Random perturbations of dynamical systems

Barbara Gentz
Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstraße 39
10117 Berlin
Germany

Lecture notes

Graduiertenkolleg Analysis, Geometrie und ihre Verbindung zu den Naturwissenschaften,
Universität Leipzig and MPI Mathematics in the Sciences, Leipzig

Summer term 2003

Version: July 18, 2003

Contents

1 Itô integrals and stochastic differential equations: A brief introduction

1.1 Itô integrals

We recall here some facts about stochastic integration with respect to Brownian motion and stochastic differential equations. We restrict our attention to those topics needed during the course. The presentation mainly follows [25]. For an introduction to the general theory of stochastic integration with respect to semimartingales, we refer to [26], [28], for example.

We consider the one-dimensional case first. Let $\{W_t\}_{t\geqslant 0}$ be a one-dimensional standard Brownian motion on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We want to give a meaning to equations of the form

$$\dot{x}_t = f(x_t, t) + F(x_t, t) \times \text{“\textit{white noise}”}, \qquad t \in [0, T]. \tag{1.1}$$

Considering the discrete-time version

$$x_{t_{k+1}} - x_{t_k} = f(x_{t_k}, t_k)\Delta t_k + F(x_{t_k}, t_k)\Delta W_k, \qquad k \in \{0, \dots, K-1\}, \tag{1.2}$$

with a partition $0 = t_0 < t_1 < \dots < t_K = T$, $\Delta t_k = t_{k+1} - t_k$ and Gaussian increments $\Delta W_k = W_{t_{k+1}} - W_{t_k}$ suggests to interpret (1.1) as an integral equation

$$x_t = x_0 + \int_0^t f(x_s, s)\,\mathrm{d}s + \int_0^t F(x_s, s)\,\mathrm{d}W_s, \tag{1.3}$$

provided the integral $\int_0^t F(x_s, s)\,\mathrm{d}W_s$ can be defined as the limit of $\sum_{k=0}^{K-1} F(x_{t_k}, t_k)\Delta W_k$ as $\Delta t_k \to 0$ in some suitable sense.

Thus we want to define integrals of the type $\int_0^t h(s, \omega)\,\mathrm{d}W_s(\omega)$ for some class of integrands $h(s, \omega)$ taking values in $\mathbb{R}$. Assume for the moment that $s \mapsto h(s, \omega)$ is continuous and of bounded variation for (almost) all $\omega \in \Omega$. Were the paths of the Brownian motion $s \mapsto W_s(\omega)$ also of finite variation, we could apply integration by parts, thereby obtaining

$$\begin{aligned}\int_0^t h(s, \omega)\,\mathrm{d}W_s(\omega) &= h(t)W_t(\omega) - h(0)W_0(\omega) - \int_0^t W_s(\omega)h(\mathrm{d}s, \omega) \\ &= h(t)W_t(\omega) - \int_0^t W_s(\omega)h(\mathrm{d}s, \omega),\end{aligned} \tag{1.4}$$

where the integral on the right-hand side is defined as a Stieltjes integral for each ω. So we can use (1.4) to define the stochastic integral $\int_0^t h(s, \omega)\,\mathrm{d}W_s(\omega)$ ω-wise by the well-defined right-hand side of (1.4) for such h. This class of possible integrands is not large enough for our purpose, because the paths of Brownian motion are almost surely *not* of finite variation and, therefore, we can not expect $s \mapsto h(s, \omega) = F(x_s(\omega), s)$ to be of finite variation.

The definition for a more general class of integrands as presented below, is due to Itô [18]. For integrands which are of finite variation it is equivalent to (1.4). Let $\{\mathcal{F}_t\}_{t\geqslant 0}$ denote the filtration generated by the Brownian motion $\{W_t\}_{t\geqslant 0}$, i. e.,

$$\mathcal{F}_t = \sigma\{W_s, s \leqslant t\}, \qquad t \geqslant 0, \tag{1.5}$$

is the σ-algebra generated by the Brownian motion up to time t.[1] First we define the stochastic integral for so-called elementary functions which are step functions in time.

Definition 1.1. *A function $h : [0,T] \times \Omega \to \mathbb{R}$ is called* elementary *if there exists a partition $0 = t_0 < t_1 < \ldots t_K = T$ such that*

$$h(t,\omega) = \sum_{k=0}^{K-1} h_k(\omega) 1_{(t_k,t_{k+1}]}(t), \qquad t \in [0,T], \tag{1.6}$$

and $\omega \mapsto h_k(\omega)$ is $\mathcal{F}_{t_k}$-measurable for all k.

For elementary integrands, the stochastic integral can be defined in a natural way by

$$\int_0^t h(s)\,\mathrm{d}W_s = \int_0^t h(s,\omega)\,\mathrm{d}W_s(\omega) = \sum_{k=0}^{K-1} h_k(\omega)[W_{t_{k+1}}(\omega) - W_{t_k}(\omega)]. \tag{1.7}$$

To extend this definition, we shall use the following isometry between Hilbert spaces.

Lemma 1.2. *Let the elementary function h be such that $h_k \in L^2(\mathbb{P})$ for all k. Then,*

$$\mathbb{E}\Big\{\Big(\int_0^t h(s)\,\mathrm{d}W_s\Big)^2\Big\} = \int_0^t \mathbb{E}\{h(s)^2\}\,\mathrm{d}s. \tag{1.8}$$

PROOF: First note that

$$\mathbb{E}\Big\{\Big(\int_0^t h(s)\,\mathrm{d}W_s\Big)^2\Big\} = \sum_{k,l=0}^{K-1} \mathbb{E}\{h_k h_l (\Delta W_k)(\Delta W_l)\}. \tag{1.9}$$

For $k < l$, $h_k h_l(\Delta W_k)$ is $\mathcal{F}_{t_l}$-measurable, while ΔW_l is independent of $\mathcal{F}_{t_l}$. Thus only the terms with $k = l$ contribute to the sum in (1.9). As h_k is $\mathcal{F}_{t_k}$-measurable while ΔW_k is independent of $\mathcal{F}_{t_k}$, $\mathbb{E}\{h_k^2(\Delta W_k)^2\} = \mathbb{E}\{h_k^2\}\Delta t_k$ follows. Now,

$$\mathbb{E}\Big\{\Big(\int_0^t h(s)\,\mathrm{d}W_s\Big)^2\Big\} = \sum_{k=0}^{K-1} \mathbb{E}\{h_k^2\}\Delta t_k = \int_0^t \mathbb{E}\{h(s)^2\}\,\mathrm{d}s \tag{1.10}$$

is immediate. □

By the preceding lemma, Definition (1.7) can be extended to the following class of functions $h : [0,T] \times \Omega \to \mathbb{R}$:

- $(t,\omega) \mapsto h(t,\omega)$ is measurable with respect to $\mathcal{B}([0,T]) \otimes \mathcal{F}$, where $\mathcal{B}([0,T])$ denotes the Borel σ-algebra on $[0,T]$;
- For each t, $\omega \mapsto h(t,\omega)$ is $\mathcal{F}_t$-measurable;
- $\int_0^T \mathbb{E}\{h(t)^2\}\,\mathrm{d}t < \infty$.

