1. An Introduction to the Large Deviation Theory

The Large Deviation Theory was largely developed in 1960s-1970s by S. R. Srinivasa Varadhan, (currently Professor in New York University, Department of Mathematics) in the United States and by Mark Freidlin (also click here) (currently Professor in the University of Maryland, Department of Mathematics) and Alexander Wentzell (currently Professor in Tulane University, Department of Mathematics) in the Soviet Union.

A major reference for the Large Deviation Theory is the book by Freidlin and Wentzel [10]. Its first edition appeared in Russian in the late 1970s, its most recent edition is the 3rd edition that appeared in 2013.

In these notes, for simplicity, we will focus on the stochastic differential equation (SDE)

\[ dx = b(x)dt + \sqrt{2\beta^{-1}}dw, \quad x \in \mathbb{R}^d \]  

(1)

where \( b(x) \) is a continuously differentiable vector field with a finite number of isolated equilibria, \( \beta^{-1} \) is a small parameter, and \( w \) is the standard Brownian motion.
1.1. **Long-time behavior of solutions of** \( \dot{x} = b(x) \). Together with Eq. (1) we also consider the unperturbed deterministic equation

\[
\frac{dx}{dt} = b(x).
\]  

We assume the vector field \( b(x) \) is such that for any initial condition \( x(0) = x_0 \), the characteristic (a.k.a. the phase trajectory) of Eq. (2) starting at \( x_0 \), will remain in some bounded region. In other words, we will consider only such vector fields \( b(x) \) that no characteristic goes to infinity at \( t \to \infty \).

The solutions of the algebraic equation \( b(x) = 0 \) are called equilibria or equilibrium points, or stationary points. An equilibrium \( x^* \) is called stable if for any its neighborhood \( U(x^*) \) one can find a neighborhood \( V(x^*) \subset U(x^*) \) such that any characteristic starting in \( V(x^*) \) will remain in \( U(x^*) \). Otherwise, the equilibrium is called unstable. If \( x^* \) is a stable equilibrium, and every characteristic \( x(t) \) starting in \( V(x^*) \) approaches \( x^* \) as \( t \to \infty \), then the equilibrium \( x^* \) is called asymptotically stable.

We assume that the equation \( b(x) = 0 \) has any a finite number of solutions. This implies, in particular, that all equilibria are isolated.

The long-time behavior of Eq. (2) is simple in 1D. For every initial condition \( x(0) = x_0 \), the characteristic \( x(t) \) tends to some asymptotically stable equilibrium. All equilibrium points are solutions of the algebraic equation \( b(x) = 0 \).

Now we discuss the long-time behavior of a characteristic in higher dimensions staying in a bounded region \( D \).

**Definition 1.** Let \( x(t) \) be a characteristic of Eq. (2) staying in a bounded region \( D \). A point \( x^* \) is a limit point of \( x(t) \) if there exists a sequence of moments of time \( \{t_n\} \to \infty \) such that

\[
x(t_n) \to x^* \text{ as } n \to \infty.
\]

The set of all limit points of a characteristic \( x(t) \) is called the \( \omega \)-limit set of \( x(t) \) or attractor.

The long-time behavior of Eq. (2) in 2D is more complicated [15]. Attracting sets are always connected and compact. Under the conditions of the vector field \( b \) mentioned above, there are three kinds of possible \( \omega \)-limit sets in 2D:

1. an equilibrium point;
2. a limit cycle (a closed characteristic corresponding to a periodic solution of Eq. (2));
3. a set consisting of equilibria and characteristics approaching equilibria at \( t \to \pm \infty \).

The long-time behavior of a characteristic in 3D can be even more complicated. Besides the attractors possible in 2D and the invariant tori, there can be also strange attractors. They are encountered, e.g., in the Lorenz system (the Lorenz attractor), Roessler system (Roessler attractor), etc. If you are interested, there is plenty of literature on this subject.
1.2. Freidlin-Wentzell Action Functional. The Freidlin-Wentzell action functional $S_T(\phi)$ is defined on all absolutely continuous paths connecting the two given points by

$$
S_T(\phi) = \frac{1}{2} \int_0^T (\dot{\phi} - b(\phi(t)))^2 dt,
$$

where $\dot{\phi} \equiv \frac{d\phi}{dt}$. If the path $\phi$ is a solution of $\dot{x} = b(x)$ then the Freidlin-Wentzel action along this path is zero. Otherwise it is positive. Let us give an insight of where the Freidlin-Wentzell action comes from.

Consider a solution $x(t)$ of Eq. (1). Let us calculate the probability that $x(t)$ nearly follows a given path $\phi(t)$. Motivated by the fact that

$$
dx - b(x)dt = \sqrt{2\beta^{-1}}dw
$$

we define another path $\xi(t)$ that should be followed by the Brownian motion:

$$
d\xi(t) = (\dot{\phi} - b(\phi)) dt, \quad x(t) = \int_0^t (\dot{\phi} - b(\phi)) dt.
$$

If the solution $x(t)$ stays close to $\phi(t)$, then the scaled Brownian motion $\sqrt{2\beta^{-1}}$ should stay close to the path $\xi(t)$. Now we outline the calculation of the probability that $j$-th component of the scaled Brownian motion $\sqrt{2\beta^{-1}}$ stays in the $\epsilon$-tube surrounding the $j$-th component of the path $\xi(t)$. This can be done using the Wiener measure as follows. We discretize the time interval $[0, T]$

$$
0 < t_1 < t_2 < \ldots < t_n = T
$$

so that the length of each subinterval is $h$. Recall that $\sqrt{2\beta^{-1}}(w(t) + h) - w(t)) \sim N(0, 2\beta^{-1}h)$. Then the probability that each component $\sqrt{2\beta^{-1}}w_j(t)$ stays within distance $\epsilon$ of $\xi_j(t)$ is given by

$$
P(\max_{1 \leq j \leq d} \max_{0 \leq t \leq T} |\sqrt{2\beta^{-1}}w_j(t) - \xi_j(t)| < \epsilon)
= \prod_{j=1}^n \lim_{n \to \infty} \int_{\xi_j(t_1) - \epsilon}^{\xi_j(t_1) + \epsilon} e^{-(y_1 - \xi_j(0))^2/4\beta^{-1}h} dy_1 \int_{\xi_j(t_2) - \epsilon}^{\xi_j(t_2) + \epsilon} e^{-(y_2 - \xi_j(1))^2/4\beta^{-1}h} dy_2 
\ldots \int_{\xi_j(t_n) - \epsilon}^{\xi_j(t_n) + \epsilon} e^{-(y_n - \xi_j(t_{n-1}))^2/4\beta^{-1}h} dy_n.
$$

(4)

For any $1 \leq k \leq n$,

$$
\frac{(y_k - y_{k-1})^2}{4\beta^{-1}h} = \frac{(\xi_j(t_k) - \xi_j(t_{k-1}) + u_k)^2}{4\beta^{-1}h},
$$

where $u_k \sim N(-\xi_j(t_k) - \xi_j(t_{k-1}), 4\beta^{-1}h)$. Note that

$$
\frac{(\xi_j(t_k) - \xi_j(t_{k-1})^2}{4\beta^{-1}h} = \frac{(\dot{\phi}_j - b_j(\phi(t)))^2}{4\beta^{-1}h} = \frac{\beta}{4}(\dot{\phi}_j - b_j(\phi(t)))^2 h.
$$
Pulling these terms outside each sub-integral, carefully taking the limits: first \( n \to \infty \), then \( \beta^{-1} \to 0 \), and finally \( \epsilon \to 0 \), one obtains that

\[
\lim_{\epsilon \to 0} P(|w_t - \xi(t)| < \epsilon) \times \exp\left\{ -\frac{\beta}{4} \int_0^T |\dot{\phi} - b(\phi(t))|^2 \, dt \right\} \approx \exp\left\{ -\frac{\beta}{2} S_T(\phi) \right\},
\]

where \( S_T \) is the Freidlin-Wentzel action Eq. (3).

1.3. Quasipotential. Now we consider the problem of exiting from a region \( D \) surrounding an attractor of Eq. (2). The expected exit times and the transition rates are defined in terms of the quasipotential. We start with the simplest case where the attractor is an asymptotically stable equilibrium \( x_0 \) of \( \dot{x} = b(x) \). The quasipotential \( U_{x_0}(x) \) at point \( x \) w.r.t. the point \( x_0 \) is defined as the infimum of the Freidlin-Wentzell action functional over all possible times \( T \) and paths \( \phi \) connecting the points \( x_0 \) and \( x \):

\[
(5) \quad U_{x_0}(x) = \inf_{\phi, T} \{ S_T(\phi) \mid \phi(0) = x_0, \phi(T) = x \}.
\]

The expected exit time \( \mathbb{E}[\tau_{exit}] \) from the basin of attraction \( B(x_0) \) of the point \( x_0 \) is logarithmically equivalent to

\[
(6) \quad \mathbb{E}[\tau_{exit}] \approx e^{\frac{\beta}{2} \min_{x \in \partial B(x_0)} U_{x_0}(x)}.
\]

In this case, according to the Large Deviation theory [10], the system also leaves the basin of attraction \( B(x_0) \) of the point \( x_0 \) along a trajectory lying near a predictable path. This path is the minimizer of the Freidlin-Wentzel action given by Eq. (3) over all paths from the point \( x_0 \) to the boundary of its basin of attraction and all times \( T \). Such a path is called the Minimum Action Path.

