); this implies that \( L^\dagger\pi(r)q^-(r) = 0 \) since \( \pi^{-1}(r) \propto e^{\beta W(r)} \) is strictly positive for all \( r \). Furthermore, \( \pi(r)q^-(r) = \pi(r) \) for \( r \in A \) and \( \pi(r)q^-(r) = 0 \) for \( r \in B \). Taken together, we see that \( \pi(r)q^-(r) \) satisfies the same well-posed boundary value problem (43) as \( p_{\text{ss}}(r) \), which implies that they are the same function, as desired. We note here that this proof does not work for stochastic processes described by the Langevin equation outside the overdamped regime.

### 3.2 An ansatz for the backward committor

Given (42), the problem of determining \( p_{\text{ss}}(r) \) now reduces to that of determining \( q^-(r) \), the backward committor. In this section, we seek an analytical expression for the backward committor near the saddle point \( r^\dagger \). To make the derivation clearer, we first perform a change of variable \( s = r - r^\dagger \) to shift the origin to the saddle point; this procedure has no effect on the functional form of any preceding expressions except for changing the independent variable from \( r \) to \( s \). At steady state, use the fact that \( p_{\text{ss}}(s) \) is time-independent and \( p_{\text{ss}}(s) = q^-(s)\pi(s) \); (12) becomes
0 = \sum_{ij} \partial_i D_{ij} \left[ \beta \partial_j W(s) + \partial_j \right] q^-(s) \pi(s) \quad (46)

0 = \sum_{ij} \partial_i D_{ij} \left[ \beta(\partial_j W(s)) q^-(s) \pi(s) + (\partial_j q^-(s)) \pi(s) - \beta \pi q^-(s) \partial_j W(s) \right] \quad (47)

0 = \sum_{ij} D_{ij} \partial_i \left[ (\partial_j q^-(s)) \pi(s) \right] \quad (48)

0 = \sum_{ij} D_{ij} \left[ \partial_i q^-(s) - \beta (\partial_i W(s)) (\partial_j q^-(s)) \right] \quad (49)

In the second equality, we used the fact that \( \partial_i \pi(s) = -\beta \pi(s) \partial_i W(s) \).

To examine the solution to (49) near the saddle point \( r^+ \), we first consider the one dimensional case. Here, (40) becomes \( W(s) = \Delta W^+ - \kappa^+ s^2/2 \) as in (25), with \( \kappa^+ \) representing the force constant for the unstable mode, and (49) reduces to

\[
\frac{d^2}{ds^2} q^-(s) + \beta \kappa^+ s \frac{d}{ds} q^-(s) = 0
\]  

Equation (50) can be solved by first solving a first-order ODE for \( \nu(s) = dq^-/ds \) using an integrating factor and then integrating both sides of the solution from 0 to \( s \). After some algebra, we arrive at the general solution in the form of an error function

\[
q^-(s) = A + \frac{B}{\sqrt{2\pi}} \int_0^s \sqrt{\beta \kappa^+} e^{-u^2/2} du
\]  

where \( A \) and \( B \) are two free constants. Since \( q^-(s) \) is a backward committor, we require that \( q^- \to 1 \) as \( s \to -\infty \) (i.e., for \( s \) in the reactant well) and \( q^- \to 0 \) as \( s \to \infty \) (i.e., for \( s \) in the product well). Substituting these limits into (51) gives \( A - B/2 = 1 \) and \( A + B/2 = 0 \); it should be obvious that \( A = 1/2 \) and \( B = -1 \), which means that the solution to (50) that satisfies the appropriate boundary conditions is

\[
q^-(s) = \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \int_0^s \sqrt{\beta \kappa^+} e^{-u^2/2} du = \frac{1}{\sqrt{2\pi}} \int_s^\infty \sqrt{\beta \kappa^+} e^{-u^2/2} du = \sqrt{\frac{3\kappa^+}{2\pi}} \int_s^\infty e^{-\beta \kappa^+ u^2/2} du
\]  

This solution indicates that the committor drops down from 1 to 0 with a sigmoidal functional form as \( s \) increases. In particular, at \( s = 0 \), \( q^- = 1/2 \).

This analysis of the one-dimensional backward committor provides important hints about the behavior of the committor in higher dimensions. In this setting, the set of all points \( s \) for which \( q^-(s) \) has the same value constitutes an isocommittor surface (or curve, in two dimensions). Near the transition state, it is not unreasonable to assume that these isocommittor surfaces can be approximated as planes (or lines, in two dimensions) if \( q^- \) is sufficiently smooth. Furthermore, within the region where the harmonic approximation is valid, this family of isocommittor planes needs to be parallel, since level surfaces cannot intersect for a well-defined function.

These considerations motivate us to devise a solution to (49) in terms of isocommittor planes. In general, a plane passing through the origin can be parameterized using its normal vector \( n \) as \( n^T s = 0 \); for planes that do not cross the origin, they can still be parameterized as \( n^T(s - s_0) = 0 \), where \( s_0 \) is a displacement vector from some point on the plane to the origin. In light of this functional form, let us define a vector \( v_+ \) that is normal to the isocommittor planes (the reason for this notation will become clear shortly); using this vector, the family of isocommittor planes can be parameterized as \( v_+^T s = u, u \in \mathbb{R} \). We orient \( v_+ \) such that \( q^- \) approaches 1 for \( u < 0 \) and approaches 0 for \( u > 0 \).

Taken together, the analysis so far suggests that the solution to (49) has the form

\[
q^-(s) = \frac{a}{\sqrt{2\pi}} \int_{v_+^T s}^\infty e^{-(au)^2/2} du
\]  

(53)
where \( a > 0 \) is a free constant to be determined. It is easy to check that with this solution form, \( q^-(s) \) is constant on each isocommitter plane, approaches 1 for \( s \) near the reactant well, and approaches 0 for \( s \) near to the product well.

One particular isocommitter plane of interest is the one that crosses the saddle point; i.e., the plane parameterized by \( u = 0 \) and where \( q^- = 1/2 \). This plane is sometimes known as the stochastic separatrix and often plays a special role in the analysis of reaction mechanisms: if we define the transition state (or more accurately, the transition state ensemble) as the set of configurations where the committor is 1/2, the transition state is the stochastic separatrix. In Kramers’ theory, this plane reduces to the peak of the barrier separating the reactant and product state. Unfortunately, the simplicity of this equivalence between the transition state and the barrier peak often does not hold in practice, depending on the choice of the reaction coordinate and the theoretical framework under consideration; see, e.g., [10, 13].

### 3.3 Properties of the vector normal to the isocommitter planes

The ansatz in (53) still leaves some questions unanswered. Specifically,

1. How does the direction of \( v_+ \) relate to the geometry at the saddle point?
2. What is the magnitude of \( v_+ \)?
3. What is the value of the free constant \( a > 0 \)?

We address these questions in this section.

#### 3.3.1 \( v_+ \) as a generalized eigenvector

To understand the direction of \( v_+ \), we first argue that \( v_+ \) is actually an eigenvector of the matrix \( H^\dagger D \). In order to show this, we need to examine the solution of (49) on the \( q^- = 1/2 \) plane.