[1] Later we will work with the filtration generated by the Brownian motion and an initial condition x_0, which is assumed to be independent of the Brownian motion $\{W_t\}_{t \geqslant 0}$. Note that all statements remain valid when using such an enlarged filtration.

Each such h can be approximated by a sequence of elementary functions $e^{(n)}$ in the following sense:

$$\int_0^T \mathbb{E}\{(h(s) - e^{(n)}(s))^2\}\,\mathrm{d}s \to 0, \qquad \text{as } n \to \infty, \tag{1.11}$$

and the basic isometry (1.8) allows to define the stochastic integral of h by setting

$$\int_0^t h(s)\,\mathrm{d}W_s = \lim_{n\to\infty} \int_0^t e^{(n)}(s)\,\mathrm{d}W_s, \qquad \text{for all } t \in [0,T], \tag{1.12}$$

where the limit exists in $L^2(\mathbb{P})$.

This defines the so-called Itô integral. It is well-defined in the sense that its value does not depend on the precise choice of the sequence of elementary functions, but note the following: By our definition of elementary functions, h is approximated by (random) step functions $e^{(n)}$ with associated partition $t_k^{(n)}$, $k = 0, \dots, K^{(n)}$. The value of $e^{(n)}(t)$ at all times $t \in [t_k^{(n)}, t_{k+1}^{(n)}]$ is chosen $\mathcal{F}_{t_k^{(n)}}$-measurable. If we were to relax this measurability condition, the definition (1.12) of the stochastic integral would yield a different value. For instance, approximating bounded, continuous functions h by elementary functions $e^{(n)}$ with $e^{(n)}(t) = h((t_k^{(n)} + t_{k+1}^{(n)})/2)$ for all $t \in [t_k^{(n)}, t_{k+1}^{(n)}]$, would yield the so-called Stratonovich integral, on which we shall comment below.

For any interval $[a,b] \subset [0,T]$, we define

$$\int_a^b h(s)\,\mathrm{d}W_s = \int_0^T 1_{[a,b]}(s)h(s)\,\mathrm{d}W_s. \tag{1.13}$$

The stochastic integral satisfies the following properties.

- Splitting of integrals:

$$\int_s^t h(s)\,\mathrm{d}W_s = \int_s^u h(s)\,\mathrm{d}W_s + \int_u^t h(s)\,\mathrm{d}W_s \qquad \text{for } 0 \leqslant s \leqslant u \leqslant t \leqslant T; \tag{1.14}$$

- Linearity:

$$\int_0^t (ch_1(s) + h_2(s))\,\mathrm{d}W_s = c\int_0^t h_1(s)\,\mathrm{d}W_s + \int_0^t h_2(s)\,\mathrm{d}W_s \qquad \text{for all constants } c; \tag{1.15}$$

- Expectation:[2]

$$\mathbb{E}\Big\{\int_0^t h(s)\,\mathrm{d}W_s\Big\} = 0; \tag{1.16}$$

- Covariance of stochastic integrals/isometry:[2]

$$\mathbb{E}\Big\{\Big(\int_0^t h_1(s)\,\mathrm{d}W_s\Big)\Big(\int_0^t h_2(s)\,\mathrm{d}W_s\Big)\Big\} = \int_0^t \mathbb{E}\{h_1(s)h_2(s)\}\,\mathrm{d}s. \tag{1.17}$$

Next, we want to consider the stochastic integral $X_t = \int_0^t h(s)\,\mathrm{d}W_s$ as a function of t. As X_t is $\mathcal{F}$-measurable, $\{X_t\}_{t\in[0,T]}$ is a (continuous-time) stochastic process. It has the following properties.

[2] Note that these properties may—and typically will—fail for the Stratonovich integral.

- X_t is $\mathcal{F}_t$-measurable;[2]
- $\{X_t\}_t$ is an $\{\mathcal{F}_t\}_t$-martingale: $\mathbb{E}\{X_t|\mathcal{F}_s\} = X_s$ for $0 \leqslant s \leqslant t \leqslant T$;[2]
- $t \mapsto X_t(\omega)$ possesses a continuous version, i.e., there exists a stochastic process Y_t with continuous sample paths, which satisfies $\mathbb{P}\{X_t \neq Y_t\} = 0$ for all t. Thus we may always assume that $t \mapsto X_t(\omega)$ is continuous for almost all ω.

Here a word of warning is due. The definition of the Itô integral can easily be extended to integrands h satisfying the same measurability assumptions as before but a weaker integrability assumption. In fact, it is sufficient to assume that

$$\mathbb{P}\biggl\{\int_0^t h(s,\omega)^2 \,\mathrm{d}s < \infty \quad \text{for all } t \geqslant 0\biggr\} = 1. \tag{1.18}$$

The stochastic integral is then defined as the limit in probability of integrals of elementary functions. Keep in mind that for such h, those of the above properties of the stochastic integral which involve expectations may fail.

Examples 1.3. (a) Let us first look at an example where the integrand is not of finite variation. A classical example is

$$\int_0^t W_s \,\mathrm{d}W_s = \frac{1}{2}W_t^2 - \frac{1}{2}t. \tag{1.19}$$

It is not difficult to calculate this integral directly, see for instance [25, Example 3.6]. Note the unexpected $-t/2$, which shows that the Itô integral can not be calculated like ordinary integrals. This correction stems from the contribution of the quadratic variation of W_t. Below we will state Itô's formula which replaces the chain rule for Riemann integrals and is very useful for calculating Itô integrals. We remark in passing that the Stratonovich integral does not require such a correction.

(b) Our second example deals with the special case of deterministic integrands. If h does not depend on ω it follows immediately from the definition of the stochastic integral, that $\int_0^t h(s)\,\mathrm{d}W_s$ is a Gaussian random variable with mean zero and variance $\int_0^t h(s)^2 \,\mathrm{d}s$.

Theorem 1.4 (Itô's formula). *Let h satisfy the measurability assumptions and* (1.18)*, and let f be another function satisfying the same measurability assumptions and the weaker integrability assumption*

$$\mathbb{P}\biggl\{\int_0^t |f(s,\omega)| \,\mathrm{d}s < \infty \quad \text{for all } t \geqslant 0\biggr\} = 1. \tag{1.20}$$

We define

$$X_t = X_0 + \int_0^t f(s)\,\mathrm{d}s + \int_0^t h(s)\,\mathrm{d}W_s. \tag{1.21}$$

Let $g : \mathbb{R} \times [0,T] \to \mathbb{R}$ be continuous and assume that the partial derivatives

$$g_t(x,t) = \frac{\partial}{\partial t}g(x,t), \qquad g_x(x,t) = \frac{\partial}{\partial x}g(x,t), \qquad g_{xx}(x,t) = \frac{\partial^2}{\partial x^2}g(x,t) \tag{1.22}$$

exist and are also continuous. Then $Y_t = g(X_t, t)$ *is given by*

$$Y_t = g(X_0, 0) + \int_0^t \Big[g_t(X_s, s) + g_x(X_s, s)f(s) + \frac{1}{2}g_{xx}(X_s, s)h(s)^2\Big]\,\mathrm{d}s + \int_0^t g_x(X_s, s)h(s)\,\mathrm{d}W_s. \tag{1.23}$$