1.4. The geometric action. We will start with eliminating time in the Freidlin-Wentzell action. Let a point \( x_0 \) be an asymptotically stable equilibrium of \( \dot{x} = b(x) \), and a point \( x \) belong to its basin of attraction. Suppose we are interested only in minimizers of the Freidlin-Wentzell action functional given by Eq. (3) with \( \phi(0) = x_0 \) and \( \phi(T) = x \) w.r.t. both the time and the paths. Then the Freidlin-Wentzell action functional can be replaced with the geometric action functional [10, 13]

\[
(7) \quad S(\psi) = \int_0^{\alpha_{\text{max}}} \{ \|b(\psi(\alpha))\| \psi_\alpha(\alpha) \| - (b(\psi(\alpha)) \cdot \psi_\alpha) \} \, d\alpha.
\]

These two actions relate as follows. Let a path \( \phi(t) \) minimize the functional (3) over all times \( T \in [0, +\infty] \) and all paths from the point \( x_0 \) to the point \( x \). Then any path \( \psi(\alpha) \) from \( x_0 \) to \( x \) that can be obtained from \( \phi(t) \) by reparametrization, minimizes the geometric
action functional (7). One can see this from the following argument [13]:

\[ S_T(\phi) = \frac{1}{2} \int_0^T |\phi - b(\phi(t))|^2 dt = \frac{1}{2} \int_0^T \left\{ |\dot{\phi}|^2 + |b(\phi(t))|^2 - 2\dot{\phi} \cdot b(\phi) \right\} dt \]

\[ \geq \int_0^T \left\{ |\dot{\phi}| b(\phi) - \dot{\phi} \cdot b(\phi) \right\} dt \]

(8)

\[ = \int_0^{\alpha_{\text{max}}} \{ |b(\psi(\alpha))| |\psi_\alpha(\alpha)| - \psi(\alpha) \cdot \psi_\alpha \} d\alpha. \]

The equality in Eq. (8) is achieved if and only if $|\dot{\phi}| = |b(\phi)|$. It follows from Eq. (8) that

\[ \inf_{T, \phi} S_T(\phi) \geq S(\psi). \]

On the other hand, if a path $\psi(\alpha)$ is the global minimizer of the geometric action (7), we can reparametrize it so that $|\dot{\phi}| = |b(\phi)|$ that makes the inequality in Eq. (8) into an equality. Strictly speaking, the global minimizer of the Freidlin-Wentzell action (3) with at least one point at a zero of $b(x)$ does not exist, as the global minimum is achieved at $T = \infty$, however, the minimizing curve $\psi(\alpha)$ between the asymptotically stable equilibrium $x_0$ and any point $x$ in its basin of attraction does exist. This is what makes the geometric action (7) attractive to work with.

We note that the value of integral in Eq. (7) is independent of the choice of parametrization of the path $\psi$. We mostly will use either the arclength parametrization or the uniform parametrization on the interval $[0, 1]$. Then the geometric action becomes

(9)

\[ S(\psi) = \int_0^L \{ |b(\psi(s))| |\psi_s| - b(\psi(s)) \cdot \psi_s \} ds \]

for the arclength parametrization, where $L$ is the length of the path, and

(10)

\[ S(\psi) = \int_0^1 \{ |b(\psi(\alpha))| |\psi_\alpha| - b(\psi(\alpha)) \cdot \psi_\alpha \} d\alpha \]

for the uniform parametrization on the interval $[0, 1]$.

1.5. The case of the overdamped Langevin dynamics. Suppose the SDE (1) is gradient, i.e., $b(x) = -\nabla V(x)$ for some function $V(x)$ and hence Eq. (1) describes the overdamped Langevin dynamics. Then the minimum and the minimizer of the geometric action have a simple characterization [13]. Suppose a path $\psi(s)$ connects two local minima $x_A$ and $x_B$ of $V(x)$ separated by a single saddle $x_{AB}$. Consider first the part of the path going uphill and set

$\psi(0) = x_A, \quad \psi(1) = x_{AB}$. 
Start from Eq. (10) and replace $b(c)$ with $-\nabla V(x)$ in it:

\[
S(\psi) = \int_0^1 \left\{ |\nabla V(\psi(\alpha)||\psi_\alpha| + \nabla V(\psi(\alpha)) \cdot \psi_\alpha \right\} d\alpha \\
= \int_0^1 \left\{ |\nabla V(\psi(\alpha)||\psi_\alpha| + \frac{dV(\alpha)}{d\alpha} \right\} d\alpha \\
= \int_0^1 \left\{ |\nabla V(\psi(\alpha)||\psi_\alpha| \right\} d\alpha + V(x_{AB}) - V(x_A).
\]

Note that

\[|\nabla V(\psi(\alpha)||\psi_\alpha| \geq \nabla V(\phi) \cdot \phi_\alpha,\]

and the equality takes place if and only if $\phi_\alpha$ is parallel to $\nabla V(\phi(\alpha))$. Hence

\[S(\psi) \geq \int_0^1 \nabla V(\psi(\alpha)) \cdot \psi_\alpha d\alpha + V(x_{AB}) - V(x_A) = 2(V(x_{AB}) - V(x_A)).\]

Therefore, in the gradient case the quasipotential with respect to a minimum $x_A$ within its basin of attraction is doubled potential. Note that traveling along the steepest descent path $\dot{\phi} = -\nabla V(\phi)$ from the saddle $x_{AB}$ to the minimum does not contribute to the geometric action.

The result that $\min_\psi S(\psi) = 2(V(x_{AB}) - V(x_A))$ where the minimum is taken over all paths starting at the local minimum $x_A$ and ending at the local minimum $x_B$, is consistent with the Arrhenius law. Indeed, the transition rate from $x_A$ to $x_B$, which is the reciprocal of the expected exit time from the basin of attraction of $x_A$ is

\[r_{AB} \propto \exp\left\{-\frac{\beta}{2}2(V(x_{AB}) - V(x_A)) \right\} = \exp\{-\beta(V(x_{AB}) - V(x_A))\} \] .

The path $\psi$ that minimizes the geometric action is the one satisfying the condition $\frac{d\psi}{d\alpha}$ is parallel to $\nabla V(\psi)$, i.e., the velocity vector of the path is parallel to the gradient of the vector field. Hence, the path $\psi$ goes either directly uphill or directly downhill. A path $\psi$ going directly uphill or directly downhill is called the Minimum Energy Path or MEP. The collection of MEPs connecting every pair of neighboring minima in the 7-well potential is shown in Fig. 1.

1.6. Quasipotential with respect to compact sets. In order to allow more complex attractors than asymptotically stable equilibrium points, we extend the definition of the quasipotential to a function defined on pairs of compact sets in $\mathbb{R}^n$.

**Definition 2.** For any pair of compact sets $X_0 \subset \mathbb{R}^n$ and $X \subset \mathbb{R}^n$ we define the quasipotential as

\[U(X_0, X) = \inf \{ S(\psi) \mid \psi(0) \in X_0, \psi(1) \in X \}.\]
Figure 1. The Minimum Energy Paths (MEPs) connecting every pair of neighboring minima in the 7-well potential.

It follows from Eqs. (9) and (11) that the quasipotential is Lipschitz-continuous function with respect to both of its arguments on every compact set $D$, and the Lipschitz constant is $2 \max_D |b|$. If $X_0$ and $X$ are points then

$$|U(X_0, X) - Y(Y_0, Y)| \leq 2 \max_D |b|(|X_0 - Y_0| + |X - Y|).$$

If $X_0$ and $X$ are compact sets, the same is true with respect to the Hausdorff distance.

An important property of the quasipotential is that it is constant on every attractor of the system. This property is summarized in the following theorem [2]

**Theorem 1.** Let $X_0$ be a compact set. Let $C = \{z(t) | t \geq 0\}$ be a trajectory of $\dot{x} = b(x)$, and $F$ be its $\omega$-limit set. Then

1. $U(X_0, y) = U(X_0, F)$ for all $y \in F$, i.e., the quasipotential $U(X_0, y)$ is constant on the $\omega$-limit set $F$ and equals to $U(X_0, F)$.
2. $U(y, X_0) = U(F, X_0)$ for all $y \in F$, i.e., the quasipotential $U(y, X_0)$ is constant on the $\omega$-limit set $F$ and equals to $U(F, X_0)$.