First, using the Leibniz integral rule, we record here the first- and second-order derivatives of (53) with respect to \( s \):

\[
\partial_i q^-(s) = -\frac{a}{\sqrt{2\pi}} e^{-(av^+_i s)^2/2} v_{+i} \quad \text{and} \quad \partial_{ij} q^-(s) = \frac{a^3}{\sqrt{2\pi}} (v^T_+ s) e^{-(av^+_i s)^2/2} v_{+i} v_{+j} \quad (54)
\]

On the \( q^- = 1/2 \) plane, \( v^T_+ s = 0 \) and (54) reduces to

\[
\partial_i q^-(s) = -\frac{a}{\sqrt{2\pi}} v_{+i} \quad \text{and} \quad \partial_{ij} q^-(s) = 0 \quad \text{when} \quad v^T_+ s = 0 \quad (55)
\]

Substituting in the derivative of \( W(s) \) in (41) and the derivatives of \( q^-(s) \) in (55), we see that (49) can be further simplified on the \( q^- = 1/2 \) plane as,

\[
0 = \sum_{ij} D_{ij} \left[ -\beta \left( \sum_k H_{ik}^k s_k \right) \left( -\frac{a}{\sqrt{2\pi}} v_{+j} \right) \right] \quad (56)
\]

\[
0 = \sum_k s_k \left[ \sum_{ij} H_{ki}^j D_{ij} v_{+j} \right] \quad (57)
\]

\[
0 = s^T (H^\dagger D v_+) \quad (58)
\]

In the second equality, we used the fact that \( H^\dagger \) is symmetric (i.e., \( H_{ik}^k = H_{ki}^k \)). This result indicates that the vector \( H^\dagger D v_+ \) is orthogonal to \( s \). Since \( v_+ \) is orthogonal to \( s \) confined on the \( q^- = 1/2 \) plane, \( H^\dagger D v_+ \) is parallel with \( v_+ \). This shows that \( v_+ \) is an eigenvector of \( H^\dagger D \). Anticipating the analysis in the next section, we denote the corresponding eigenvalue as \( -\lambda_+ \) (with \( \lambda_+ > 0 \)).
Note that we can write the eigenvector equation $H^\dagger D v_+ = -\lambda_+ v_+$ as

$$H^{-\dagger} v_+ = -\lambda_+^{-1} D v_+$$

(59)

where, through an abuse of notation, we have written $(H^\dagger)^{-1} = H^{-\dagger}$. Equations of this form are called generalized eigenvalue problems whenever $H^{-\dagger}$ is symmetric and $D$ is symmetric positive definite. The collection of all such generalized eigenvectors form a nonsingular matrix $V$ that simultaneously diagonalizes $H^\dagger$ and $D$, in the sense that

$$V^T D V = I \quad \text{and} \quad V^T H^{-\dagger} V = \Lambda^{-1}$$

(60)

where $\Lambda^{-1}$ is a diagonal matrix containing the corresponding generalized eigenvalues. The vectors $v_i$’s in $V$ are orthogonal with respect to the inner product induced by $D$; i.e.,

$$\langle v_i, v_j \rangle_D = v_i^T D v_j = \lambda_j v_i^T H^{-\dagger} v_j = \lambda_j \delta_{ij} = \delta_{ij}$$

(61)

where $\delta_{ij}$ is the Kronecker delta function.

### 3.3.2 The value of the free constant $a$

The magnitudes of $a$ and $v_+$ are related; first we determine the value of $a$ in terms of $v_+$. To do so, we consider the solution to (49) outside the committor 1/2 plane where $v_+^T s \neq 0$:

$$0 = \sum_{ij} D_{ij} \left[ \frac{a^3}{\sqrt{2\pi}} (v_+^T s) e^{-(a v_+^T s)^2/2 v_+^T v_+ - \beta \left( \sum_k H^\dagger_{ik} s_k \right)} \left( -\frac{a}{\sqrt{2\pi}} e^{-(a v_+^T s)^2/2 v_+^T s} \right) \right]$$

(62)

$$0 = \sum_{ij} D_{ij} \left[ a^2 (v_+^T s) v_+^T v_+ + \beta \left( \sum_k H^\dagger_{ik} s_k \right) v_+^T v_+ \right]$$

(63)

$$0 = a^2 (v_+^T s) \sum_{ij} v_+^T D_{ij} v_+ + \beta \sum_{jik} v_+^T D_{jk} H^\dagger_{ik} s_k$$

(64)

$$0 = a^2 (v_+^T s) v_+^T D v_+ + \beta (v_+^T D H^\dagger) s$$

(65)

$$0 = a^2 (v_+^T s) v_+^T D v_+ - \beta (\lambda_+ v_+^T) s$$

(66)

$$0 = a^2 v_+^T D v_+ - \beta \lambda_+$$

(67)

$$a = \sqrt{\beta \lambda_+/v_+^T D v_+}$$

(68)

Applying (68) to (53) gives

$$q^{-}(s) = \sqrt{\frac{\beta \lambda_+}{2\pi v_+^T D v_+}} \int_{v_+^T s}^{\infty} \exp \left( -\frac{1}{2} v_+^T D v_+ \right) du$$

(69)

Note that (68) implies that $-\lambda_+$ is a negative eigenvalue of $H^\dagger D$: because $\beta > 0$ and $v_+^T D v_+ > 0$ (since $D$ is symmetric positive definite), $\lambda_+$ also needs to be positive so that $a > 0$. What does this result mean for the direction represented by the corresponding eigenvector $v_+$. When $D$ can be written in the form of $\sigma^2 I$, where $I$ is an identity matrix, the diffusion on the potential (of mean force) is isotropic (i.e., same diffusivity in every direction). In this case, $v_+$ is simply an eigenvector of $H^\dagger$,

$$H^\dagger v_+ = -\lambda_+ \sigma^{-2} v_+$$

(70)

that is, $v_+$ is an eigenvector associated with the only negative eigenvalue of $H^\dagger$ corresponding to the unstable mode separating the reactant and product well. In the case of anisotropic diffusion, $H^\dagger D$ still has a single unique unstable mode represented by $v_+$: because $H^\dagger$ is symmetric and $D$ is symmetric positive definite, a result in linear algebra states that the eigenvalues of $H^\dagger D$ have the same signs as those of $H^\dagger$. As such, $v_+$ is still an eigenvector associated with the only negative eigenvalue of $H^\dagger D$ in the anisotropic case. However, the presence of the matrix $D$ effectively stretches and rotates the dynamics on the surface of the potential; as such, the unstable mode of $H^\dagger D$ no longer necessarily points in the same direction as the unstable mode of $H^\dagger$ (see, e.g., Figure 3A).
3.3.3 The magnitude of $v_+$

The magnitude of $v_+$ determines how fast the committor $q^-(s)$ decays to zero from the reactant well to the product well. To determine this quantity, we first note that the derivative of the one-dimensional backward committor in (52) at the saddle point $s = 0$ is given by

$$\left. \frac{d}{ds} q^-(s) \right|_{s=0} = -\sqrt{\frac{\beta K^\dagger}{2\pi}} e^{-\beta \kappa^s s^2/2} \bigg|_{s=0} = -\sqrt{\frac{\beta K^\dagger}{2\pi}}$$

(71)

While the gradient of the multi-dimensional backward committor in (69) at the saddle point $s = 0$ is given by

$$\left. \nabla_s q^-(s) \right|_{s=0} = -\sqrt{\frac{\beta \lambda_+}{2\pi v_+^T D v_+}} \exp \left(-\frac{1}{2} \frac{\beta \lambda_+}{v_+^T D v_+} (v_+^T s)^2 \right) \bigg|_{s=0} v_+ = -\sqrt{\frac{\beta \lambda_+}{2\pi v_+^T D v_+}} v_+$$

(72)

Comparing these two expressions, we can identify $\kappa^\dagger$ with $\lambda_+$ since both are “eigenvalues” corresponding to the unstable mode at the saddle point. In order for the absolute value of (71) and the norm of (72) (induced by $D$) to be equal, we require that

$$v_+^T D v_+ = 1$$

(73)

Recall that the expression $v_+^T D v_+$ is the inner product of $v_+$ with itself induced by $D$ and defines a vector norm via $\|v_+\|_D = \sqrt{v_+^T D v_+}$. Taken together, we see that the solution to (49) is

$$q^-(s) = \sqrt{\frac{\beta \lambda_+}{2\pi}} \int_{v_+^T s}^{\infty} e^{-\beta \lambda_+ u^2/2} du$$

(74)

See Figure 3B and 3C for a comparison of (74) with the exact solution to (44) for the example double-well potential shown in 3A.