Denoting (1.21) briefly as $\mathrm{d}X_t = f\,\mathrm{d}t + h\,\mathrm{d}W_t$, Itô's formula can be written as

$$\mathrm{d}Y_t = g_t\,\mathrm{d}t + g_x\,\mathrm{d}X_t + \frac{1}{2}g_{xx}(\mathrm{d}X_t)^2, \tag{1.24}$$

where $(\mathrm{d}X_t)^2$ is calculated according to the scheme

$$(\mathrm{d}t)^2 = (\mathrm{d}t)(\mathrm{d}W_t) = (\mathrm{d}W_t)(\mathrm{d}t) = 0, \qquad (\mathrm{d}W_t)^2 = \mathrm{d}t. \tag{1.25}$$

Examples 1.5. (a) Using Itô's formula, we can easily calculate $\int_0^t s\,\mathrm{d}W_s$. Set $g(x,t) = t \cdot x$ and $Y_t = g(W_t, t)$. Then $\mathrm{d}Y_t = W_t\,\mathrm{d}t + t\,\mathrm{d}W_t + \frac{1}{2}0\,\mathrm{d}t$, and, therefore,

$$\int_0^t s\,\mathrm{d}W_s = Y_t - Y_0 - \int_0^t W_s\,\mathrm{d}s = tW_t - \int_0^t W_s\,\mathrm{d}s. \tag{1.26}$$

Note that this is an integration-by-parts formula, cf. (1.4). Similarly, by setting $g(x,t) = h(t)\cdot x$, (1.4) can be established for suitable h.[3]

(b) Applying Itô's formula to $g(W_t)$ for $g(x) = x^2/2$, we easily obtain (1.19).

(c) Applying Itô's formula with $g(x,t) = \mathrm{e}^{x-t/2}$ shows that $Z_t = Z_0 g(W_t, t)$ satisfies $\mathrm{d}Z_t = Z_t\,\mathrm{d}W_t$. The stochastic process Z_t is the so-called Doléans exponential of W_t. Note that Z_t is "the solution of the stochastic differential equation $\mathrm{d}Z_t = Z_t\,\mathrm{d}W_t$ ", which explains why Z_t is considered as the exponential of W_t.

Remark 1.6 (The multidimensional case). The definition of the stochastic integral in $\mathbb{R}$ can be easily extended to multidimensional Brownian motion and stochastic integrals taking values in some $\mathbb{R}^n$. Let $W_t = (W_t^{(1)}, \dots, W_t^{(k)})$ be a k-dimensional standard Brownian motion and assume that $h(s,\omega) = (h_{ij}(s,\omega))_{i\leqslant n, j\leqslant k}$ is a matrix-valued function, taking values in the set of $(n \times k)$-matrices. Then we can rely on our previous definition of the stochastic integral and define the ith component of the n-dimensional stochastic integral by

$$\sum_{j=1}^k \int_0^t h_{ij}(s)\,\mathrm{d}W_s^{(j)}, \tag{1.27}$$

provided each h_{ij} allows for stochastic integration in $\mathbb{R}$. The above mentioned properties of stochastic integrals carry over in the natural way. In particular, the isometry (1.8) now reads

$$\mathbb{E}\Big\{\Big\langle \int_0^t f(s)\,\mathrm{d}W_s, \int_0^t g(s)\,\mathrm{d}W_s\Big\rangle\Big\} = \int_0^t \mathbb{E}\big\{\mathrm{Tr}(f(s)g(s)^T)\big\}\,\mathrm{d}s \tag{1.28}$$

for integrands f and g taking values in the same space, and the covariance of stochastic integrals can be calculated as

$$\mathbb{E}\Big\{\Big(\int_0^t f(s)\,\mathrm{d}W_s\Big)\Big(\int_0^t g(s)\,\mathrm{d}W_s\Big)^T\Big\} = \int_0^t \mathbb{E}\big\{f(s)g(s)^T\big\}\,\mathrm{d}s. \tag{1.29}$$

[3]Note that a particularly elegant proof of Itô's formula is based on an integration-by-parts formula such as (1.4) and (1.35). In that case, integration-by-parts has to be established first.

We conclude our discussion of the multidimensional case by stating the corresponding version of Itô's formula. As the multidimensional integral can be defined componentwise, it is sufficient to consider $Y_t = g(X_t, t)$ for multidimensional X_t and one-dimensional Y_t.

Theorem 1.7 (Itô's formula, multidimensional case). *Let $h : [0,\infty) \times \Omega \to \mathbb{R}^{n\times k}$ satisfy the measurability integrability assumptions for $\int_0^t h(s)\,\mathrm{d}W_s$ to be defined, and let $f : [0,\infty) \times \Omega \to \mathbb{R}^n$ be another function satisfying the same measurability assumptions and the weaker integrability assumption*

$$\mathbb{P}\left\{\int_0^t \|f(s,\omega)\|\,\mathrm{d}s < \infty \quad \text{for all } t \geqslant 0\right\} = 1. \tag{1.30}$$

We define X_t by

$$\mathrm{d}X_t = f(t)\,\mathrm{d}t + h(t)\,\mathrm{d}W_t, \tag{1.31}$$

and write $X_t^{(i)}$ for its ith component. Let $g : \mathbb{R}^n \times [0,T] \to \mathbb{R}$ be continuous and assume that the partial derivatives

$$g_t(x,t) = \frac{\partial}{\partial t} g(x,t), \qquad g_{x_i}(x,t) = \frac{\partial}{\partial x_i} g(x,t), \qquad g_{x_i x_j}(x,t) = \frac{\partial^2}{\partial x_i \partial x_j} g(x,t) \tag{1.32}$$

exist and are also continuous. Then $Y_t = g(X_t, t)$ is given by

$$\mathrm{d}Y_t = g_t(X_t,t)\,\mathrm{d}t + \sum_{i=1}^n g_{x_i}(X_t,t)\,\mathrm{d}X_t^{(i)} + \frac{1}{2}\sum_{i,j=1}^n g_{x_i x_j}(X_t,t)(\mathrm{d}X_t^{(i)})(\mathrm{d}X_t^{(j)}), \tag{1.33}$$

where $(\mathrm{d}X_t^{(i)})(\mathrm{d}X_t^{(j)})$ is calculated according to the scheme

$$(\mathrm{d}t)^2 = (\mathrm{d}t)(\mathrm{d}W_t^{(\mu)}) = (\mathrm{d}W_t^{(\mu)})(\mathrm{d}t) = 0, \qquad (\mathrm{d}W_t^{(\mu)})(\mathrm{d}W_t^{(\nu)}) = \delta_{\mu\nu}\,\mathrm{d}t. \tag{1.34}$$

As a consequence of this multidimensional version of Itô's formula, we immediately obtain the following integration-by-parts formula.

Lemma 1.8. *Let $\mathrm{d}X_t^{(i)} = f_i\,\mathrm{d}t + h_i\,\mathrm{d}W_t$, $i = 1,2$. Then*

$$X_t^{(1)}X_t^{(2)} = X_0^{(1)}X_0^{(2)} + \int_0^t X_s^{(1)}\,\mathrm{d}X_s^{(2)} + \int_0^t X_s^{(2)}\,\mathrm{d}X_s^{(1)} + \int_0^t h_1(s)h_2(s)\,\mathrm{d}s. \tag{1.35}$$

Finally, there is also a version of Fubini's theorem for stochastic integrals.