1.7. **The Hamilton-Jacobi-Bellman equation for the quasipotential.** In this section we derive the Hamilton-Jacobi-Bellman equation from the Bellman Principle of Optimality [1]. This equation will provide us with the relationship between the velocity vector $\psi_s$ of the Minimum Action Path and the gradient of the quasipotential that will enable us to compute the Minimum Action Path once we have found the quasipotential.
Let us fix some compact set $X_0$ and consider the quasipotential $U(x) \equiv U(X_0, x)$ as a function of the point $x$. Then the function $U(x)$ is given by

$$
U(x) = \left\{ \inf_{\psi} \int_0^L \{|b(\psi(s)\|\psi_s| - b(\psi(s)) \cdot \psi_s\} ds \mid \psi(0) \in X_0, \psi(L) = x \right\}.
$$

**Bellman’s Principle of Optimality:** An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

We will treat the quasipotential $U(x)$ as the value function, the integrand in Eq. (12) as the cost function, and unit velocity vector $\psi_s$ as the control. The Bellman optimality principle reads

$$
U(x) = \inf_{\psi_s} \left\{ \int_0^l |\psi(s)| - b(\psi(s)) \cdot \psi_s ds + U \left( x - \int_0^l \psi_s(s) ds \right) \right\}.
$$

We fix a small $\epsilon > 0$. Then Eq. (13) becomes

$$
U(x) = \inf_{\psi_s \in S^{n-1}} \left\{ \{|b(\psi)| - b(\psi) \cdot \psi_s\} + U(x) - (\nabla U(x) \cdot \psi_s)\epsilon + O(\epsilon^2) \right\},
$$

where $S^{n-1}$ is the unit sphere in $\mathbb{R}^n$. Canceling $U(x)$, dividing by $\epsilon$ and taking the limit $\epsilon \to 0$ we obtain the static Hamilton-Jacobi-Bellman equation

$$
H(x, \nabla U) := \inf_{\psi_s \in S^{n-1}} \left\{|b(x)| - (b(x) + \nabla U(x)) \cdot \psi_s\right\} = 0.
$$

The control $\psi_s \in S^{n-1}$ that gives the infimum in Eq. (15) can be found explicitly. We observe that $-|b + \nabla U| \leq (b + \nabla U) \cdot \psi_s \leq |b + \nabla U|$. Therefore, the infimum is achieved if

$$
\psi_s = \frac{b + \nabla U}{|b + \nabla U|}.
$$

Furthermore, this infimum equals zero if and only if $|b + \nabla U| = |b|$. Hence the optimal control $\psi_s$ and the gradient of the value function $\nabla U$ are related via

$$
\psi_s = \frac{b + \nabla U}{|b|}, \quad \nabla U = |b|\psi_s - b.
$$

The last equation shows that $\nabla U(x)$ belongs to the sphere of radius $|b(x)|$ centered at $b(x)$ as shown in Fig. 2, and the direction of the Minimum Action Path $\psi_s$ at a point $x$ is collinear to $b(x) + \nabla U(x)$. Using Eq. (16) the Hamilton-Jacobi-Bellman equation (15) can be rewritten as

$$
H(x, \nabla U) = |\nabla U|^2 + 2\nabla U \cdot b(x) = 0, \quad U|_{X_0} = 0.
$$

Now the question is what is the relationship between the quasipotential defined as the solution of the minimization problem (12) and solutions of Eq. (17). By construction, the quasipotential $U(x)$ is a weak solution of Eq. (17). On the other hand, Eq. (17) can have infinitely many weak solutions and several classic solutions, one of which is $U(x) \equiv 0$. Moreover, the concept of the viscosity solution cannot help to eliminate the solutions of
Figure 2. The relationship between the force field \( b(x) \), the gradient of the quasipotential \( \nabla U(x) \), and the direction of the Minimum Action path \( \psi_s \parallel b + \nabla U \).

(17) that are not the quasipotential, i.e., the solution of Eq. (12). We will illustrate it on the following 1D example.

**Example 1** Suppose \( x \in \mathbb{R}^1 \). Then \( b(x) = -V'(x) \) for some \( V(x) \).

Let Therefore, Eq. (17) becomes \((U'')^2 - 2U'V' = 0\). Any continuously differentiable \( U(x) \) that satisfies \( U' = 0 \) or \( U' = 2V' \) is its classic solution. For any non-constant \( V(x) \) one immediately sees two solutions: \( U(x) \equiv 0 \) and \( U(x) = 2V(x) - V(x_0) \). If \( V' \) has several zeros, one can obtain several continuously differentiable solutions by gluing pieces with \( U' = 0 \) and \( U' = 2V' \) at the points where \( V' = 0 \) as shown in Fig. 3. Totally there are \( 2^N \) continuously differentiable solutions where \( N \) is the number of zeros of \( V' \).

A solution obtained by gluing pieces with \( U' = 0 \) and \( U' = 2V' \) is \( C^p \) if \( V \) is \( C^p \) and its first \( p \) derivatives are zero at the points of gluing. Each of these continuously differentiable solutions is a viscosity solution. Therefore, the concept of viscosity solution may not eliminate the solutions that are not the quasipotential.

The Hamilton-Jacobi-Bellman equation (17) implies that the vector field \( b(x) \) can be decomposed into the potential component and the orthogonal to it rotational component wherever \( U(x) \) is continuously differentiable as

\[
(18) \quad b(x) = -\frac{1}{2} \nabla U(x) + \left( \frac{1}{2} \nabla U(x) + b(x) \right) \equiv -\frac{1}{2} \nabla U(x) + l(x).
\]

The orthogonality of \( \nabla U(x) \) and \( l(x) \) can be shown from Eq. (17) by rewriting it as

\[
0 = 2\nabla U \cdot \left( \frac{1}{2} \nabla U + b(x) \right) = 2\nabla U \cdot l(x).
\]

Eq. (16) implies that the Minimal Action paths are the flowlines of the field \( \frac{1}{2} \nabla U + l(x) \), i.e., the Minimum Action Paths follow the rotational component of the field but go against its potential component as in an example provided in [10] (Chapter 4, Theorem 3.1, p. 118).
2. Analytical examples

In this section, we explore the behavior of the quasipotential in the 2D case \[2\]. As it is well-known, there are only three kinds of the \(\omega\)-limit sets of trajectories \(x(t), t \geq 0\) of \(\dot{x} = b(x)\) confined in a compact sets sets in 2D. They are (i) the asymptotically stable equilibrium point, (ii) the limit cycle, and (iii) a set consisting of a finite number of equilibria and heteroclinic characteristics connecting them.

2.1. Periodic trajectories. In this section we show that if an equilibrium is such that all trajectories in some its neighborhood \(D\) are periodic then the quasipotential \(U(x, y) = 0\) for any \(x, y \in D\). The Minimum Action Path between the points \(x\) and \(y\) does not exist.

**Theorem 2.** Let \(x_0\) be an equilibrium of \(\dot{x} = b(x), x \in \mathbb{R}^2\) and all of trajectories in some neighborhood \(D\) of \(x_0\) containing no other critical points are periodic. Then \(U(x, y) = 0\) for all \(x \in D, y \in D\).

**Proof.** Let \(C_1\) and \(C_2\) be two close periodic trajectories lying in the domain \(D\). \(C_1\) lies inside the region bounded by \(C_2\).

We fix some \(\epsilon > 0\) and consider a path \(\psi(s)\) starting at some point of the trajectory \(C_1\) and ending at some point of trajectory \(C_2\) whose velocity vector is given by (see Fig. 4)

\[
\psi_s = \frac{b(\psi) + \epsilon b^\perp(\psi)}{|b + \epsilon b^\perp|}.
\]

Here \(b^\perp(x)\) is a vector field orthogonal to the vector field \(b(x)\) such that \(|b^\perp| = |b|\), and \(b^\perp\) at the trajectory \(C_1\) is directed toward the trajectory \(C_2\).

Then we consider the geometric action \(S(\psi)\) for this path. Using the orthogonality of \(b\) and \(b^\perp\) we have

\[
b \cdot \psi_s = \frac{|b|^2 + \epsilon (b \cdot b^\perp)}{|b| \sqrt{1 + \epsilon^2}} = |b| \left(1 - \frac{1}{2} \epsilon^2 + O(\epsilon^4)\right).
\]

The length of the path \(\psi\) can be bounded from above as follows. Let us consider an orthogonal coordinate system \((l, y)\) in the neighborhood of \(x_0\) whose coordinate curves...
Figure 4. Illustration for Theorem 2. Two closed trajectories $C_1$ and $C_2$ and a path $\psi$ connecting them.