Lastly, we note here that, with the help of (54) (after setting $a = \sqrt{\beta \lambda_+}$), it is fairly straightforward to check that (74) is in fact an exact solution to (49) under the harmonic approximation; i.e., (74) satisfies

$$\sum_{ij} D_{ij} \partial_{ij} q^-(s) - \beta \sum_{ijk} s_k H_{ki}^j D_{ij} \partial_j q^-(s) = 0$$

(75)

in addition to satisfying the boundary conditions listed in (44). In other words, if the potential surface $W(s)$ is exactly harmonic, the isocommittor surfaces are indeed parallel planes parametrized by $v_+^T s = u$ for $u \in \mathbb{R}$.

3.4 Properties of the steady-state flux vector field

Now with an expression for the backward committor in hand, we are ready to write down an expression for the steady-state flux and analyze its behavior near the saddle point.

3.4.1 The flux vector field has a constant direction

Substituting (42) and (74) in the multidimensional Fokker-Planck equation (10), we see that the steady-state flux is

$$J(s) = -D \pi(s) \nabla_s q^-(s)$$

$$= -D \pi(s) \left( -\sqrt{\frac{\beta \lambda_+}{2\pi}} e^{-\beta \lambda_+ (v_+^T s)^2/2} \right) \nabla_s (v_+^T s)$$

(76)

(77)

$$= \sqrt{\frac{\beta \lambda_+}{2\pi}} e^{-\beta \lambda_+ (v_+^T s)^2/2} \pi(s) D v_+$$

(78)

One immediate observation here is that near the saddle point, the flux has a constant direction $D v_+$. 
3.4.2 The flux vector field is constant along its streamlines

Not only does \( J(s) \) have a constant direction \( Dv_+ \), it also has a constant magnitude along that direction. To see why, we first parameterize \( s \) using an orthogonal decomposition

\[
 s(u) = \Delta s + uDv_+ \tag{79}
\]

where \( \Delta s \) is a constant that measures the displacement from the origin (i.e., the saddle point) orthogonal to the direction defined by \( Dv_+ \), while \( u \in \mathbb{R} \) is the independent variable that measures progression along the \( Dv_+ \) direction. In (78), the only factor dependent on \( s \) is \( e^{-\beta \lambda_+ (v^+_+ s)^2/2} \pi(s) \). Using the harmonic approximation of \( \pi(s) \) at near the saddle point in (40), this expression can be written as

\[
 \exp \left[ -\frac{\beta}{2} s^T \left( H^+ + \lambda_+ v_+ v_+^T \right)s \right] \tag{80}
\]

Using the parameterization \( s(u) \), the quadratic form in this expression becomes \( (\Delta s + uDv_+)^T (H^+ + \lambda_+ v_+ v_+^T)(\Delta s + uDv_+). \) This expression can be expanded, but some of the terms after the expansion are constants independent of \( u \). Keeping only the \( u \)-dependent terms, we get

\[
u \Delta s^T H^+ Dv_+ + u \lambda_+ \Delta s^T v_+^T Dv_+ + uv_+^T Dv_+^T H^+ \Delta s + uv_+^T Dv_+ Dv_+^T H^+ \Delta s + \frac{1}{2} \lambda_+ v_+^T Dv_+ Dv_+^T + \frac{1}{2} u^2 \lambda_+ v_+^T Dv_+^T v_+ + \frac{1}{2} \lambda_+ v_+^T Dv_+ Dv_+^T \tag{81}\]

Using the facts that \( H^+ \) and \( D \) are symmetric, \( (H^+ + \lambda_+ v_+ v_+^T)(\Delta s + uDv_+) \) is a divergence-less vector field in \( \Omega \). Not only does \( J(s) \) have a constant direction \( Dv_+ \), but also a constant magnitude along that direction. To see why, we first parameterize \( s(u) \) using an orthogonal decomposition

\[
 s(u) = \Delta s + uDv_+ \tag{79}
\]

where \( \Delta s \) is a constant that measures the displacement from the origin (i.e., the saddle point) orthogonal to the direction defined by \( Dv_+ \), while \( u \in \mathbb{R} \) is the independent variable that measures progression along the \( Dv_+ \) direction. In (78), the only factor dependent on \( s \) is \( e^{-\beta \lambda_+ (v^+_+ s)^2/2} \pi(s) \). Using the harmonic approximation of \( \pi(s) \) at near the saddle point in (40), this expression can be written as

\[
 \exp \left[ -\frac{\beta}{2} s^T \left( H^+ + \lambda_+ v_+ v_+^T \right)s \right] \tag{80}
\]

Using the parameterization \( s(u) \), the quadratic form in this expression becomes \( (\Delta s + uDv_+)^T (H^+ + \lambda_+ v_+ v_+^T)(\Delta s + uDv_+). \) This expression can be expanded, but some of the terms after the expansion are constants independent of \( u \). Keeping only the \( u \)-dependent terms, we get

\[
u \Delta s^T H^+ Dv_+ + u \lambda_+ \Delta s^T v_+^T Dv_+ + uv_+^T Dv_+^T H^+ \Delta s + uv_+^T Dv_+ Dv_+^T H^+ \Delta s + \frac{1}{2} \lambda_+ v_+^T Dv_+ Dv_+^T + \frac{1}{2} u^2 \lambda_+ v_+^T Dv_+^T v_+ + \frac{1}{2} \lambda_+ v_+^T Dv_+ Dv_+^T \tag{81}\]

Using the facts that \( H^+ \) and \( D \) are symmetric, \( (H^+ + \lambda_+ v_+ v_+^T)(\Delta s + uDv_+) \) is a divergence-less vector field in \( \Omega \). Not only does \( J(s) \) have a constant direction \( Dv_+ \), but also a constant magnitude along that direction.

3.4.3 Flux surface integrals are constant over dividing surfaces

As in the one-dimensional case, we need the total flux from the reactant well to the product well, which gives the reaction rate. This task is made somewhat easy by a few considerations. First, we note that, for a double-well potential, the flux-over-population steady-state flux \( J(s) \) has exactly one source in the reactant well \( A \) and one sink in the product well \( B \); that is, \( J(s) \) is a divergence-less vector field in \( \Omega \setminus (A \cup B) \) at the steady state (this should be obvious by setting time derivative to zero in (10)). A consequence of this property is that the total flux from \( A \) to \( B \) is given by the surface integral of \( J(s) \) over a dividing surface between \( A \) and \( B \) (i.e., a surface that partitions \( \Omega \) into two sets, one of which contains \( A \) and the other contains \( B \)). Importantly, the particular choice of the dividing surface is irrelevant; if this integral differs on two dividing surfaces, it would imply that there is a net sink or source located in the in-between region. Second, since the probability flow is concentrated near the saddle point, one can restrict the flux surface integral further to regions close to the saddle point where (78) is valid.