Lemma 1.9. *Let $h : [0,T] \times \mathbb{R} \times \Omega \to \mathbb{R}$ be $\mathcal{B}([0,T]) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{F}$-measurable, and assume that for any $u \in \mathbb{R}$, the map $\omega \mapsto h(t,u,\omega)$ is $\mathcal{F}_t$-measurable for all $t \in [0,T]$. If*

$$\int_0^T \int_{\mathbb{R}} \mathbb{E}\{h(s,u)^2\}\,\mathrm{d}u\,\mathrm{d}s < \infty, \tag{1.36}$$

then the integrals

$$\int_{\mathbb{R}} \int_0^T h(s,u)\,\mathrm{d}W_s\,\mathrm{d}u \qquad \text{and} \qquad \int_0^T \int_{\mathbb{R}} h(s,u)\,\mathrm{d}u\,\mathrm{d}W_s \tag{1.37}$$

are well-defined and almost surely equal.

1.2 Stochastic differential equations

Let us now turn to our original motivation for defining stochastic integrals which was to give a meaning to equations such as

$$\dot{x}_t = f(x_t, t) + F(x_t, t)\dot{W}_t, \qquad t \in [0, T], \tag{1.38}$$

where $\dot{W}_t$ denotes white noise, as integral equations

$$x_t = x_0 + \int_0^t f(x_s, s)\,\mathrm{d}s + \int_0^t F(x_s, s)\,\mathrm{d}W_s, \tag{1.39}$$

where $\{W_t\}_{t\geqslant 0}$ is a Brownian motion and $\int_0^t F(x_s, s)\,\mathrm{d}W_s$ is the Itô integral. We shall focus on so-called *strong solutions*, i. e., we try to find a stochastic process $\{x_t\}_{t\geqslant 0}$ which solves (1.39) for a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a given Brownian motion $\{W_t\}_{t\geqslant 0}$. We will briefly state the usual conditions for the existence of such a solution and its uniqueness. For further reading we refer to the broad literature on the subject, see for instance the brief introduction in [25], and the corresponding chapter in [26].

Let $f, F : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ be jointly measurable deterministic functions of x and t. We assume that the initial condition x_0 is deterministic or is independent of the Brownian motion $\{W_t\}_{t\geqslant 0}$ and satisfies $\mathbb{E}\{x_0^2\} < \infty$. For $t \in [0, T]$, let $\mathcal{F}_t = \sigma\{x_0, W_s, s \leqslant t\}$.

Definition 1.10. *A stochastic process $\{x_t\}_{t\in[0,T]}$ is called a* (strong) solution *of the stochastic differential equation (SDE)* (1.39)*, if*

- *x_t is $\mathcal{F}_t$-measurable;*
- $\mathbb{P}\{\int_0^T \|f(x_s(\omega), s)\|\,\mathrm{d}s < \infty\} = 1$ *and* $\mathbb{P}\{\int_0^T \|F(x_s(\omega), s)\|^2\,\mathrm{d}s < \infty\} = 1$*;*
- *for any $t \in [0, T]$,* (1.39) *holds with probability* 1.

Note that the $\mathcal{F}_t$-measurability of x_t implies the $\mathcal{F}_t$-measurability of $\omega \mapsto f(x_t(\omega), t)$ and $\omega \mapsto F(x_t(\omega), t)$, so that in particular the stochastic integral in (1.39) is well-defined.

The following theorem states the standard result on existence and uniqueness of such a solution.

Theorem 1.11 (Existence and uniqueness of strong solutions). *Assume that there exists a constant K such that the following holds for all $t \in [0, T]$ and all $x, y \in \mathbb{R}$:*

- Lipschitz condition:

$$\|f(x, t) - f(y, t)\| + \|F(x, t) - F(y, t)\| \leqslant K\|x - y\|; \tag{1.40}$$

- bounded-growth condition:

$$\|f(x, t)\| + \|F(x, t)\| \leqslant K(1 + \|x\|). \tag{1.41}$$

Then the SDE (1.39) *has a (pathwise) unique almost surely continuous solution $\{x_t\}_{t\in[0,T]}$. Here uniqueness means that for any two almost surely continuous solutions $\{x_t\}_{t\in[0,T]}$ and $\{y_t\}_{t\in[0,T]}$,*

$$\mathbb{P}\left\{\sup_{0\leqslant t\leqslant T} \|x_t - y_t\| > 0\right\} = 0. \tag{1.42}$$

Remark 1.12. The bounded-growth condition (1.41) excludes explosions of the solution, as in the case of ordinary differential equations. Uniqueness can be proved using only the Lipschitz condition (1.40). The result remains valid when the Lipschitz condition is replaced by a local analogue. As in the deterministic case, these conditions are sufficient but not necessary, and many interesting problems are actually not covered by this simplest of all existence and uniqueness theorems. It can be proved by a stochastic version of Picard–Lindelöf iterations.

Example 1.13 (Linear SDEs). We will frequently approximate solutions of stochastic differential equations locally by solutions of a linearized equation. Linear SDEs can be solved easily: Consider for example the one-dimensional SDE

$$\mathrm{d}x_t = \big[a(t)x_t + b(t)\big]\,\mathrm{d}t + F(t)\,\mathrm{d}W_t. \tag{1.43}$$

Using the notations $\alpha(t,s) = \int_s^t a(u)\,\mathrm{d}u$ and $\alpha(t) = \alpha(t,0)$, we can write its solution as

$$x_t = x_0\,\mathrm{e}^{\alpha(t)} + \int_0^t b(s)\,\mathrm{e}^{\alpha(t,s)}\,\mathrm{d}s + \int_0^t F(s)\,\mathrm{e}^{\alpha(t,s)}\,\mathrm{d}W_s. \tag{1.44}$$

Indeed, (1.44) can be established by an application of Itô's formula: Let x_t be defined by (1.43) and set $y_t = \mathrm{e}^{-\alpha(t)}\,x_t$. Then $\mathrm{d}y_t = b(t)\,\mathrm{e}^{-\alpha(t)}\,\mathrm{d}t + F(t)\,\mathrm{e}^{-\alpha(t)}\,\mathrm{d}W_t$, which determines y_t, and (1.44) for x_t follows.

If x_0 is either deterministic or Gaussian, then x_t is Gaussian for all t. In any case,

$$\mathbb{E}\{x_t\} = \mathbb{E}\{x_0\}\,\mathrm{e}^{\alpha(t)} + \int_0^t b(s)\,\mathrm{e}^{\alpha(t,s)}\,\mathrm{d}s, \tag{1.45}$$

$$\mathrm{Var}\,\{x_t\} = \mathrm{Var}\,\{x_0\}\,\mathrm{e}^{2\alpha(t)} + \int_0^t F(s)^2\,\mathrm{e}^{2\alpha(t,s)}\,\mathrm{d}s, \tag{1.46}$$

where we used the fact that we always assume the initial condition to be independent of the Brownian motion.

As a special case, we obtain that

$$\mathrm{d}x_t = -a_0 x_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t \tag{1.47}$$

has the unique almost surely continuous solution

$$x_t = x_0\,\mathrm{e}^{-a_0 t} + \sigma\int_0^t \mathrm{e}^{-a_0(t-s)}\,\mathrm{d}W_s, \tag{1.48}$$

which is known as Ornstein–Uhlenbeck (velocity) process, modelling the velocity of a Brownian particle. In this context, $-a_0 x_t$ is the damping or frictional force.