$y = \text{const}$ are the periodic trajectories of $\dot{x} = b(x)$ and hence the orthogonal curves are the flowlines of $\dot{x} = b^\perp(x)$. Let $y_1$ and $y_2$ be the $y$-coordinates of the curves $C_1$ and $C_2$. Let $\gamma$ be the Euclidean length along each curve $l = \text{const}$. At every point of the path $\psi$ the component of the velocity in the direction of $y$ is

$$\frac{\partial \psi}{\partial \gamma} = \frac{\epsilon b^\perp}{\sqrt{|b^2| + \epsilon^2 |b^\perp|^2}} = \frac{\epsilon}{\sqrt{1 + \epsilon^2}}$$

Since the path $\psi$ ends as soon as it reaches $C_2$, its length $L$ is bounded by

$$L \equiv \int_0^L |\psi_s| ds = \int_{y_1}^{y_2} |\psi_y| dy = \int_0^{\gamma_{\text{max}}} |\psi_\gamma| d\gamma = \int_0^{\gamma_{\text{max}}} \frac{\sqrt{1 + \epsilon^2}}{\epsilon} d\gamma$$

$$= \frac{\sqrt{1 + \epsilon^2}}{\epsilon} \gamma_{\text{max}} < \frac{A}{\epsilon},$$

where $A$ is some constant that gives the uniform bound for all paths $\psi$ from $C_1$ to $C_2$ with the velocity vector given by Eq. (19). E.g. one can take $A = 2 |y_2 - y_1| \max \frac{|d^2 x|}{|d^2 y|}$, where the maximum is taken in the whole closed region bounded by $C_1$ and $C_2$. Therefore,

$$S(\psi) \leq \int_0^{A/\epsilon} |b| - b \cdot \psi_s ds = \int_0^{A/\epsilon} |b| - \frac{|b|^2 + \epsilon b \cdot b^\perp}{|b| \sqrt{1 + \epsilon^2}} ds$$

$$= \int_0^{A/\epsilon} |b| - |b| \left(1 - \frac{\epsilon^2}{2} + O(\epsilon^4)\right) ds < \int_0^{A/\epsilon} \epsilon^2 |b| ds$$

$$\leq \epsilon^2 \max_{x \in \psi} |b(x)| \frac{A}{\epsilon} = \epsilon A \max_{x \in \psi} |b(x)|$$

Since this argument is valid for any $\epsilon > 0$, $U(C_1, C_2) = 0$. Since this is true for any pair of trajectories in $D$, $U(x, y) = 0$ for any pair of points $x \in D \setminus \{x_0\}$ and $y \in D \setminus \{x_0\}$. Since $U(x, y)$ is continuous function of both arguments, we conclude that $U(x, y) = 0$ for all $x \in D$, $y \in D$. \qed
2.2. Linear SDEs. In this section we show how the quasipotential can be computed exactly for linear systems with asymptotically stable equilibria.

**Theorem 3.** Let the origin $x_0$ be an asymptotically stable equilibrium of $\dot{x} = Ax$, $x = (x_1, x_2)^T$, where

$$Ax = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$  

Let $[Ax]^\perp = \begin{bmatrix} -c & -d \\ a & b \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Then the vector field $b(x)$ can be decomposed as

$$Ax = -\frac{1}{2} \nabla U(x) + l(x),$$

where

$$-\frac{1}{2} \nabla U(x) = Ax \sin^2 \theta - [Ax]^\perp \sin \theta \cos \theta$$

$$l(x) = Ax \cos^2 \theta + [Ax]^\perp \sin \theta \cos \theta$$

where $\theta = \arccot \left( \frac{c - b}{a + d} \right)$.

$U(x)$ is the quasipotential for the linear SDE

$$dx = Ax + \sqrt{2} \beta^{-1} dw$$

w.r.t. its asymptotically stable equilibrium at the origin. It is given by

$$U(x) = A x_1^2 + 2B x_1 x_2 + C x_2^2,$$

where

$$A = -\left( a \sin^2 \theta + c \sin \theta \cos \theta \right),$$

$$B = -\left( b \sin^2 \theta + d \sin \theta \cos \theta \right) = -\left( c \sin^2 \theta - a \sin \theta \cos \theta \right),$$

$$C = -\left( d \sin^2 \theta - b \sin \theta \cos \theta \right),$$

and $\theta$ is given by Eq. (23).

**Proof.** First we observe that since the origin $x_0$ is asymptotically stable, the eigenvalues $\lambda_1$ and $\lambda_2$ of the matrix of the system $\dot{x} = Ax$ must have negative real parts, and their sum and product fulfill $\lambda_1 + \lambda_2 = a + d < 0$, and $\lambda_1 \lambda_2 = ad - bc > 0$.

Next, we modify the equation $\dot{x} = Ax$ by rotating the force field $Ax$ by angle $\theta$

$$\dot{x} = Ax \cos \theta + [Ax]^\perp \sin \theta$$

where $\theta \in \left[ -\frac{\pi}{2}, 0 \right) \cup \left( 0, \frac{\pi}{2} \right]$ is to be defined so that the trajectories of the new equation are closed curves. Therefore, the eigenvalues $\mu_1$ and $\mu_2$ of the matrix of Eq. (30) need to be complex conjugate and purely imaginary. Therefore

$$0 = \mu_1 + \mu_2 = a \cos \theta - c \sin \theta + d \cos \theta + b \sin \theta = (a + d) \cos \theta - (c - b) \sin \theta.$$
Since $a + d < 0$, we can define the angle $\theta$ by
\begin{equation}
\cot \theta = \frac{c - b}{a + d}.
\end{equation}

We denote the vector field in the right-hand side of Eq. (30) by $\tilde{l}(x)$ and introduce an orthogonal to it field
\[ u(x) = -Ax \sin \theta + [Ax]^{\perp} \cos \theta. \]
Now we decompose the field $Ax$ as
\begin{equation}
Ax = \alpha_u u(x) + \alpha_l \tilde{l}(x)
\end{equation}
where $\alpha_u$ and $\alpha_l$ are constants to be determined. From Eq. (33) we have
\[ Ax = Ax \left( \alpha_l \cos \theta - \alpha_u \sin \theta \right) + [Ax]^{\perp} \left( \alpha_u \cos \theta + \alpha_l \sin \theta \right). \]
Therefore
\begin{equation}
\alpha_u = -\sin \theta, \quad \alpha_l = \cos \theta.
\end{equation}
We set $-\frac{1}{2} \nabla U(x) = \alpha_u u(x)$ and $\tilde{l}(x) = \alpha_l \tilde{l}(x)$. Combining Eqs. (33) and (34) with the definitions of the fields $u(s)$ and $\tilde{l}(x)$ we obtain the right-hand sides of Eqs. (21) and (22).

Now it remains to show that
\[ u(x) = -Ax \sin \theta + [Ax]^{\perp} \cos \theta = -\frac{1}{2} \nabla U \]
for the potential $U$ given by Eq. (25). Integrating $-u(x)$ we obtain that $U(x)$ must satisfy
\[
U(x) = -x_1^2 \left( a \sin^2 \theta + c \sin \theta \cos \theta \right) - 2x_1x_2 \left( b \sin^2 \theta + d \sin \theta \cos \theta \right) + f_2(x_2)
\]
\[
= -x_2^2 \left( d \sin^2 \theta - b \sin \theta \cos \theta \right) - 2x_1x_2 \left( c \sin^2 \theta - a \sin \theta \cos \theta \right) + f_1(x_1),
\]
where $f_1(x_1)$ and $f_2(x_2)$ are arbitrary functions. The equality of the coefficients of $x_1x_2$ follows from the definition of the angle $\theta$. Indeed, multiplying Eq. (32) by the nonzero constant $(a + d) \sin^2 \theta$ we get
\[
(a + d) \sin \theta \cos \theta = (c - b) \sin^2 \theta,
\]
that implies the equality of the coefficients of $x_1x_2$. Finally, using Eq. (35) and setting $U = 0$ at the origin we obtain Eq. (25).

**Remark** The level sets of the quasipotential constructed in Theorem 3 are the closed trajectories of the equation $\dot{x} = l(x)$.

**Remark** Eq. (25) for the quasipotential can be compared with the formula obtained in [5, 6] for the equation $\dot{x} = Ax$ in arbitrary dimension
\[
U(x) = \frac{1}{2} \left( \left( \int_0^\infty e^{At} e^{A^*t} \, dt \right)^{-1} x, x \right).
\]
This equation can be simplified if the matrix $A^T A = AA^T$ is normal.

**Remark** Unfortunately, the construction of the quasipotential conducted in the proof of Theorem 3 is not applicable for linear systems in $\mathbb{R}^n$ for $n \geq 3$ or for an arbitrary nonlinear system in $\mathbb{R}^2$ because the angle $\theta$ between the field $b(x)$ and the gradient of the quasipotential is not generally constant in these cases.

**Example 2** For the linear SDE of the form of Eq. (24) with
\[
A = \begin{pmatrix}
-1 & -1 \\
0 & -1
\end{pmatrix},
\]
the angle $\theta$ is determined by $\cot \theta = -\frac{1}{2}$. Hence $\sin \theta = -\frac{2}{\sqrt{5}}$ and $\cos \theta = \frac{1}{\sqrt{5}}$.

Then
\[
l(x) = \begin{bmatrix}
\frac{1}{5} [-1 & -1] - \frac{2}{5} [0 & 1] \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix}
\frac{1}{5} [-1 & -3] \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},
\]
\[
-\frac{\nabla U(x)}{2} = \begin{bmatrix}
\frac{4}{5} [-1 & -1] + \frac{2}{5} [0 & 1] \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix}
\frac{1}{5} [4 & 2] \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.
\]
Hence the quasipotential is
\[
U(x) = \frac{1}{5} \left( 4x_1^2 + 4x_1x_2 + 6x_2^2 \right).
\]
2.2.1. Equilibrium probability density for linear SDEs. The quasipotential allows to estimate the equilibrium probability density $m(x)$ (or the invariant probability measure) in the case where the vector field $b$ has a unique stable equilibrium $x_0$ [10] (Chapter 4, Theorem 4.3):

$$m(x) \asymp e^{-\frac{\beta}{2}U(x_0, x)}.$$  

We remind that the symbol $\asymp$ means logarithmic equivalence. Hence this estimate gives the exponent of the equilibrium density but not its prefactor.