With the help of (108), a result we will prove in Section 4, a particularly simple choice of dividing surface is the \( q^{-} = 1/2 \) plane. The total flux of (78) over the stochastic separatrix is

\[
 J = \int \|v_+\|^{-1} v_+^T J(s) d\sigma(s) \tag{82}
\]

\[
 = \frac{1}{2} \int v_+^T J(s) \delta(v_+^T s) ds \tag{83}
\]

\[
 = \sqrt{\frac{\beta \lambda_+}{2\pi}} v_+^T Dv_+ Z^{-1/2} e^{-\beta \Delta W^T} \int e^{-\beta s^T H^+ s/2} \delta(v_+^T s) ds \tag{84}
\]

\[
 = \sqrt{\frac{\beta \lambda_+}{2\pi}} Z^{-1/2} e^{-\beta \Delta W^T} \left[ \left( \frac{1}{\beta} \right)^N v_+^T H^{-1} v_+ \det H^+ \right]^{-1/2} \tag{85}
\]

\[
 = \left( \frac{2\pi}{\beta} \right)^{N/2} Z^{-1/2} e^{-\beta \Delta W^T} \frac{\beta \lambda_+}{2\pi} \left| \det H^+ \right|^{-1/2} \tag{86}
\]
Here, equality (1) follows from the relation \( d\sigma(s) = \delta(v_x^T s)\|\nabla_s v_x^T s\| \, ds \), which is a consequence of what is known as the coarea formula; note that \( \|\nabla_s v_x^T s\| = \|v_x\| \). The volume integrals are over the entire \( \Omega = \mathbb{R}^N \) reaction coordinate space of the system; this will be the assumed domain of integration for the rest of the note unless otherwise specified. Equality (2) follows from the facts that \( H^{-1}v_+ = -\lambda_+ Dv_+ \) and \( \det H^\dagger < 0 \).

In the rest of this section, we show and confirm that the total flux can be obtained using a surface integral over an arbitrary plane near the saddle point, which gives the same result as (86). Let us parameterize such a plane by \( n^T s = \theta \), where \( n \) is the normal vector (not necessarily normalized) and \( \theta \in \mathbb{R} \). The only condition we impose on the plane is that \( n \) is not orthogonal to \( Dv_+ \) (i.e., the plane parameterized by \( n^T s = \theta \) is not parallel to \( Dv_+ \), the direction of the flux). Furthermore, without loss of generality we assume that \( n \) is oriented such that \( n^T Dv_+ > 0 \). Using (78) and the coarea formula, the integral of \( J(s) \) over this plane is given by

\[
J = \int \|n\|^{-1} n^T J(s) \, d\sigma(s) \tag{87}
\]

\[
= \int n^T J(s) \delta(\theta - n^T s) \, ds \tag{88}
\]

\[
= \sqrt{\frac{\beta \lambda_+}{2\pi}} n^T Dv_+ Z^{-1} e^{-\beta \Delta W^\dagger} \int e^{-\beta s^T (H^\dagger + \lambda_+ v_+ v_+^T)s/2} \delta(\theta - n^T s) \, ds \tag{89}
\]

As with most Gaussian integrals, we seek a change of variable that makes the volume integral in (89) separable. For a typical Gaussian integral with an integrand of the form \( e^{-x^TAx/2} \), where \( A \) is a symmetric matrix, there is an orthogonal matrix \( Q \) that diagonalizes \( A \) (i.e., \( Q^T AQ = M \) for some diagonal matrix \( M \)) and enables a change of variable \( x = Q y \) that renders the integral separable. In the current case, recall our discussion of the generalized eigenvector problem in Section 3.3, where we have defined the matrix \( V \) that simultaneously diagonalizes \( H^\dagger \) and \( D \) according to (60). Comparing \( V \) and \( Q \), we introduce the change of variable \( t = V^T s \) (as well as \( m = V^T n \)). The Jacobian for this transformation is \( \det V^{-1} \), which is not necessarily one because \( V \) may not be orthogonal. After this transformation, \( v_+ \) relates to a new basis vector \( e_+ \) by \( e_+ = V^{-1} v_+ \); to see why, note that together with (60), \( H^\dagger Dv_+ = -\lambda_+ v_+ \) implies

\[
\Lambda V^{-1} v_+ = -\lambda_+ V^{-1} v_+ \tag{90}
\]

that is, \( V^{-1} v_+ \) is an eigenvector of the diagonal matrix \( \Lambda \) corresponding to its unique negative eigenvalue \( -\lambda_+ \). Since \( -\lambda_+ \) has an algebraic multiplicity of one, the eigenvector \( V^{-1} v_+ \) must have only one nonzero element. Furthermore, the norm of \( V^{-1} v_+ \) is fixed by \( \|V^{-1} v_+\| = v_+ V^{-T} V^{-1} v_+ = \|v_+\| Dv_+ = 1 \). Since \( v_+ \) is a column vector of \( V \), the nonzero element of \( e_+ \) must be positive, otherwise \( V e_+ = -v_+ \). Together, we have shown that \( e_+ \) is a normalized standard basis vector in the new coordinate system.

With this change of variable, (89) becomes

\[
J = \sqrt{\frac{\beta \lambda_+}{2\pi}} n^T Dv_+ Z^{-1} e^{-\beta \Delta W^\dagger} \int e^{-\beta t^T (\Lambda + \lambda_+ e_+ e_+^T)t/2} \delta(\theta - m^T V^{-1} V^{-T} t) \, |\det V^{-1}| \, dt \tag{91}
\]

\[
= \sqrt{\frac{\beta \lambda_+}{2\pi}} n^T Dv_+ Z^{-1} e^{-\beta \Delta W^\dagger} |\det V^{-1}| \int e^{-\beta t^T (\Lambda + \lambda_+ e_+ e_+^T)t/2} \delta(\theta - m^T D t) \, dt \tag{92}
\]

Before further simplifications of (92), a comment about the term \( t^T (\Lambda + \lambda_+ e_+ e_+^T) t \) is in order. In the new basis, \( H^\dagger \) becomes \( \Lambda \), while \( v_+ v_+^T \) becomes \( e_+ e_+^T \). In particular, \( \Lambda + \lambda_+ e_+ e_+^T \) is a diagonal matrix whose diagonal elements are the eigenvalues \( \lambda_j \)'s except for \( -\lambda_+ \), which has been deleted from \( \Lambda \) by \( \lambda_+ e_+ e_+^T \). This observation should become especially obvious using the outer product form of the spectral theorem, which gives \( \Lambda = \sum_j \lambda_j e_je_j^T \); in other words, compared to \( \Lambda \) (or \( H^\dagger \)), \( \Lambda + \lambda_+ e_+ e_+^T \) (or \( H^\dagger + \lambda_+ v_+ v_+^T \)) is rank-deficient because one of its eigenvalues corresponding to the unstable mode has been deleted. As a result, the quadratic form \( t^T (\Lambda + \lambda_+ e_+ e_+^T) t \) sums over all \( \lambda_j t_j^2 \) except for \( j \) corresponding to \( -\lambda_+ \). In light of this analysis, we will use notations \( \sum_{j \neq +} \) and \( \prod_{j \neq +} \) to denote summation or product over \( j \)'s except for \( j \) corresponding to \( -\lambda_+ \) and \( e_+ \).
Let us denote the volume integral in (92) as $I$. We will simplify $I$ with the following strategy. First, we convert the Dirac delta function in (92) into its integral representation; i.e.,

$$
\delta(f(x)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikf(x)} \, dk
$$

for any sufficiently smooth test function $f(x)$. Second, we integrate over all $t_j$’s for which $j \neq +$, then over $k$, and then over $t_+$. This order of operation ensures that every integral is an analytically tractable Gaussian integral with a linear term, which can be solved by first completing the squares:

$$
I = \frac{1}{2\pi} \int \int e^{-\beta t^T (\lambda + \lambda_+ e_+ e_+^T) t/2} e^{i k (\theta - m^T D t)} \, dk \, dt
$$

$$
= \frac{1}{2\pi} \int \int e^{i k \theta} e^{-ik \sum m_i D_{i+} t_+} \prod_{j \neq +} \left( \int e^{-\beta \lambda_j t_j^2/2} e^{-ik \sum m_i D_{ij} t_j} \, dt_j \right) \, dk \, dt_+
$$

$$
= \frac{1}{2\pi} \int \int e^{i k \theta} e^{-ik \sum m_i D_{i+} t_+} \prod_{j \neq +} \left( \sqrt{\frac{2\pi}{\beta \lambda_j}} e^{-k^2 (\sum m_i D_{ij})^2/2\beta \lambda_j} \right) \, dk \, dt_+
$$

$$
= \frac{1}{2\pi} \left( \prod_{j \neq +} \sqrt{\frac{2\pi}{\beta \lambda_j}} \right) e^{i k \theta} \frac{2\pi \beta}{\beta \lambda_j} \left( \sum_{j \neq +} \frac{1}{\lambda_j} (\sum m_i D_{ij})^2 \right)^{1/2} \left( \sum_{j \neq +} \frac{1}{\lambda_j} (\sum m_i D_{ij})^2 \right)^{1/2} \left( \frac{2\pi \sum_{j \neq +} \lambda_j^{-1} (\sum m_i D_{ij})^2}{\beta} \right)^{1/2} \left( \frac{2\pi \sum_{j \neq +} \lambda_j^{-1} (\sum m_i D_{ij})^2}{\beta} \right)^{1/2}
$$

$$
= \left( \prod_{j \neq +} \sqrt{\frac{2\pi}{\beta \lambda_j}} \right) |m^T D e_+|^{-1}
$$