2 Sample-path large deviations

2.1 Introduction

We are interested in the behaviour of solutions to SDEs in $\mathbb{R}^d$ of the form

$$\mathrm{d}x_t^\varepsilon = b(x_t^\varepsilon)\,\mathrm{d}t + \sqrt{\varepsilon}\,\mathrm{d}W_t, \qquad x_0^\varepsilon = x_0, \tag{2.1}$$

where the noise intensity $\sqrt{\varepsilon}$ is considered a small parameter. One would expect that for sufficiently small noise, x_t^ε is close to the solution of the corresponding deterministic ordinary differential equation (ODE)

$$\dot{x}_t = b(x_t) \tag{2.2}$$

with the same initial condition. Indeed, if b is Lipschitz continuous with Lipschitz constant L_b, then

$$\|x_t^\varepsilon - x_t\| \leqslant L_b \int_0^t \|x_s^\varepsilon - x_s\|\,\mathrm{d}s + \sqrt{\varepsilon}\|W_t\|, \tag{2.3}$$

and Gronwall's lemma shows that

$$\sup_{0\leqslant s\leqslant t} \|x_s^\varepsilon - x_s\| \leqslant \sqrt{\varepsilon} \sup_{0\leqslant s\leqslant t} \|W_s\|\,\mathrm{e}^{L_b t}. \tag{2.4}$$

Thus is is sufficient to estimate the probability that the Brownian motion W_s leaves a ball of some radius r before time t, an estimate which is standard. Using the reflection principle for the one-dimensional Brownian motion, we find that the probability of a sample path x_s^ε leaving a δ-neighbourhood of the deterministic solution x_s satisfies

$$\mathbb{P}\Big\{\sup_{0\leqslant s\leqslant t} \|x_s^\varepsilon - x_s\| \geqslant \delta\Big\} \leqslant \mathbb{P}\Big\{\sup_{0\leqslant s\leqslant t} \|W_s\| \geqslant \frac{\delta}{\sqrt{\varepsilon}}\,\mathrm{e}^{-L_b t}\Big\} \leqslant 2d\exp\Big\{-\frac{\delta^2\,\mathrm{e}^{-2L_b t}}{2\varepsilon d t}\Big\}. \tag{2.5}$$

(Note that the dependence on the dimension d in this estimate is not optimal and can easily be improved.) As expected, for fixed ε, the probability of leaving a neighbourhood of the deterministic solution increases with t, as typical path will eventually leave any fixed neighbourhood, and it decreases with increasing δ as it is more difficult to leave larger neighbourhoods. We also see that the smaller the noise intensity $\sqrt{\varepsilon}$, the more difficult it is to leave a neighbourhood of x_t. In order for the probability to be small, we need δ^2/ε to be larger than $(1/t)\,\mathrm{e}^{2L_b t}$ (neglecting the dependence on the dimension d). Thus we observe an exponential decay of the probability of leaving a neighbourhood of x_s.

Let now A be a measurable subset of the set $\mathcal{C} = \mathcal{C}([0,T],\mathbb{R}^d)$ of all continuous functions $[0,T] \to \mathbb{R}^d$, where measurable refers to the Borel σ-algebra on $\mathcal{C}$. Assume that A does *not* contain the δ-neighbourhood of the deterministic solution considered above. Then we know that the path $x^\varepsilon = (x_s^\varepsilon)_{0\leqslant s\leqslant T}$ satisfies

$$\mathbb{P}\{x^\varepsilon \in A\} \leqslant 2d\exp\Big\{-\frac{\delta^2\,\mathrm{e}^{-2L_b T}}{2\varepsilon d T}\Big\} \to 0 \qquad \text{as } \varepsilon \to 0, \tag{2.6}$$

and we consider the event that $x^\varepsilon \in A$ as a *large deviation* as x^ε typically fluctuates around the deterministic solution and we are now looking at atypical behaviour. Our aim is to obtain the *rate* at which the probability in (2.6) tends to zero as a function of the set A. (Note that we used a rough estimate, neglecting the actual choice of A.) In general it is not possible to obtain the exact rate, but at least the exponential rate can be found. Thus we are looking for a *rate function* or *(normalized) action functional* $I : \mathcal{C} \to [0, \infty]$ such that

$$\mathbb{P}\{\|x^\varepsilon - \varphi\|_\infty < \delta\} \approx \mathrm{e}^{-I(\varphi)/\varepsilon}, \tag{2.7}$$

where $\|\cdot\|_\infty$ denotes the supremum norm on the time interval $[0, T]$. This describes the case of A being a small ball around some continuous function φ.

Below we will first consider the special case of a scaled Brownian motion $x_t^\varepsilon = \sqrt{\varepsilon} W_t$, then discuss some general principles and finally generalize to the situation (2.1), where we will actually allow for a larger class of diffusion coefficients. Our presentation mainly follows [15] and [13]. We will also comment on the different approach chosen in [12].

2.2 Sample-path large deviations for Brownian motion: Schilder's Theorem

In this section we consider the special case $x_t^\varepsilon = \sqrt{\varepsilon} W_t$ of a scaled Brownian motion (in $\mathbb{R}^d$), with $x_0^\varepsilon = 0$. We will use the following notations:

Notations 2.1. Fix a time interval $[0,T]$.

- $\mathcal{C} := \mathcal{C}([0,T], \mathbb{R}^d)$ for the set of all continuous functions $[0,T] \to \mathbb{R}^d$,
- $\mathcal{C}_0 := \{\varphi \in \mathcal{C} \colon \varphi_0 = 0\}$,
- $\|\varphi\|_\infty = \|\varphi\|_{[0,T]} := \sup_{t\in[0,T]} \|\varphi_t\|$ for $\varphi \in \mathcal{C}$ or $\mathcal{C}_0$,
- $\mathcal{L}_2 = \mathcal{L}_2([0,T], \mathbb{R}^d)$ for the set of all square-integrable functions $[0,T] \to \mathbb{R}^d$,
- $H_1 = H_1([0,T], \mathbb{R}^d) := \left\{ \int_0^T f(s)\, \mathrm{d}s \colon f \in \mathcal{L}_2 \right\}$
  for the set of all absolutely continuous functions $\varphi : [0,T] \to \mathbb{R}^d$, with square-integrable derivative $\dot\varphi$ and $\varphi_0 = 0$,
- $\|\varphi\|_{H_1} := \left(\int_0^T \|\dot\varphi_s\|^2 \,\mathrm{d}s \right)^{1/2}$ for $\varphi \in H_1$.