The quasipotential given by Eq. (25) allows to find exactly the equilibrium probability density for the linear diffusion process $dx = Ax + \sqrt{2\beta^{-1}}dw, x \in \mathbb{R}^2$:

$$m(x) = Z^{-1}e^{-\frac{\beta}{2}U(x)},$$

where $U(x)$ is given by Eq. (25) and $Z$ is the normalization constant.

**Exercise** Prove that the equilibrium probability density is given by Eq. (37) whenever quasipotential is continuously differentiable and the rotational component $l(x)$ of the vector field $b(x)$ is divergence-free, i.e., $\nabla \cdot l(x) = 0$.

We remind that a function $m(x)$ is the equilibrium probability density for the diffusion process $dx = b(x)dt + \sqrt{2\beta^{-1}}dw, x \in \mathbb{R}^n$, if and only if it satisfies the condition $\int_{\mathbb{R}^n} m(x)dx = 1$ and the stationary forward Kolmogorov equation

$$\nabla \cdot J(x) = 0,$$

where $J := J(x) = b(x)m(x) - \beta^{-1}\nabla m(x)$

is the probability current. Let $C$ be an attractor of $\dot{x} = b(x)$ and $U(C, x)$ be the quasipotential. Suppose $U(C, x)$ is continuously differentiable. Plugging $m(x)$ given by in Eq. (37) into the expression for the probability current $J(x)$ and using the decomposition $b(x) = -\frac{1}{2}\nabla U(C, x) + l(x)$, we get

$$J(x) = b(x)m(x) - \beta^{-1}\nabla m(x)$$

$$= \left(-\frac{1}{2}\nabla U(C, x) + l(x)\right) m(x) + \frac{1}{2}\nabla U(C, x)m(x) = l(x)m(x).$$

Then its divergence is

$$\nabla \cdot J(x) = \left(\nabla \cdot l(x) - \frac{\beta}{2}l(x) \cdot \nabla U(C, x)\right)m(x) = m(x)\nabla \cdot l(x)$$

due to the orthogonality of $l(x)$ and $\nabla U(C, x)$. Therefore, $\nabla \cdot J(x) = 0$ if and only if $U(x)$ is continuously differentiable and $l(x)$ is divergence-free.

Show that $l(x)$ given by Eq. (22) is divergence-free. Indeed,

$$\nabla \cdot l(x) = \nabla \cdot \left\{ \left( \begin{array}{c} ax_1 + bx_2 \\ cx_1 + dx_2 \end{array} \right) \cos^2 \theta + \left( \begin{array}{c} -cx_1 - dx_2 \\ ax_1 + bx_2 \end{array} \right) \cos \theta \sin \theta \right\}$$

$$= a \cos^2 \theta - c \cos \theta \sin \theta + d \cos^2 \theta + b \cos \theta \sin \theta$$

$$= \left( \cot \theta - \frac{c-b}{a+d} \right) (a+d) \cos \theta \sin \theta = 0.$$
2.3. An example with a limit cycle. We would like to give an example of an equation with a stable limit cycle where the quasipotential can be found analytically. We consider the SDE made up from an example of a system with a stable limit cycle from [15]

\begin{align}
\frac{dx}{dt} &= (y + x(1 - x^2 - y^2))dt + \sqrt{2\beta^{-1}}dw_1 \\
\frac{dy}{dt} &= (-x + y(1 - x^2 - y^2))dt + \sqrt{2\beta^{-1}}dw_2.
\end{align}

The corresponding equation without the stochastic term takes the form

\begin{align}
\frac{dr}{dt} = r(1 - r^2), \quad \frac{d\phi}{dt} = -1
\end{align}

in the polar coordinates, and its general solution is

\begin{align}
\phi(t) &= -(t - t_0), \\
r(t) &= \frac{1}{\sqrt{1 + ke^{-2t}}},
\end{align}

where \(k\) and \(t_0\) are arbitrary constants. One can see that this equation has the stable limit cycle \(r(t) = 1\) for \(k = 0\), and all other trajectories approach it from inside or from outside as \(t \to +\infty\). It is easy to decompose the vector field in Eq. (40) into the potential and rotational components

\begin{align}
b(x, y) &= -\frac{1}{2}\nabla U + l(x, y)
\end{align}

just by glance. One can check that

\begin{align}
U &= \frac{1}{2} (x^4 + 2x^2y^2 + y^4) - x^2 - y^2 + \frac{1}{2} (r^2 - 1)^2, \\
l &= \begin{pmatrix} y \\ -x \end{pmatrix}.
\end{align}

One can see from Eq. (41) that the quasipotential \(U(r)\) has a minimum at the limit cycle \(r = 1\) as \(U'(r) = r(r^2 - 1)\) is zero at \(r = 1\), and \(U''(r) = 3r^2 - 1\) is \(2 > 0\) at \(r = 1\). Furthermore, we can calculate the angle \(\theta(x, y)\) between the vector fields \(b(x, y)\) and \(l(x, y)\):

\begin{align}
\cos \theta = \frac{b \cdot l}{|b||l|} = \frac{x^2 + y^2}{r\sqrt{r^6 - 2r^4 + 2r^2}} = \frac{1}{\sqrt{(r^2 - 1)^2 + 1}}.
\end{align}

Therefore, \(\theta\) depends only on \(r\) and \(\theta(r) \to \arccos(1) = 0\) as \(r \to 1\).

We also point out, that the equilibrium probability density for the diffusion process given by Eq. (40) can be found exactly by

\begin{align}
m(x_1, x_2) = Z^{-1}e^{-\frac{\beta}{2}U(x, y)},
\end{align}

where \(U(x, y)\) is given by Eq. (41) and \(Z\) is the normalization constant. Indeed, we have shown in Section 2.2.1 that it suffices to check whether the rotational component \(l(x, y)\) given by Eq. (42) is divergence-free. Obviously, this is the case.
3. Numerical examples

3.1. Test example. In this section, we test the Ordered Upwind Method on the examples with known answer. We consider equations of the form

\[ \begin{align*}
  dx &= (-2x - ay)dt + \sqrt{2\beta^{-1}}dw_1 \\
  dy &= (2ax - y)dt + \sqrt{2\beta^{-1}}dw_2,
\end{align*} \]

(44)

where the parameter \( a \) is the ratio of the length of the vector \( l(x) \), the rotational component to the length of the vector \( \nabla U_2 \), the potential component. It is easy to check that the quasipotential equals \( U = 2x^2 + y^2 \). The numerical results for \( a = 1 \), \( a = 4 \), and \( a = 10 \) are shown in Figs. 6 (a), (c), and (e) respectively.

3.2. Maier-Stein model. In this section we present the quasipotential and the Minimum Action path for the Maier-Stein model [17] with the parameters \( \mu = 1 \) and \( \alpha = 10 \)

\[ \begin{align*}
  dx &= (x - x^3 - 10xy^2)dt + \sqrt{2\beta^{-1}}dw_1 \\
  dy &= -(1 + x^2)ydt + \sqrt{2\beta^{-1}}dw_2.
\end{align*} \]

(45)

Eq. (45) has three equilibria: \((\pm 1, 0)\) are stable nodal points (star points for the linearized equation), and \((0, 0)\) is a saddle. The quasipotential is non-differentiable along some interval in the \( x \)-axis.

3.3. A system with two stable limit cycles. In this section we demonstrate that one can use the ordered upwind method for computation of the quasipotential w.r.t. limit cycles and for finding the Minimum Action paths between them. We consider the equation

\[ \begin{align*}
  dx &= \left( x - \frac{x^3}{3} + y - \frac{y^3}{9} \right) dt + \sqrt{2\beta^{-1}}dw_1 \\
  dy &= (x + a)dt + \sqrt{2\beta^{-1}}dw_2.
\end{align*} \]

(46)

For various values of the parameter \( a \) this system has from one to three equilibria. The value of \( x \) in any critical point is always \(-a\). The diagram of the critical points depending on the parameter \( a \) is shown in Fig. 8 (a). For \(|a| < 1\) this system has three critical points, one saddle and two unstable spiral points. The diagram at Fig. 8 (b) shows the existence of stable limit cycles. We have plotted the values of \( y \) on the stable orbits for \( x = -a \). Comparing Figs. (a) and (b) we see that for \(|a| \geq 2.15\) there is only one stable orbit that is a stable node. For \( 2.15 > |a| > 1 \) there are two stable nodes. For \( 1 < |a| < 0.75 \) there are two limit cycles. For \(|a| < 0.75 \) there is only one limit cycle. For \(|a| < 0.5 \) all of the equilibria lie inside it.