Substituting (99) back into (92) gives

$$
J = \sqrt{\frac{\beta \lambda_+}{2\pi}} \left( \prod_{j \neq +} \sqrt{\frac{2\pi}{\beta \lambda_j}} \right) Z^{-1} e^{-\beta \Delta W^2} | \det V^{-1} | \frac{n^T D v_+}{|m^T D e_+|}
$$

$$
= \left( \frac{2\pi}{\beta} \right)^{N/2} \frac{\beta \lambda_+}{2\pi} Z^{-1} e^{-\beta \Delta W^2} \frac{| \det V^{-1} |}{| \det H^T D |^{1/2}}
$$

$$
= \left( \frac{2\pi}{\beta} \right)^{N/2} Z^{-1} e^{-\beta \Delta W^2} \frac{\beta \lambda_+}{2\pi} | \det H^T |^{-1/2}
$$

In equality (1), we have used the fact that

$$
n^T D v_+ = m^T V^{-1} D V e_+ = m^T V^{-1} (V^T V^{-1}) V e_+ = m^T (V^{-1} V^{-T}) e_+ = m^T D^T e_+ = m^T D e_+ > 0
$$

In equality (2), we have used the fact that

$$
| \det H^T D | = | \det H^T | \det D \quad \text{and} \quad \det D = \det V^{-T} V^{-1} = (\det V^{-1})^2
$$

both of which follow directly from (60). It should be obvious that (102) is independent of either $\theta$ or $n$ used to parameterize the dividing plane and is thus an intrinsic property of the flux. The result in (102) is the same as (86), as expected. However, the choice of $q^- = 1/2$ in the derivation of (86) is special in that the quadratic form $s^T (H^T + \lambda_+ v_+ v_+^T) s$ is simply $s^T H^T s$, which is not rank-deficient.
3.5 Langer’s rate constant

According to the flux-over-population method, the only quantity left to determine for computing the rate constant is the steady-state population \( n \) in the reactant well. This is given by

\[
n = Z^{-1} \int e^{-\beta (r-r_A)^T H_A (r-r_A)/2} dr = Z^{-1} \prod_i \int_{-\infty}^{\infty} e^{-\beta \mu_i t_i^2/2} dt_i
\]

\[= Z^{-1} (2\pi/\beta)^{N/2} \prod_i \mu_i^{-1/2} = Z^{-1} (2\pi/\beta)^{N/2} (\det H_A)^{-1/2} \quad (105)\]

where \( r - r_A = Qt \) with \( Q \) being an orthogonal matrix that diagonalizes \( H_A \), and \( \mu_i \)'s are the eigenvalues of \( H_A \). Recall that \( H_A \) is assumed to be symmetric positive definite, and thus \( \mu_i > 0 \) for all \( i \)'s.

At last, combining (86) and (105), we see that the rate constant is

\[
k_{AB} = J/n = \frac{\beta \lambda_+}{2\pi} \sqrt{\frac{\det H_A}{|\det H^\dagger|}} e^{-\beta \Delta W^\dagger} \quad \text{where} \quad H^\dagger D v_+ = -\lambda_+ v_+ \quad (106)\]

Three brief comments about this result are in order. First, we note here that it is also common in the literature to define \(-\lambda_+\) as an eigenvalue of the matrix \( \beta H^\dagger D \), while the magnitude of \( v_+ \) is still fixed by \( v_+^T D v_+ = 1 \). One can retrace our derivations and see that, starting from (65), this has the effect of replacing all occurrences of \( \beta \lambda_+ \) by \( \lambda_+ \), in which case the rate constant now reads

\[
k_{AB} = \frac{\lambda_+}{2\pi} \sqrt{\frac{\det H_A}{|\det H^\dagger|}} e^{-\beta \Delta W^\dagger} \quad \text{where} \quad \beta H^\dagger D v_+ = -\lambda_+ v_+ \quad (107)\]

Second, it is easy to check that the multidimensional result can be reduced to Kramers’ rate constant (28) in the case of a single reaction coordinate. By comparing (25) to (39) and (40), we identify \( H_A \) as \( \kappa_A \) and \( H^\dagger \) as \(-\kappa^\dagger\). Next, since \(-\lambda_+\) is the eigenvalue corresponding to the unstable mode of \( H^\dagger D \), and the eigenvalue of a \( 1 \times 1 \) “matrix” is the matrix element itself, it follows that in the one-dimensional case where only the unstable mode is considered, \(-\lambda_+\) can be identified as the product of \( H^\dagger \) (i.e., \(-\kappa^\dagger\)) and \( D \), the one-dimensional diffusion constant. Making these substitutions reduce (106) to (28).

Lastly, we briefly comment on some computational aspects of the theory. Using molecular dynamics simulations, the application of (106) requires the determination of the activation free energy \( \Delta W^\dagger \) and the diffusion matrix \( D \). The activation free energy can be obtained through a variety of well-established equilibrium techniques such as umbrella sampling, or more recent nonequilibrium techniques based on Jarzynski equality [14]. Some methods for computing the diffusion matrix are described in, e.g., [15–19].

3.6 Limitation of Langer’s theory

Some caution should be exercised in applying Langer’s theory; in this section we briefly discuss two such pitfalls. First, Langer’s result can fail in some unexpected cases where the assumption of separation of timescales implied by the flux-over-population procedure is violated. In the case of highly anisotropic diffusion, it is possible to arrange the relative positions of the reactant and product wells such that barrier (re)crossing over the saddle point can take place on a timescale much shorter than the time required to relax within either well due to slower within-well diffusion. If the diffusion anisotropy is extreme, the overall kinetics could become nonexponential, depending on the initial preparation of the system within the reactant well. In such cases, Langer’s theory predicts a rate constant that cannot properly account for the dynamics of within-well equilibration. This issue was first described in [20], and a corrected analytical expression for the rate constant in the two-dimensional case was derived in [21]. More generally, anisotropy in the diffusion matrix and/or potential surface can lead to deviations from Langer’s theory; some case studies have been documented in, e.g., [22, 23].

Second, for complex systems typically encountered in condensed-phase chemistry and biophysics, the difficulty with Langer’s theory is not the calculations entailed by the theory itself, but rather the question of
how to select a small number of reaction coordinates that can provide a “complete” description of an activated rate process. This question is an active area of research (see, e.g., [24, 25]), and in practice answering this question is often more akin to an art that relies on domain-specific knowledge and intuition. Typically, an incomplete description of the reaction leads to a loss of Markovianity of the projected stochastic process in the reaction coordinate space; mathematically, this prevents us from simplifying the generalized Langevin equation into the Langevin equation (see Section 1.1), which was the starting point for the derivation of Kramers’ theory. When the memory kernel can be reasonably approximated, one approach to treating the non-Markovian dynamics in the reaction coordinate space is the Grote-Hynes theory based on the stable states picture [26], which is akin to Kramers’ theory but with the generalized Langevin equation as its starting point; a multidimensional generalization was developed in [27]. Recent theoretical development has focused on alternative frameworks such as transition path theory [28], milestoning [29], weighted ensemble [30], and path sampling methods [31, 32] that circumvent these difficulties to various extent, usually at the expense of increased computational costs.