Our aim is the following result, due to Schilder (1966):

Theorem 2.2 (Schilder's Theorem). *The family of induced measures* $\{\mathbb{P}(x^\varepsilon)^{-1}\}_{\varepsilon>0}$ *on* $\mathcal{C}_0$*, equipped with the the Borel* σ*-algebra, satisfies a* large deviation principle *with the* good *rate function*

$$I(\varphi) = I^{\mathrm{BM}}_{[0,T],0}(\varphi) = \begin{cases} \frac12 \|\varphi\|^2_{H_1}, & \text{if } \varphi \in H_1, \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.8}$$

That is, the relation

$$-\inf_{\Gamma^\circ} I \leqslant \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon \in \Gamma\} \leqslant \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon \in \Gamma\} \leqslant -\inf_{\overline{\Gamma}} I, \tag{2.9}$$

holds for all Borel sets $\Gamma \subset \mathcal{C}_0$*, where* Γ° *and* $\overline{\Gamma}$ *denote interior or closure of* Γ*, respectively, and* I *is lower semi-continuous with* compact *level sets.*

Remark 2.3.

(a) Relation (2.9) is called a large deviation principle (LDP). It consists of two parts, a lower and an upper bound, the middle one of the estimates being trivial. A function $I : \mathcal{C}_0 \to [0, +\infty]$ is called a rate function if it is lower semi-continuous. Then the level sets $\{\varphi \colon I(\varphi) \leqslant \alpha\}$ are closed. If the level sets are compact, we call I a good rate function.

(b) Note that $x^\varepsilon \notin H_1$ almost surely, as the paths of a Brownian motion are almost surely of unbounded variation. Thus $I(x^\varepsilon) = +\infty$ almost surely, for *all* $\varepsilon > 0$. Let $\Gamma = \mathcal{C}_0$ for the moment. Then $\inf_\Gamma I = 0$ and I being a good rate function implies that there exists a φ such that $I(\varphi) = 0$. Of course, in the context of Schilder's Theorem, this is trivial as we can choose $\varphi_t \equiv 0$. Thus we see again that x^ε concentrates near the function which vanishes everywhere. The LDP is useful to estimate the probability of rare events, namely that x^ε is *not* close to $\varphi_t \equiv 0$.

(c) Note also that in general, the limits in (2.9) are not equal, so that

$$\lim_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon \in \Gamma\} \tag{2.10}$$

does not necessarily exist and we cannot do better than stating bounds. If the limit

$$\lim_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon \in \Gamma\} = -\inf_{\Gamma} I, \tag{2.11}$$

does exist, then Γ is called an I-continuity set.

(d) In the statement of the LDP (2.9), the interior and the closure of Γ are needed for lower and upper bound to hold. It is easy to see that the interior is actually needed for the lower bound: For non-atomic measures such as $\mathbb{P}(x^\varepsilon)^{-1}$, $\mathbb{P}\{x^\varepsilon = \varphi\} = 0$ holds for all φ. Would the lower bound be valid with Γ° replaced by Γ, then

$$-I(\varphi) = -\inf_{\{\varphi\}} I \leqslant \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon = \varphi\} = -\infty \tag{2.12}$$

would imply $I(\varphi) = +\infty$ for all φ, which contradicts the upper bound, as

$$0 = \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon \in \mathcal{C}_0\} \leqslant -\inf_{\mathcal{C}_0} I = -\infty \tag{2.13}$$

cannot hold. Note that this argument uses only that $\mathbb{P}(x^\varepsilon)^{-1}$ is non-atomic.

Example 2.4. As an example let us consider the probability that x^ε leaves a ball of radius δ around the origin. As the typical spreading of Brownian motion scales with $t^{1/2}$, we expect x^ε to remain in $B(0,\delta)$, provided $\delta^2/\varepsilon \gg T$. Let us make this precise. Set $\Gamma = \{\varphi \in \mathcal{C}_0 \colon \|\varphi\|_\infty < \delta\} = B(0,\delta)$. Then $\inf_{\Gamma^c} I$ is obtained for any φ of the form $\varphi_s = sx/T$ with $\|x\| = \delta$, and the LDP states that $\mathbb{P}\{x^\varepsilon \notin \Gamma\}$ decays like $\mathrm{e}^{-\delta^2/2\varepsilon T}$, which is small for $\delta^2/\varepsilon \gg T$, as expected.

As it is easy to see that I as defined in (2.8) is a good rate function, i.e., I is lower semi-continuous with compact level sets, we will focus on the proof of relation (2.9). Our presentation follows [15, Section 3.2]. We will establish two lemmas, providing the fundamental estimates for upper and lower bound, respectively, and show how Schilder's theorem follows from these lemmas.

The first of these lemmas gives a lower bound on the probability of x^ε remaining in a ball. Note that this bound depends only on the centre of the ball (and on ε).

Lemma 2.5. *For all $\delta > 0$, all $\gamma > 0$ and all $K > 0$, there exists an $\varepsilon_0 = \varepsilon_0(\delta,\gamma,K,T) > 0$ such that for all $\varepsilon \leq \varepsilon_0$ and all $\varphi \in \mathcal{C}_0$ with $I(\varphi) < K$, we have*

$$\mathbb{P}\{\|x^\varepsilon - \varphi\|_\infty < \delta\} \geqslant \mathrm{e}^{-[I(\varphi)+\gamma]/\varepsilon}. \tag{2.14}$$

Before proving the preceding lemma, let us show that it implies the lower bound in Schilder's theorem. (Actually, Lemma 2.5 is equivalent to the lower bound in (2.9).)

PROOF OF THE LOWER BOUND IN THEOREM 2.2. Let G be an arbitrary open set in $\mathcal{C}_0$. If $\inf_G I = +\infty$, the lower bound in (2.9) is trivial. Thus we may assume that $\inf_G I < \infty$, in which case we may choose a $\varphi \in G$ such that $I(\varphi) < \infty$. In addition, we find a radius $r_\varphi > 0$ such that the ball $B(\varphi, r_\varphi)$ of radius r_φ, centred in φ, is contained in G. Now

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon \in G\} \geqslant \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon \in B(\varphi, r_\varphi)\} \geqslant -I(\varphi) \tag{2.15}$$

follows directly from (2.14). All that remains to do is to take the infimum over all $\varphi \in G$. □

The proof of Lemma 2.5 is standard. It makes use of the Cameron–Martin or Girsanov formula which allows to change the drift term in an SDE by a suitable change of measure. Here we only need a particularly simple version which states that for $h \in \mathcal{L}_2$ the distribution of the stochastic process $\widehat{W}_t$, given by

$$\mathrm{d}\widehat{W}_t = \mathrm{d}W_t - h(t)\,\mathrm{d}t \qquad \text{resp.} \qquad \widehat{W}_t = W_t - \int_0^t h(s)\,\mathrm{d}s, \tag{2.16}$$

considered on the σ-algebra $\mathcal{F}_t = \sigma\{W_s \colon s \leqslant t\}$, generated by the Brownian motion W_s up to time t, has a density with respect to $\mathbb{P}$ which is given by the Radon–Nikodym derivative

$$\frac{\mathrm{d}\mathbb{P}\widehat{W}^{-1}}{\mathrm{d}\mathbb{P}} = \exp\biggl\{-\int_0^t \langle h(s), \mathrm{d}W_s\rangle - \frac{1}{2}\int_0^t \|h(s)\|^2\,\mathrm{d}s\biggr\} \qquad \text{on } \mathcal{F}_t. \tag{2.17}$$