We choose \( a = 0.9 \) in order to have two stable limit cycles in Eq. (46). We denote the upper limit cycle via \( C_1 \), and the lower one via \( C_2 \). In Fig. 9 (a) and (c) we show the level sets of the quasipotentials \( U(C_i; x, y), i = 1, 2 \), and the Minimum Action paths from \( C_1 \) to \( C_2 \) and from \( C_2 \) to \( C_1 \) respectively. We have found the quasipotential barriers \( \Delta U_1 = 0.1567 \) between the limit cycles \( C_1 \) and \( C_2 \), and \( \Delta U_2 = 0.0289 \) between the limit cycles \( C_2 \) and \( C_1 \).
Figure 6. The test example. The quasipotential for Eq. (44) for $a = 1$ (a), $a = 4$ (b), and $a = 10$ (c) respectively. The solid thick black curves are the level sets of the computed quasipotential, the dashed red curves are the level sets of the exact quasipotential. The solid thick curve curve from the origin to the point $P = (0, 1)$ is the computed Minimum Action Path, while the dashed thick black curve from the point $P = (0, 1)$ to the origin is the “exact” trajectory.

4. Methods for finding Minimum Energy Paths and saddle points

4.1. The string method. The first version of the string method was proposed by E, Ren and Vanden-Eijnden in 2002 [7]. An improved and simplified version was introduced by
the same authors in 2007 [8]. An analysis of the curve evolution under the string method was conducted in [4].

4.1.1. Curve evolution under the string method. The string method is a numerical procedure accomplishing a curve evolution described below in $\mathbb{R}^n$ under the influence of a potential $V : \mathbb{R}^n \to \mathbb{R}$. A curve is a continuous image of a unit interval in $\mathbb{R}^n$. Moreover, the ends of the curve are fixed at two distinct critical points of the potential. Suppose the curve is differentiable and its configuration at time $t$ is parametrized nondegenerately by $\alpha \in (A(t), B(t))$, i.e., the curve at time $t$ is $\{\phi(\alpha, t) \mid \alpha \in (A(t), B(t))\}$ and its unit tangent is $\tilde{\tau}(\alpha, t) = \frac{\phi_{\alpha}}{|\phi_{\alpha}|}$. Then the string method evolves it so that the normal velocity is the projection of $\nabla V$ normal to the curve:

$$
\phi_t(\alpha, t) = -\nabla V_{\perp} + \lambda \tilde{\tau},
$$

where

$$
\nabla V_{\perp} = \nabla V(\phi(\alpha, t)) - (\nabla V(\phi(\alpha, t)) \cdot \tilde{\tau}(\alpha, t)) \tilde{\tau}(\alpha, t).
$$

The evolution as a geometric curve is well defined, although $\lambda = \lambda(\alpha, t)$ is not unique. Different choices of $\lambda$ correspond to different parametrizations. This is easy to see using the chain rule: if $\tilde{\phi}(\alpha, t)$ and $\tilde{\phi}(\tilde{\alpha}, t)$ represent the same evolving curve, then $\phi(\alpha, t) = \tilde{\phi}(\tilde{\alpha}(\alpha, t), t)$ and differentiation gives

$$
\phi_t = \tilde{\phi}_t + \tilde{\phi}_{\tilde{\alpha}} \tilde{\alpha}_t = \tilde{\phi}_t + \mu \tilde{\tau} \text{ with } \mu = |\tilde{\phi}_{\tilde{\alpha}}| \tilde{\alpha}_t.
$$

It is often convenient to choose a particular parametrization. For a robust numerical solution, a good choice is the unit-speed parametrization, i.e., $|\phi_{\alpha}| = 1$. Then $\lambda(\alpha, t)$ is
Figure 8. The equilibria (a) and the stable limit cycles (b) of Eq. (46) depending on the parameter \( a \). (a): The \( y \)-coordinate of the equilibria \((x = -a)\) at each of the equilibria. Black: saddle point; Blue: stable node; Green: stable spiral point; Magenta: unstable spiral point. (b): The \( y \)-coordinates of the points of intersection of stable orbits with the line \( x = -a \). The different colors (red and blue) are used only to distinguish the separate orbits. Comparing Figs. (a) and (b) we see that for \( |a| \geq \sim 2.15 \) there is only one stable orbit that is a stable node. For \( \sim 2.15 > |a| > 1 \) there are two stable nodes. For \( 1 > |a| < \sim 0.75 \) there are two limit cycles. For \( |a| < \sim 0.75 \) there is only one limit cycle. For \( |a| < \sim 0.5 \) all of the equilibria lie inside it.

fully determined, and \( \alpha \) ranges over \((0,l(t))\) where \( l(t) \) is the length of the curve at time \( t \). Moreover, it is advantageous to avoid evaluation of \( \nabla V \perp \). If we set 

\[
\beta = \lambda + (\nabla V(\phi(\alpha,t)) \cdot \hat{\tau}(\alpha,t)),
\]

then the evolution law is given by 

\[
(49) \quad \phi_t(\alpha,t) = -\nabla V(\phi(\alpha,t)) + \beta(\alpha,t)\hat{\tau},
\]

where \( \beta(\alpha,t) \) is again uniquely determined by the requirement that \( |\phi_\alpha| = 1 \). To be stationary under Eqs. (47) or (49), a curve must satisfy \( \nabla V \perp = 0 \) pointwise. When this happens, Eq. (47) says that \( \phi_t \) is everywhere tangent to the curve. Therefore, its image (as a geometric curve) doesn’t change, although the parametrization may change in time. Thus: a piecewise smooth curve passing through a sequence of critical points \( x_1, \ldots, x_N \) (the points where \( \nabla V(x) = 0 \)) is a stationary state of the string method if the curve
The quasipotentials $U(C_1; x, y)$ (a) and $U(C_2; x, y)$ (b), where $C_1$ is the upper limit cycle, and $C_2$ is the lower limit cycle in the system given by Eq. (46) with $a = 0.9$ respectively. The quasipotential barriers found are $\Delta U_1 = 0.1567$ between the limit cycles $C_1$ and $C_2$, and $\Delta U_2 = 0.0289$ between the limit cycles $C_2$ and $C_1$. The Minimum Action paths are shown with thick white curves. The limit cycles $C_1$ and $C_2$ are shown with thick yellow curves.

is everywhere tangent to $\nabla V$. Put differently: the curve is stationary if it consists of a sequence of critical points connected by heteroclinic orbits (solutions of $\dot{x} = -\nabla V(x)$ traced forward or backward in time). As noted previously, such curves are called Minimum Energy Paths.

We remark that the curve evolution under the string method is mathematically equivalent to the one under the Nudged Elastic Band method [16]. However, this curve evolution is accomplished by different numerical procedures.

4.1.2. Numerical implementation. Let $x_A$ and $x_B$ be two potential minima or two points lying close to two isolated potential minima. Let $\phi_0(\alpha) \in [0, 1]$ be an initial curve connecting $x_A$ and $x_B$, i.e., $\phi_0(0) = x_A$, $\phi_0(1) = x_B$. For example, $\phi_0(\alpha)$ can be a segment of strength line connecting $x_A$ and $x_B$. We discretize the curve, i.e., represent it as a sequence of $N$ points (these points are called "images" by chemical physicists) $\{\phi(\alpha_j)\}_{j=1}^N = \{\phi_j\}_{j=1}^N$, $\alpha_j = j/(N - 1)$. Each iteration of the string method evolving a curve $\phi$ according to Eq. (49) consists in two steps. Pick a time step $h$.

**Gradient descent:** Move images according to the gradient descent $\phi_j^* = \phi_j^k - h\nabla V(\phi_j)$, $1 \leq j \leq N$. 

![Figure 9](image-url)
Reparametrization: Define a new continuous curve $\phi^*$ by interpolation between the updated images $\{\phi_j^*\}_{j=1}^N$. Distribute the new images $\phi_j^{new}$ uniformly along the curve $\phi^*$.

Stop if the curve almost stops moving, i.e. if $\max_{1 \leq j \leq N} \|\phi_j^{new} - \phi_j\| < tol$ where $tol$ is the user-prescribed tolerance.

The string method was used to obtain the Minimum Energy Paths in Fig. 1 connecting all neighboring minima of the 7-well potential.

4.2. The dimer method. The dimer method for finding saddle points was proposed by Henkelman and Jonsson in 1999 [12]. Later, Du and Zhang modified it into the shrinking dimer method and analyzed its dynamics and convergence properties [18]. Suppose we want to find a saddle point of Morse index 1 in the potential energy landscape $V(x)$, $x \in \mathbb{R}^n$.

Here we will present a basic version of the shrinking dimer method.

4.2.1. Shrinking dimer dynamics. A saddle of Morse index 1 is a saddle at which the Hessian matrix of $V$ has a unique negative eigenvalue. A dimer is the segment of the straight line connecting a pair of points $x_1, x_2 \in \mathbb{R}^n$. Its length is

$$l = \|x_2 - x_1\|.$$ 

Its center is the point

$$x := \frac{x_1 + x_2}{2}.$$ 

Its direction is characterized by the unit vector

$$v := \frac{x_2 - x_1}{l}.$$ 

Given the position of the center $x$ and the direction $v$ one can restore the positions of the endpoints of the dimer:

$$x_1 = x - \frac{l}{2}v, \quad x_2 = x + \frac{l}{2}v.$$ 

The shrinking dimer dynamics is designed so that if the initial approximation is good enough, the dimer gets oriented so that the vector $v$ is parallel to the eigenvector corresponding to the only negative eigenvalue, climbed to the saddle along the direction of this eigenvector (this is the gentlest ascend direction), as shrinks to a point while approaching the saddle.