4 Simplification of Langer’s theory by Berezhkovskii and Szabo

It is often claimed that the committor is the ideal reaction coordinate to describe a reaction. In cases where Langer’s theory is adequate, this claim is supported by a striking result [3], which we will demonstrate in this section; namely, that projecting the system dynamics along the direction of \( v_+ \) and then applying Kramers’ theory gives the same rate constant as Langer’s result, while projections along any other directions either give an overestimation or outright do not converge. Recall from our extensive discussion in Section 3 that \( v_+ \) is the vector normal to the isocommittor planes near the saddle point and that \( v_+ \) is the eigenvector of \( H^\dagger D \) corresponding to the unique unstable mode at the saddle point.

4.1 Gaussian surface integral over a plane

Before we discuss the main result of this section, we take a detour here and show that for any vectors \( x, x_0, n \in \mathbb{R}^N \) and an invertible, symmetric matrix \( A \in \mathbb{R}^{N \times N} \),

\[
\int_{\Omega} \delta (\theta - n^T x) e^{-\beta (x-x_0)^T A (x-x_0)/2} \, dx = \frac{e^{-\beta (\theta - n^T x_0)^2/2n^T A^{-1} n} \sqrt{(2\pi/\beta)^{1-N}(n^T A^{-1} n) \det A}}{\sum_{m} \delta (\theta - n^T x_m) e^{-\beta (\theta - n^T x_0)^2/2n^T A^{-1} n} \sqrt{(2\pi/\beta)^{1-N}(n^T A^{-1} n) \det A}}
\]  

whenever \( n^T A^{-1} n \det A > 0 \). This is a Gaussian integral restricted to a plane parameterized by \( n^T x = \theta \), with \( n \) being the normal vector.

Our strategy for evaluating (108) is somewhat similar to the approach taken in Section 3.4 for evaluating the surface integral of the steady-state flux vector; one major difference here is that (108) does not involve a rank-deficient matrix. Again, we start the derivation by seeking a change of variable that makes the volume integral in (108) separable. Since \( A \) is symmetric, there exists a diagonal matrix \( \Lambda \) and orthogonal matrix \( Q \) such that \( Q^T A Q = \Lambda \). Here, the diagonal elements in \( \Lambda \) contain the eigenvalues of \( A \) and the columns of \( Q \) are the corresponding eigenvectors. Let us define the change of variable \( x = Qy \), along with \( x_0 = Qy_0 \) and \( n = Qm \). The Jacobian of this transformation is \( \det Q = 1 \).

Before we apply the change of variable to (108), let us pause for a moment and consider the meaning of the condition \( n^T A^{-1} n \det A > 0 \). Since \( A \) is invertible and symmetric, all the eigenvalues of \( A \) are real and nonzero. If \( A \) is furthermore positive definite, then \( \det A = \prod \lambda_j > 0 \) and \( n^T A^{-1} n > 0 \) and thus (108) converges for any nonzero \( n \in \mathbb{R}^N \). We are more interested in the scenario where \( A \) has a single unstable mode. Let us denote this eigenvector as \( v_j \) and the corresponding negative eigenvalue as \( \lambda_j \). In this case, \( \det A < 0 \) and thus we require \( n^T A^{-1} n < 0 \); this implies that \( n \) should not be too close to being perpendicular to \( v_j \) (i.e., the integration should not be over a plane close to being parallel to the unstable mode). To see why, let us write \( n = \sum_j m_j v_j \) using the eigenvectors \( v_j \)’s in \( Q \) as the basis vectors. With this representation, \( n^T A^{-1} n = \sum_j \lambda_j^{-1} m_j^2 \); this sum is negative only if \( m_j \) is large enough such that \( |\lambda_j^{-1} m_j^2| > \sum_{j \neq j'} \lambda_j^{-1} m_j^2 \).
In the following analysis we will prove (108) in the case of a single unstable mode. For readers not interested in the algebraic details, the rest of the section can be skipped without loss of continuity.

Now, let us denote the integral in (108) as $I$. With the change of variable, we see that

$$I = \int \delta(\theta - m^TQ^TQy)e^{-\beta(y-y_0)^2} dy$$

(109)

$$= \frac{1}{2\pi} e^{-\beta y_0^2 \Lambda y_0/2} \int e^{ik(\theta - m^TQy)} e^{-\beta y^2 \Lambda y/2} e^{\beta y_0^2 \Lambda y} dy dk$$

(110)

$$= \frac{1}{2\pi} e^{-\beta y_0^2 \Lambda y_0/2} \int e^{ik\theta} \int e^{-\beta y^2 \Lambda y/2} e^{(\beta y_0 - ik\Lambda - m^TQy) \Lambda y} dy dk$$

(111)

$$= \frac{1}{2\pi} e^{-\beta y_0^2 \Lambda y_0/2} \int e^{ik\theta} \int e^{-\frac{\lambda_j^2}{2\beta} [\beta(y_j-y_0) + ik\lambda_j^{-1}m_j]} \frac{\lambda_j}{2\pi} \left[ik\lambda_j^{-1}m_j - \beta y_0\right]^2 dy dk$$

(112)

$$= \frac{1}{2\pi} e^{-\beta y_0^2 \Lambda y_0/2} \int e^{ik\theta} \int e^{-\frac{\lambda_j^2}{2\beta} [\beta(y_j-y_0) + ik\lambda_j^{-1}m_j]} \left(\int e^{-\frac{\lambda_j^2}{2\beta} [\beta(y_j-y_0) + ik\lambda_j^{-1}m_j]} dy dk\right)$$

(113)

The second equality follows by using the integral representation of the Dirac delta function as in (93).

At this point, we split the integrand of $\int dk$ in (113) into the product of two expressions. The first involves all $j$’s for which $\lambda_j > 0$, and the second consists of $j'$ for which $\lambda_j < 0$. The first group is

$$\prod_{j \neq j'} \int e^{-\frac{\lambda_j^2}{2\beta} [\beta(y_j-y_0) + ik\lambda_j^{-1}m_j]} dy_j = \prod_{j \neq j'} \int e^{-\frac{\lambda_j^2}{2\beta} [\beta(y_j-y_0) + ik\lambda_j^{-1}m_j]} \sqrt{\frac{2\pi}{\beta \lambda_j}}$$

(114)

$$= \left(\prod_{j \neq j'} \sqrt{\frac{2\pi}{\beta \lambda_j}}\right) e^{\frac{\beta}{2} \sum_{j \neq j'} \lambda_j y_{0j}^2 e^{\sum_{j \neq j'} (-k^2 \lambda_j^{-1}m_j^2/2\beta - ikm_jy_0)}}$$

While the second group is simply

$$\int e^{ik\theta - ikm_jy_j - \beta \lambda_j y_j^2/2 + \beta \lambda_j y_0j y_j} dy_j$$

(115)

The original integral in (108) can now be rewritten as

$$I = \frac{1}{2\pi} e^{-\frac{\beta}{2} y_0^2 \Lambda y_0} \left(\prod_{j \neq j'} \sqrt{\frac{2\pi}{\beta \lambda_j}}\right) e^{\frac{\beta}{2} \sum_{j \neq j'} \lambda_j y_{0j}^2} \int \int J(y_{j'}) dy_{j'} dk$$

(116)

$$= \frac{1}{2\pi} e^{-\frac{\beta}{2} \lambda_j y_j^2 (2\pi/\beta)^{-1} \left(\prod_{j \neq j'} \lambda_j^{-1/2}\right)} \int J(y_{j'}) \int J(k) dk$$