PROOF OF LEMMA 2.5. Fix $\delta, \gamma, K > 0$ and a function φ with $I(\varphi) \leqslant K$. We want to estimate

$$\mathbb{P}\{\|x^\varepsilon - \varphi\|_\infty < \delta\} = \mathbb{P}\Bigl\{\Bigl\|W - \frac{\varphi}{\sqrt{\varepsilon}}\Bigr\|_\infty < \frac{\delta}{\sqrt{\varepsilon}}\Bigr\} = \mathbb{P}\{\|\widehat{W}\|_\infty < \delta/\sqrt{\varepsilon}\}, \tag{2.18}$$

with $\widehat{W}_t = W_t - \varphi_t/\sqrt{\varepsilon}$. By the Cameron–Martin formula (2.17), we can rewrite the probability as

$$\mathbb{P}\{\|x^\varepsilon - \varphi\|_\infty < \delta\} = \mathrm{e}^{-\frac{1}{2\varepsilon}\int_0^T \|\dot\varphi_s\|^2\,\mathrm{d}s} \int_{\{W \in B(0,\delta/\sqrt{\varepsilon})\}} \mathrm{e}^{-\frac{1}{\sqrt{\varepsilon}}\int_0^T \langle\dot\varphi_s, \mathrm{d}W_s\rangle}\,\mathrm{d}\mathbb{P}. \tag{2.19}$$

Now we split the domain of integration into a "good" part and a "bad" one. As we want a sufficiently precise lower estimate, the "bad" set is the one where the exponent becomes unusually small. For the bad set we choose

$$A_x := \biggl\{-\frac{1}{\sqrt{\varepsilon}}\int_0^T \langle\dot\varphi_s, \mathrm{d}W_s\rangle \leqslant -x\biggr\} \tag{2.20}$$

with $x = \sqrt{4I(\varphi)/\varepsilon}$. In order to be able to neglect A_x, we need an estimate on its probability:

$$\mathbb{P}\{A_x\} = \frac{1}{2}\mathbb{P}\biggl\{\biggl|\frac{1}{\sqrt{\varepsilon}}\int_0^T \langle\dot\varphi_s, \mathrm{d}W_s\rangle\biggr| \geqslant x\biggr\} \leqslant \frac{1}{2\varepsilon x^2}\mathbb{E}\biggl\{\biggl(\int_0^T \langle\dot\varphi_s, \mathrm{d}W_s\rangle\biggr)^2\biggr\} \tag{2.21}$$

by Tshebychev's inequality, and

$$\mathbb{P}\{A_x\} \leqslant \frac{1}{2\varepsilon x^2}\int_0^T \|\dot\varphi_s\|^2\,\mathrm{d}s = \frac{1}{\varepsilon x^2} I(\varphi) = \frac{1}{4} \tag{2.22}$$

follows by our choice of x. Now, on A_x^{c},

$$\begin{aligned}\mathbb{P}\{\|x^\varepsilon - \varphi\|_\infty < \delta\} &\geqslant \mathrm{e}^{-I(\varphi)/\varepsilon}\,\mathrm{e}^{-x}\,\mathbb{P}\{W \in B(0,\delta/\sqrt{\varepsilon}) \cap A_x^{\mathrm{c}}\} \\ &\geqslant \mathrm{e}^{-I(\varphi)/\varepsilon}\,\mathrm{e}^{-x}\bigl[\mathbb{P}\{A_x^{\mathrm{c}}\} - \mathbb{P}\{W \in B(0,\delta/\sqrt{\varepsilon})^{\mathrm{c}}\}\bigr].\end{aligned} \tag{2.23}$$

As $\mathbb{P}\{A_x^{\mathrm{c}}\} \geqslant 3/4$, choosing ε small enough assures that $\mathbb{P}\{A_x^{\mathrm{c}}\} - \mathbb{P}\{W \in B(0,\delta/\sqrt{\varepsilon})^{\mathrm{c}}\} \geqslant 1/2$, and by our choice of x,

$$\mathbb{P}\{\|x^\varepsilon - \varphi\|_\infty < \delta\} \geqslant \mathrm{e}^{-[I(\varphi)+\gamma]/\varepsilon} \tag{2.24}$$

follows for all $\varepsilon \leqslant \varepsilon_0(\delta, \gamma, K, T)$. □

Let us now turn to the upper bound in Schilder's Theorem. The upper bound is a direct consequence on the following upper bound on the probability that x^ε leaves a neighbourhood of a level set

$$\Phi(\alpha) := \{\varphi \in \mathcal{C}_0 \colon I(\varphi) \leqslant \alpha\} \tag{2.25}$$

of the rate function I. Level sets are special neighbourhoods of the function vanishing everywhere, so the event of x^ε leaving a neighbourhood of a level set represents the type of events we are interested in.

Lemma 2.6. *For all $\delta > 0$, all $\gamma > 0$ and all $\alpha_0 > 0$, there exists an $\varepsilon_0 > 0$ such that for all $\varepsilon \leqslant \varepsilon_0$ and all $\alpha \leqslant \alpha_0$,*

$$\mathbb{P}\{\operatorname{dist}(x^\varepsilon, \Phi(\alpha)) \geqslant \delta\} \leqslant \mathrm{e}^{-[\alpha-\gamma]/\varepsilon}, \tag{2.26}$$

where $\operatorname{dist}(\varphi, \Phi(\alpha))$ *denotes the distance of a function φ from the level set $\Phi(\alpha)$, measured with respect to the supremum norm.*

Choosing $F = \{\varphi \colon \operatorname{dist}(\varphi, \Phi(\alpha)) \geqslant \delta\}$ shows that the upper bound in Schilder's Theorem implies the preceding lemma. All what is needed is to pay some attention to the statement's uniformity in α. The inverse is also true (for good rate functions):

PROOF OF THE UPPER BOUND IN THEOREM 2.2. Choose an arbitrary closed set F. Without loss of generality, we may assume that $\inf_F I > 0$, so that we may choose a $\gamma > 0$ such that $\alpha := \inf_F I - \gamma > 0$. As I is a good rate function, the level set $\Phi(\alpha)$ is compact and, by our choice of α, its intersection $\Phi(\alpha) \cap F$ with the closed set F is empty. Thus $\delta := \operatorname{dist}(\Phi(\alpha), F) > 0$. Now Lemma 2.6 implies

$$\mathbb{P}\{x^\varepsilon \in F\} \leqslant \mathbb{P}\{\operatorname{dist}(x^\varepsilon, \Phi(\alpha)) \geqslant \delta\} \leqslant \mathrm{e}^{-[\inf_F I - 2\gamma]/\varepsilon}, \tag{2.27}$$

which completes the proof. □

Let us now prove Lemma 2.6 which estimates the probability that x^ε leaves a neighbourhood of a level set. The main problem in the proof is the fact that $I(x^\varepsilon) = +\infty$. Thus we need to approximate x^ε by smoother functions. We will choose random polygons. Note that we might also choose random step functions, for instance.