Let $F_1 = -\nabla V(x_1)$ and $x_2 = -V(x_2)$ be the natural forces acting on the ends of the dimer. The force acting on the center of the dimer, $F$ is defined as the average of these forces:

$$F = \frac{F_1 + F_2}{2}.$$ 

The difference of the forces acting on $x_1$ and $x_2$ is

$$\Delta F := F_2 - F_1.$$
The time evolution of the dimer is defined by the following system of ODEs:

\begin{align}
\mu_1 \dot{x} &= (I - 2vv^T)F, \\
\mu_2 \dot{v} &= (I - vv^T)\Delta F/l, \\
\mu_3 \dot{l} &= -l.
\end{align}

The coefficients $\mu_1$, $\mu_2$ and $\mu_3$ allow us to adjust the relationships between the speed of the motion of the center of the dimer, the speed of its rotation, and the speed of its shrinking.

The operator $(I - 2vv^T)$ in Eq. (50) reflects the vector $F$ with respect to the hyperplane orthogonal to the unit vector $v$ along the direction of the dimer. Note that if the dimer is parallel to the eigenvector $u$ corresponding to the only negative eigenvalue of the Hessian matrix of the potential $V$, and the center of the dimer $x$ is located so that the radius vector from it to the saddle is also parallel to this eigenvector then

$$(I - 2vv^T)F = (I - 2vv^T)(-|F|v) = v|F| = -F,$$

i.e., the dimers moved directly uphill toward the saddle.

The operator $(I - vv^T)$ in Eq. (51) subtracts from the vector $\Delta F$ its component parallel to the direction of the dimer $v$. Hence $\dot{v}$ is orthogonal to $v$ at all times. Therefore, the unit length of $v$ is preserved.

The right-hand side of Eq. (52) is completely determined by $l$. Here we have picked the exponential shrinking law. One can pick another law.

The global convergence of the shrinking dimer method was proven in [18] for the quadratic potential of the form

$$V(z) = -\frac{z_1^2}{2} + \frac{1}{2} \sum_{i=2}^{n} z_i^2.$$ 

Precisely, for any initial condition $x_0, v_0 \in \mathbb{R}^n$ such that $v_0$ is not perpendicular to the vector $(1, 0, ..., 0)$, the solution of Eqs. (50)-(52) given by (2.4) will converge to $(x^*, v^*, l^*)$ as $t \to \infty$, where $x^* = (0, ..., 0)$, $v^* = (1, 0, ..., 0)$ or $v^* = ?(1, 0, ..., 0)$, and $l^* = 0$.

4.2.2. Numerical implementation. Pick a time step $h$. Pick the initial position of the center $x_0$, the initial direction of the dimer $v_0$, and the initial length $l_0$. Then for $k = 0, 1, 2, \ldots$
iterate:

\begin{align*}
F_1 &= -\nabla V(x_k - lv_k/2), \quad F_2 = -\nabla V(x_k + lv_k/2), \\
F &= \frac{1}{2}(F_1 + f_2), \quad \Delta F = F_2 - F_1, \\
x_{k+1} &= x_k + \frac{h}{\mu_1} (I - 2v_k v_k^T) F, \\
w &= v_k + \frac{h}{\mu_2 l_k} (I - v_k v_k^T) \Delta F, \\
v_{k+1} &= \frac{w}{\|w\|}, \\
l_{k+1} &= \frac{l_k}{1 + h/\mu_3}.
\end{align*}

This scheme is referred to as the Modified Euler scheme in [18]. The first two equations perform Forward Euler steps. Then \(v\) is normalized. \(l\) is updated according to the Backward Euler scheme. As a stopping criterion, you can use, for example, \(|\nabla V(x)| < \text{tol}\) where \(\text{tol}\) is some reasonable tolerance.

Time evolution of the shrinking dimer method is illustrated in Fig. 10.

5. Freidlin’s cycles

Freidlin introduced a hierarchy of cycles to describe the dynamics in the case where the system evolving according to \(dx = b(x) dt + \sqrt{2\beta^{-1}} dw\) has multiple attractors, and \(\beta^{-1}\) is close to zero [9], [10] (Chapter 6), [11]. In the case of the overdamped Langevin dynamics

\[ dx = -\nabla V(x) dt + \sqrt{2\beta^{-1}} dw \]

the construction of the hierarchy of cycles can be simplified [3]. For simplicity, here we present the construction of the hierarchy of Freidlin’s cycles for the overdamped Langevin dynamics on the example of a three-well potential.

Suppose two minima \(A\) and \(B\) are separated by a single mountain pass \(x^*\) (a Morse index 1 saddle). According to the Large Deviation Theory, the most likely transitions between the minima \(A\) and \(B\) occur in a small tube surrounding the global minimizer of the Freidlin-Wentzell action over all travel times \(\tau\) and all paths \(\phi(t)\) with the boundary conditions \(\phi(0) = A, \phi(\tau) = B\). The infimum of the Freidlin-Wentzell action is

\[ \inf_{\tau, \phi} S_\tau(\phi) = 2[V(x^*) - V(A)] \]

and it is achieved on the Minimum Energy Path (MEP) (i.e., a path \(\phi\) such that \(\phi' \parallel \nabla V\)), going from the minimum \(A\) directly uphill to the saddle \(x^*\), and then going directly downhill from the saddle \(x^*\) to the minimum \(B\).

However, if the potential energy landscape has multiple minima and mountain passes, the problem of finding the most likely transition paths between two given minima becomes more complicated. The global minimizer of the Freidlin-Wentzel action between two given minima does not necessarily indicate the most likely transition path (see the example with
Figure 10. Illustration of one step of the shrinking dimer method on two examples of dimer. (a), (b), (c): Example 1. (d), (e), (f): Example 2. (a), (c): Calculation of the vector \((I - 2vv^T)F\). (b), (e): Calculation of the vector \((I - vv^T)\Delta F\). (c), (f): The updated dimers.
a three-well potential ahead). The most likely transition path in the limit of $\beta^{-1} \to 0$ (the asymptotic zero-temperature path) is the MinMax path, i.e., the lowest possible transition path connecting these two minima. This path can be obtained from the construction of the hierarchy of Freidlin’s cycles.

We will start with a simple and illustrative example that motivates the general construction.

5.1. **Example: a three-well potential.** We consider a potential with three minima $A$, $B$, and $C$ and three saddles $S_{AB}$, $S_{BC}$ and $S_{AC}$ shown in Fig. 2. For brevity, we will denote the values of the potential at the minima and the saddles as $V_A \equiv V(A)$, $V_B \equiv V(B)$, $V_C \equiv V(C)$, $V_{AB} \equiv V(S_{AB})$, $V_{AC} \equiv V(S_{AC})$, $V_{BC} \equiv V(S_{BC})$.

The values of the potential at the saddles satisfy

$$V_{AB} > V_{AC} > V_{BC} \quad \text{and} \quad V_{AC} + V_{BC} > V_{AB}. \quad (60)$$

Suppose we are interested in transitions from $A$ to $B$. The global minimizer of the Freidlin-Wentzell action with endpoints at $A$ and $B$ is achieved on the MEP such that the sum

$$\sum_{\text{path}} (V_{\text{saddle}} - V_{\text{min}}) \quad (61)$$

is minimal along it. It follows from Eq. (61) that the global minimizer of the Freidlin-Wentzell action is the upper MEP passing through the highest saddle $S_{AB}$. Below we will show that this MEP does not describe the transition process correctly as the temperature tends to zero.
Asymptotically, as the temperature tends to zero, the stochastic dynamics can be described by a discrete-space continuous-time Markov chain \([9, 10, 11]\). The local equilibria \(A, B,\) and \(C\) are called the zero order cycles. The transition rates between the states \(A, B,\) and \(C\) are given by the generator matrix (or the rate matrix) \(Q\). The off-diagonal entries of \(Q\) are the pairwise transition rates defined by
\[
q_{ij} = \lim_{t \to \infty} \frac{N_{ij}(t)}{\tau_i}, \quad i, j \in \{A, B, C\},
\]
where \(N_{ij}\) is the number of transitions from the basin of attraction of the minimum \(i\) to the basin of attraction of the minimum \(j\) that have happened over the time interval \([0, t]\), and \(\tau_i\) is the total time spent in the basin of attraction of the minimum \(i\) during the time interval \([0, t]\). The diagonal entries of \(Q\) are defined by
\[
q_{ii} = -\sum_{j \neq i} q_{ij},
\]
so that the sum of each row of \(Q\) is zero. Taking into account the result of the Large Deviation Theory \([10]\) we get,
\[
Q = \begin{bmatrix}
-q_{AB} - q_{AC} & q_{AB} & q_{AC} \\
q_{BA} & -q_{BA} - q_{BC} & q_{BC} \\
q_{CA} & q_{CB} & -q_{CA} - q_{CB}
\end{bmatrix},
\]
where \(q_{ij} \propto e^{-\beta(V_{ij} - V_i)}\), \(i, j \in \{A, B, C\}\).