(117)

with

$$J(k) = e^{ik\theta - ikm_jy_j + \sum_{j \neq j'} (-k^2 \lambda_j^{-1}m_j^2/2\beta - ikm_jy_0j)}$$

(118)

$$J(y_{j'}) = e^{-\frac{\beta}{2} \lambda_j y_j^2 + \beta \lambda_j y_0j y_j}$$

(119)

To further simplify this expression, we first integrate over $k$, which is a Gaussian integral with a linear term,

$$\int J(k) dk = \left(\frac{2\pi \beta}{\sum_{j \neq j'} \lambda_j^{-1}m_j^2}\right)^{1/2} e^{-\frac{\beta}{2} \left(\frac{\sum_{j \neq j'} \lambda_j^{-1}m_j^2}{\sum_{j \neq j'} \lambda_j^{-1}m_j^2} - \theta\right)^2}$$

(120)
Then we integrate over $y_j'$, which is again a Gaussian integral with a linear term,

\[
\int I(y_j') \int I(k) \, dk \, dy_j' = \left( \frac{2\pi \beta}{\sum_{j \neq j'} \lambda_j^{-1} m_j^2} \right)^{1/2} \int e^{-\frac{2}{\beta} \lambda_j y_j'^2 + \beta \lambda_j y_0 y_j' - \frac{\beta}{2} \frac{\big( \Sigma_{j \neq j'} m_j y_0 j + m_j y_j' \theta \big)^2}{\Sigma_{j \neq j'} \lambda_j^{-1} m_j^2}} \, dy_j' \tag{121}
\]

\[
= \left( \frac{2\pi \beta}{\sum_{j \neq j'} \lambda_j^{-1} m_j^2} \cdot 2\pi \lambda_j^{-1} \frac{\sum_{j \neq j'} \lambda_j^{-1} m_j^2}{\beta} \right)^{1/2} e^{\frac{2}{\beta} \lambda_j y_0 y_j' e^{-\frac{\beta}{2} \frac{\big( \Sigma_{j \neq j'} m_j y_0 j + m_j y_j' \theta \big)^2}{\Sigma_{j \neq j'} \lambda_j^{-1} m_j^2}}} \tag{122}
\]

\[
= 2\pi \left( \lambda_j n^T A^{-1} n \right)^{-1/2} e^{\frac{2}{\beta} \lambda_j y_0 y_j' e^{-\beta (\theta - n^T x_0) / 2n^T A^{-1} n}} \tag{123}
\]

The last equality follows from the fact that $n^T A^{-1} n = m^T A^{-1} m = \sum_j \lambda_j^{-1} m_j^2$.

Taken together, we see that

\[
I = e^{-\frac{\beta}{2} \lambda_j y_0 y_j' (2\pi / \beta)^{-1}} \left( \prod_{j \neq j'} \lambda_j^{-1/2} \right) \left( \lambda_j n^T A^{-1} n \right)^{-1/2} e^{\frac{2}{\beta} \lambda_j y_0 y_j' e^{-\beta (\theta - n^T x_0) / 2n^T A^{-1} n}} \tag{124}
\]

\[
= (2\pi / \beta)^{-1} \left( \prod_{j} \lambda_j^{-1/2} \right) \left( n^T A^{-1} n \right)^{-1/2} e^{-\beta (\theta - n^T x_0) / 2n^T A^{-1} n} \tag{125}
\]

\[
= \frac{e^{-\beta (\theta - n^T x_0) / 2n^T A^{-1} n}}{\sqrt{(2\pi / \beta)^{1-N} (n^T A^{-1} n) \det A}} \tag{126}
\]

as desired.

### 4.2 Projection of Langer’s result to one dimension

Let us consider the same $N$-dimensional potential of mean force $W(r)$ as in Section 3. Here, we are interested in further projecting $W(r)$ to a one-dimensional potential of mean force. Let us denote this direction by a vector $n$. Integrating away the degrees of freedom orthogonal to $n$ is equivalent to performing a surface integral of the probability density $\pi(r)$ over a family of planes that are orthogonal to $n$. Let us parameterize this family of planes as $n^T r = \theta$ for $\theta \in \mathbb{R}$. After having worked through our derivation of Langer’s result, it should be obvious to the reader that the one-dimensional potential of mean force is

\[
e^{-\beta F(\theta)} = Z^{-1} \int \delta(\theta - n^T r) e^{-\beta W(r)} \, dr \tag{127}
\]

Adapting Kramers’ result in (24) to $F(\theta)$, the one-dimensional rate constant is

\[
k(n) = \left( \int_{\theta_A}^{\theta_B} \frac{e^{\beta F(\theta)}}{n^T Dn} \, d\theta \int_{-\infty}^{\theta_B} e^{-\beta F(\theta)} \, d\theta \right)^{-1} \tag{128}
\]

where $\theta_A$ and $\theta_B$ are the positions of free energy minima at the reactant and product well, respectively, and $\theta^*$ is the position of the saddle point in the one-dimensional projection. The expression $n^T Dn$ is the one-dimensional diffusion constant along $n$. To understand this expression, first note that since the diffusion matrix represents a physical property of the system, the representation of $D$ should change with coordinate transformations in such a way as to leave the underlying physics invariant. This makes $D$ a tensor, specifically a second-order contravariant tensor. Under the transformation $r \mapsto \theta$, the tensor $D_{ij}$ transforms correspondingly to a zeroth-order tensor (i.e., a constant) by

\[
\sum_{ij} D_{ij} \frac{\partial \theta}{\partial r_i} \frac{\partial \theta}{\partial r_j} = \sum_{ij} n_i D_{ij} n_j = n^T Dn \tag{129}
\]
as desired.

Using (108) and the harmonic approximations at the saddle point (40), the first integral in (128) evaluates to

$$\int_{\theta_A}^{\theta_B} e^{\beta F(\theta)} \frac{e^{\beta \Delta W^\dagger}}{n^T Dn} d\theta \approx Z e^{\beta \Delta W^\dagger} \int \left( \int \delta(\theta - n^T r) e^{-\beta (r-r^\dagger)^T H^H (r-r^\dagger) / 2} dr \right)^{-1} d\theta$$

$$= Z e^{\beta \Delta W^\dagger} \sqrt{(2\pi / \beta)^{1-N} (n^T H^{-4} n) \det H^H \sqrt{(2\pi / \beta) |n^T H^{-4} n|}}$$

$$= (2\pi / \beta)^{1-N/2} e^{\beta \Delta W^\dagger} |n^T H^{-4} n|^{1/2} \det H^H$$

Note that since $H^H < 0$, the use of (108) requires that $n^T H^{-4} n < 0$; otherwise the integral diverges as we discussed in the previous section. Similarly, with the harmonic approximation at the reactant well (39), the second integral in (128) evaluates to

$$\int_{-\infty}^{\theta_1} e^{-\beta F(\theta)} d\theta \approx Z^{-1} \int \int \delta(\theta - n^T r) e^{-\beta (r-r^\dagger)^T H_A (r-r^\dagger) / 2} dr d\theta$$

$$= Z^{-1} \sqrt{(2\pi / \beta)^{1-N} (n^T H^{-4} n) \det H_A (131)}$$

$$= Z^{-1} (2\pi / \beta)^{N/2} (\det H_A)^{-1/2}$$

The application of (108) here does not impose any further conditions on $n$, since $H_A$ is symmetric positive definite. Together, the rate constant is

$$k(n) = \frac{\beta}{2\pi} \sqrt{\det H_A \frac{n^T Dn}{|n^T H^{-4} n|} e^{-\beta \Delta W^\dagger}}$$

(136)

The one-dimensional rate constant in (136) is reminiscent of Langer’s result in (106), except with $\lambda_+$ replaced by $n^T Dn / |n^T H^{-4} n|$. In fact, $n^T Dn / |n^T H^{-4} n| = \lambda_+$ when $n$ is parallel to $v_+$ (this should be clear from the fact that $H^{-4} v_+ = -\lambda_+ D v_+$). Furthermore $k(v_+)$ is the minimum of $k(n)$ (more strictly speaking, any vector proportional to $v_+$ will do). To see why, recall the earlier discussion related to the simultaneous diagonalization of $H^H$ and $D$ by $V$ in (60). With a change of variable $n = Vm$,

$$\frac{n^T Dn}{|n^T H^{-4} n|} = \frac{m^T V^T D V m}{|m^T V^T H^{-4} V m|} = \frac{m^T m}{|m^T \Lambda^{-1} m|} = \frac{\|m\|^2}{|m^T \Lambda^{-1} m|}$$

(137)

The expression $m^T \Lambda^{-1} m / m^T m$ is an example of a Rayleigh quotient. In the absence of any constraint, the quotient is bounded between the largest and smallest eigenvalues of $\Lambda^{-1}$. This bound is unfortunately not helpful, since we require that $n^T H^{-4} n$ (and thus $m^T \Lambda^{-1} m$) be negative. For any fixed length $\|m\|$, the expression $| - \lambda_+^{-1} m_+^2 + \sum_{j \neq +} \lambda_j^{-1} m_j^2 |$ in the denominator is maximized under this constraint whenever the positive terms in the sum $\sum_{j \neq +} \lambda_j^{-1} m_j^2$ are minimized. This is achieved by setting all $m_j = 0$ except for $|m_+| = \|m\|$; in other words, $m$ is proportional to $e_+$, and thus $n$ is proportional to $v_+$ (the change of variable described here is the same as that in Section 3.4.3, but one should be careful not to confuse the vectors $n$ and $m$ defined in this section with those defined in Section 3.4.3). Taken together, we have shown that $n = v_+$ minimizes (136) and the minimum is equivalent to (106), as desired.

5 Acknowledgement

I would like to thank Robert Alberstein for his critical reading of the note, and I would like to acknowledge support through Chan Zuckerberg Biohub Investigator funds to Dr. Tanja Kortemme (UCSF).
References

[1] Langer JS (1969) Statistical theory of the decay of metastable states. *Annals of Physics* 54(2):258--275.

[2] Kramers HA (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7(4):284--304.

[3] Berezhkovskii A, Szabo A (2004) One-dimensional reaction coordinates for diffusive activated rate processes in many dimensions. *The Journal of Chemical Physics* 122(1):014503.

[4] Hänggi P, Talkner P, Borkovec M (1990) Reaction-rate theory: Fifty years after Kramers. *Reviews of Modern Physics* 62(2):251--341.

[5] Zhou HX (2010) Rate theories for biologists. *Quarterly Reviews of Biophysics* 43(2):219--293.

[6] Peters B (2016) Reaction Coordinates and Mechanistic Hypothesis Tests. *Annual Review of Physical Chemistry* 67(1):669--690.

[7] Peters B (2017) *Reaction Rate Theory and Rare Events*. (Elsevier), 1st edition.

[8] E W, Vanden-Eijnden E (2010) Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events. *Annual Review of Physical Chemistry* 61(1):391--420.

[9] Tuckerman ME (2010) *Statistical Mechanics: Theory and Molecular Simulation*. (Oxford University Press, Oxford ; New York), Illustrated edition.

[10] Vanden-Eijnden E, Tal FA (2005) Transition state theory: Variational formulation, dynamical corrections, and error estimates. *The Journal of Chemical Physics* 123(18):184103.

[11] Reimann P, Schmid GJ, Hänggi P (1999) Universal equivalence of mean first-passage time and Kramers rate. *Physical Review E* 60(1):R1--R4.

[12] Metzner P, Schütte C, Vanden-Eijnden E (2006) Illustration of transition path theory on a collection of simple examples. *The Journal of Chemical Physics* 125(8):084110.

[13] Dellago C, Bolhuis P, Geissler P (2006) Transition Path Sampling Methods in *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*, eds. Ferrario M, Ciccotti G, Binder K. (Springer, Berlin, Heidelberg), pp. 349--391.

[14] Dellago C, Hummer G (2014) Computing Equilibrium Free Energies Using Non-Equilibrium Molecular Dynamics. *Entropy* 16(1):41--61.

[15] Im W, Roux B (2002) Ions and Counterions in a Biological Channel: A Molecular Dynamics Simulation of OmpF Porin from Escherichia coli in an Explicit Membrane with 1M KCl Aqueous Salt Solution. *Journal of Molecular Biology* 319(5):1177--1197.

[16] Liu P, Harder E, Berne BJ (2004) On the Calculation of Diffusion Coefficients in Confined Fluids and Interfaces with an Application to the Liquid-Vapor Interface of Water. *The Journal of Physical Chemistry B* 108(21):6595--6602.

[17] Hummer G (2005) Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *New Journal of Physics* 7(1):34.

[18] Ma A, Nag A, Dinner AR (2006) Dynamic coupling between coordinates in a model for biomolecular isomerization. *The Journal of Chemical Physics* 124(14):144911.

[19] Peters B (2009) Competing nucleation pathways in a mixture of oppositely charged colloids: Out-of-equilibrium nucleation revisited. *The Journal of Chemical Physics* 131(24):244103.
[20] Berezhkovskii AM, Zitserman VY (1989) Anomalous regime for decay of the metastable state: An extension of multidimensional Kramers theory. *Chemical Physics Letters* 158(5):369–374.

[21] Berezhkovskii AM, Szabo A, Greives N, Zhou HX (2014) Multidimensional reaction rate theory with anisotropic diffusion. *The Journal of Chemical Physics* 141(20):204106.

[22] Northrup SH, McCammon JA (1983) Saddle-point avoidance in diffusional reactions. *The Journal of Chemical Physics* 78(2):987–989.

[23] Klosek-Dygas MM, et al. (1989) Diffusion theory of multidimensional activated rate processes: The role of anisotropy. *The Journal of Chemical Physics* 90(2):1141–1148.

[24] Chen W, Ferguson AL (2018) Molecular enhanced sampling with autoencoders: On-the-fly collective variable discovery and accelerated free energy landscape exploration. *Journal of Computational Chemistry* 39(25):2079–2102.

[25] Ribeiro JML, Bravo P, Wang Y, Tiwary P (2018) Reweighted autoencoded variational Bayes for enhanced sampling (RAVE). *The Journal of Chemical Physics* 149(7):072301.

[26] Grote RF, Hynes JT (1980) The stable states picture of chemical reactions. II. Rate constants for condensed and gas phase reaction models. *The Journal of Chemical Physics* 73(6):2715–2732.

[27] Berezhkovskii AM, Pollak E, Zitserman VY (1992) Activated rate processes: Generalization of the Kramers–Grote–Hynes and Langer theories. *The Journal of Chemical Physics* 97(4):2422–2437.

[28] Metzner P, Schütte C, Vanden-Eijnden E (2009) Transition Path Theory for Markov Jump Processes. *Multiscale Modeling & Simulation* 7(3):1192–1219.

[29] Bello-Rivas JM, Elber R (2015) Exact milestoning. *The Journal of Chemical Physics* 142(9):094102.

[30] Zuckerman DM, Chong LT (2017) Weighted Ensemble Simulation: Review of Methodology, Applications, and Software. *Annual Review of Biophysics* 46(1):43–57.

[31] Dellago C, Bolhuis PG, Csaika FS, Chandler D (1998) Transition path sampling and the calculation of rate constants. *The Journal of Chemical Physics* 108(5):1964–1977.

[32] van Erp TS, Bolhuis PG (2005) Elaborating transition interface sampling methods. *Journal of Computational Physics* 205(1):157–181.