The construction of the hierarchy will be done in 2 stages.

**Stage 1:** First we introduce a new notation for the generator matrix \(Q\): \(Q \equiv Q^{(1)}\). The purpose of this will become clear shortly. Next, we convert the generator matrix \(Q^{(1)}\) into the jump matrix \(\Pi^{(1)}\) and then take the zero-temperature limit, i.e., let \(\beta \to \infty\). The jump matrix \(\Pi^{(1)}\) for the generator matrix \(Q^{(1)}\) is defined by the rule: \(\Pi^{(1)}_{ii} = 0\) if \(q_{ii} \neq 0\) (our case), and \(\Pi^{(1)}_{ij} = -q_{ij}/q_{ii}\) for \(j \neq i\). This rule leads to cancellation of \(V_A, V_B\) and \(V_C\). We note that the jump matrix coincides with the weighted adjacency matrix introduced in \([7]\). We get
\[
\Pi^{(1)} = \begin{bmatrix}
0 & \Pi^{(1)}_{AB} & \Pi^{(1)}_{AC} \\
\Pi^{(1)}_{BA} & 0 & \Pi^{(1)}_{BC} \\
\Pi^{(1)}_{CA} & \Pi^{(1)}_{CB} & 0
\end{bmatrix},
\]
where
\[
\Pi^{(1)}_{AB} \propto \frac{e^{-\beta V_{AB}}}{e^{-\beta V_{AB}} + e^{-\beta V_{AC}}}, \quad \Pi^{(1)}_{BA} \propto \frac{e^{-\beta V_{AB}}}{e^{-\beta V_{AB}} + e^{-\beta V_{BC}}},
\]
\[
\Pi^{(1)}_{AC} \propto \frac{e^{-\beta V_{AC}}}{e^{-\beta V_{AB}} + e^{-\beta V_{AC}}}, \quad \Pi^{(1)}_{CA} \propto \frac{e^{-\beta V_{AC}}}{e^{-\beta V_{AC}} + e^{-\beta V_{BC}}},
\]
\[
\Pi^{(1)}_{BC} \propto \frac{e^{-\beta V_{BC}}}{e^{-\beta V_{AB}} + e^{-\beta V_{BC}}}, \quad \Pi^{(1)}_{CB} \propto \frac{e^{-\beta V_{BC}}}{e^{-\beta V_{BC}} + e^{-\beta V_{BC}}}
\]
An off-diagonal entry $\Pi_{ij}^{(1)}$ is equal to the probability that being at the state $i$, the system will jump to the state $j$. A zero diagonal entry $\Pi_{ii}^{(1)} = 0$ indicates that the system will eventually escape from the state $i$ with probability 1. Taking the limit $\beta \to \infty$ and keeping in mind that $V_{AB} > V_{AC} > V_{BC}$, we obtain a limiting jump matrix

$$\text{(64)}$$

$$\Pi_0^{(1)} := \lim_{\beta \to \infty} \Pi^{(1)} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$ 

Eq. (63) shows that the limiting jump matrix (64) depends on the values of the potential at the saddles only. The matrix $\Pi_0^{(1)}$ corresponds to the Markov jump process shown in Fig. 3(a). In this process, the state $A$ is transient, while the states $B$ and $C$ are recurrent.

From $\text{Stage 1}$ we get

$$P(B \mid BC) = \frac{\frac{\tau_B}{\tau_B + \tau_C} \times e^{\beta(V_{BC}-V_B)}}{e^{-\beta V_B} + e^{-\beta V_C}}$$

and

$$P(C \mid BC) = \frac{\frac{\tau_C}{\tau_B + \tau_C} \times e^{\beta(V_{BC}-V_C)}}{e^{-\beta V_B} + e^{-\beta V_C}},$$

respectively. Therefore, the transition rates from $\{B, C\}$ to $A$ via $B$ and $C$ are

$$k_{BA} \equiv \lim_{t \to \infty} \frac{N_{BA}}{\tau_B + \tau_C} = P(B \mid BC)Q_{BA}^{(1)} \times \frac{e^{-\beta V_{AB}}}{e^{-\beta V_B} + e^{-\beta V_C}},$$

and

$$k_{CA} \equiv \lim_{t \to \infty} \frac{N_{CA}}{\tau_B + \tau_C} = P(C \mid BC)Q_{CA}^{(1)} \times \frac{e^{-\beta V_{AC}}}{e^{-\beta V_B} + e^{-\beta V_C}}.$$  

We define the generalized rate matrix $Q^{(2)}$ as follows. The first row of $Q^{(2)}$ coincides with the first row of $Q^{(1)}$ as the state $A$ does not belong to a first order cycle. The entries $Q_{AB}^{(2)}$ and $Q_{AC}^{(2)}$ are the transition rates from $A$ to $B$ and from $A$ to $C$ respectively. The entries $Q_{BA}^{(2)}$ and $Q_{CA}^{(2)}$ are the transition rates from the cycle $\{B, C\}$ to $A$ via $B$ and $C$ defined in
Eqs. (65) and (66) respectively. Finally, 

\[ Q_{ij}^{(2)} = -\sum_{k \in \{B,C\}, l \notin \{B,C\}} Q_{kl}^{(2)}, \]

\[ i, j \in \{B,C\}. \]

Thus,

\[
Q^{(2)} = \begin{bmatrix}
-q_{AB} - q_{AC} & q_{AB} & q_{AC} \\
k_{BA} & -k_{BA} - k_{CA} & -k_{BA} - k_{CA} \\
k_{CA} & -k_{BA} - k_{CA} & -k_{BA} - k_{CA}
\end{bmatrix},
\]

where \( k_{BA} \) and \( k_{CA} \) are given by Eqs. (65) and (66) respectively. Strictly speaking, \( Q^{(2)} \) is not a generator matrix, as the sums of its entries in the second and the third rows are not zero. It could be converted into a \( 2 \times 2 \) generator matrix of a continuous time Markov chain with two states: \( A \) and \( \{B,C\} \). But by doing this we would lose the information via what sub-state, \( B \) or \( C \), the system enters and leaves the cycle \( \{B,C\} \).

The generalized jump matrix \( \Pi^{(2)} \) is obtained from the matrix \( Q^{(2)} \) by the rule

\[
\Pi^{(2)}_{ij} = \begin{cases} 
0, & \text{if } i \text{ and } j \text{ belong to} \\
-(Q^{(2)})_{ij} / (Q^{(2)})_{ii}, & \text{otherwise}.
\end{cases}
\]

We get

\[
\Pi^{(2)} = \begin{bmatrix}
0 & \Pi^{(2)}_{AB} & \Pi^{(2)}_{AC} \\
\Pi^{(2)}_{BA} & 0 & 0 \\
\Pi^{(2)}_{CA} & 0 & 0
\end{bmatrix},
\]

where

\[
\Pi^{(2)}_{AB} \propto \frac{e^{-\beta V_{AB}}}{e^{-\beta V_{AB}} + e^{-\beta V_{AC}}} \quad \Pi^{(2)}_{BA} \propto \frac{e^{-\beta V_{AB}}}{e^{-\beta V_{AB}} + e^{-\beta V_{AC}}},
\]

\[
\Pi^{(2)}_{AC} \propto \Pi^{(2)}_{CA} \propto \frac{e^{-\beta V_{AC}}}{e^{-\beta V_{AB}} + e^{-\beta V_{AC}}},
\]

Letting \( \beta \to \infty \) and keeping in mind that \( V_{AC} < V_{AB} \), we obtain

\[
\Pi^{(2)}_0 := \lim_{\beta \to \infty} \Pi^{(2)} = \begin{bmatrix}
0 & 0 & 1 \\
0 & 0 & 0 \\
1 & 0 & 0
\end{bmatrix}.
\]

The matrix \( \Pi^{(2)}_0 \) shows that if the system is at the macro-state \( \{B,C\} \), it jumps to \( A \) from \( C \), and if it is at \( A \), it jumps to the macro-state \( \{B,C\} \) via \( C \). The set \( \{A,C\} \) is a reduced second order cycle corresponding to the full second order cycle \( \{A,B,C\} \). The expected exit time from the cycle \( \{B,C\} \) is given by \( (Q^{(2)}_{CA})^{-1} \). The corresponding discrete-space discrete-time Markov chain is shown in Fig. 3(b).

Therefore, the system switches from state \( A \) to state \( B \) and back via state \( C \). The MEP corresponding to this transition process (the lower MEP in Fig. 2) goes from the minimum \( A \) to the minimum \( B \) via the minimum \( C \).

This example illustrates the phenomenon that the system exits from each cycle via the lowest saddle adjacent to it. Furthermore, a counterintuitive at the first glance conclusion is that as soon as \( V_{AC} < V_{AB} \) and \( V_{BC} < V_{AB} \), the system switches from \( A \) to \( B \) and
Figure 12. Illustration to the 3-well potential example in Section 5.1. The Markov chain corresponding to the matrices (a): $\Pi_0^{(1)}$ and (b): $\Pi_0^{(2)}$.

back via the intermediate minimum $C$, regardless of how deep the minimum $C$ is, if the temperature is low enough.

References