PROOF OF LEMMA 2.6. As already mentioned, we want to approximate x^ε by a random polygon. For a given spacing $\Delta > 0$ to be chosen later, we denote by $x^{n,\varepsilon}$ the polygon with vertices

$$(0,0), (\Delta, x^{n,\varepsilon}_\Delta), (2\Delta, x^{n,\varepsilon}_{2\Delta}), \dots, (T, x^{n,\varepsilon}_T). \tag{2.28}$$

For convenience, we assume that Δ has been chosen such that $T/\Delta \in \mathbb{N}$. When estimating the probability that x^ε leaves a neighbourhood of a level set of the rate function, there are two possibilities to be taken into account: either $x^{n,\varepsilon}$ is not a good approximation to x^ε, which we will show to be unlikely, or $x^{n,\varepsilon}$ is a good approximation and then x^ε leaving a neighbourhood of a level set implies that $x^{n,\varepsilon}$ at least leaves the level set itself. Thus,

$$\mathbb{P}\{\operatorname{dist}(x^\varepsilon, \Phi(\alpha)) \geqslant \delta\} \leqslant \mathbb{P}\{\|x^\varepsilon - x^{n,\varepsilon}\|_\infty \geqslant \delta\} + \mathbb{P}\{I(x^{n,\varepsilon}) > \alpha\}. \tag{2.29}$$

Using the fact that the increments $x_s^\varepsilon - x_s^{n,\varepsilon}$, considered on different time intervals $[k\Delta, (k+1)\Delta)$, are identically distributed and combining a crude and a standard estimate, we find the bound

$$\begin{aligned}
\mathbb{P}\{\|x^\varepsilon - x^{n,\varepsilon}\|_\infty \geqslant \delta\} &= \mathbb{P}\Big\{\sup_{0\leqslant s\leqslant T}\|x_s^\varepsilon - x_s^{n,\varepsilon}\| \geqslant \delta\Big\} \leqslant \frac{T}{\Delta}\mathbb{P}\Big\{\sup_{0\leqslant s\leqslant \Delta}\|x_s^\varepsilon - x_s^{n,\varepsilon}\| \geqslant \delta\Big\} \\
&\leqslant \frac{T}{\Delta}\mathbb{P}\Big\{\sup_{0\leqslant s\leqslant \Delta}\|x_s^\varepsilon\| \geqslant \delta\Big\} = \frac{T}{\Delta}\mathbb{P}\Big\{\sup_{0\leqslant s\leqslant \Delta}\|W_s\| \geqslant \frac{\delta}{\sqrt{\varepsilon}}\Big\} \qquad (2.30)\\
&\leqslant \frac{2dT}{\Delta}\,\mathrm{e}^{-\delta^2/2d\varepsilon\Delta} \qquad (2.31)
\end{aligned}$$

on the quality of the approximation. Choosing $\Delta = \delta^2/2d\alpha_0$,

$$\mathbb{P}\{\|x^\varepsilon - x^{n,\varepsilon}\|_\infty \geqslant \delta\} \leqslant \frac{1}{2}\,\mathrm{e}^{-[\alpha_0-\gamma]/\varepsilon} \qquad (2.32)$$

follows for all $\varepsilon \leqslant \varepsilon_0(T, \delta, \gamma, \alpha_0)$.

Now we will also estimate the second term in (2.29) which gives the probability that the approximation $x^{n,\varepsilon}$ leaves the level set. Using the fact that $x^{n,\varepsilon}$ is a polygon, we see that

$$I(x^{n,\varepsilon}) = \frac{1}{2}\sum_{l=1}^{T/\Delta}\int_{(l-1)\Delta}^{l\Delta}\frac{\|\sqrt{\varepsilon}W_{l\Delta} - \sqrt{\varepsilon}W_{(l-1)\Delta}\|^2}{\Delta^2}\,\mathrm{d}s = \frac{\varepsilon}{2}\sum_{l=1}^{T/\Delta}\frac{\|W_{l\Delta} - W_{(l-1)\Delta}\|^2}{\Delta}. \qquad (2.33)$$

The sum on the right-hand side equals in distribution the sum $\sum \xi_i^2$ over the squares of dT/Δ independent one-dimensional standard-normal random variables ξ_i, which can be estimated by Tshebychev's inequality, yielding

$$\mathbb{P}\{I(x^{n,\varepsilon}) > \alpha\} = \mathbb{P}\Big\{\sum_{i=1}^{dT/\Delta}\xi_i^2 > \frac{2\alpha}{\varepsilon}\Big\} \leqslant \mathrm{e}^{-2\kappa\alpha/\varepsilon}\big(\mathbb{E}e^{\kappa\xi_1^2}\big)^{dT/\Delta} = (1-2\kappa)^{-dT/2\Delta}\,\mathrm{e}^{-2\kappa\alpha/\varepsilon} \qquad (2.34)$$

for any $\kappa < 1/2$. We choose $\kappa = (1-\gamma/2\alpha)/2$. Then, for ε small enough,

$$\mathbb{P}\{I(x^{n,\varepsilon}) > \alpha\} \leqslant \frac{1}{2}\,\mathrm{e}^{-[\alpha-\gamma]/\varepsilon} \qquad (2.35)$$

follows, and the lemma is proved. □

In [12], a slightly different proof is chosen, namely, $x^\varepsilon = \sqrt{\varepsilon}W$ is approximated by random step functions $t \mapsto \widehat{x}_t^\varepsilon := \sqrt{\varepsilon}W_{\varepsilon[t/\varepsilon]}$. Then the proof of Schilder's Theorem reduces to showing that the approximation is sufficiently good, i. e.,

$$\limsup_{\varepsilon\to 0}\varepsilon\log\mathbb{P}\{\|x^\varepsilon - \widehat{x}^\varepsilon\|_\infty \geqslant \delta\} = -\infty, \qquad (2.36)$$

to allow to extend Cramer's Theorem which states a LDP for sums of independent identically distributed random variables.

A more abstract proof is presented in [13], which relies more heavily on properties of Wiener measure and is based on the use of logarithmic moment-generating functions, the Fenchel–Legendre transform of which are natural candidates for rate functions, see the remarks on general concepts in the next section.

2.3 Some general remarks on large deviation principles

In this section we summarize some aspects of large deviation theory in order to shed some light on how the notions and results presented above are special cases of general concepts. For details and proofs see [13], [12].

Let E be a separable metric space, equipped with a σ-algebra $\mathcal{E}$. We do *not* necessarily assume that $\mathcal{E}$ is the Borel σ-algebra. On $(E, \mathcal{E})$ we consider a family of probability measures $\{\mu_\varepsilon\}_{\varepsilon>0}$.

Definition 2.7.

- *A function $I : E \to [0, \infty]$ is called* rate function *if it is lower semi-continuous.*
- *A rate function I is called* good *if its level sets $\{x \in E \colon I(x) \leqslant \alpha\}$ are compact.*
- *The family $\{\mu_\varepsilon\}_{\varepsilon>0}$ of probability measures is said to satisfy the* (full) large deviation principle *with rate function I, if*

$$-\inf_{\Gamma^\circ} I \leqslant \liminf_{\varepsilon\to0} \varepsilon \log \mu_\varepsilon(\Gamma) \leqslant \limsup_{\varepsilon\to0} \varepsilon \log \mu_\varepsilon(\Gamma) \leqslant -\inf_{\overline{\Gamma}} I \tag{2.37}$$

  holds for all $\Gamma \in \mathcal{E}$, where Γ° and $\overline{\Gamma}$ denote the interior and the closure of Γ, respectively.

- *The family $\{\mu_\varepsilon\}_{\varepsilon>0}$ of probability measures is said to satisfy a* weak *large deviation principle with rate function I, if the upper bound in* (2.37) *holds for all (pre-)compact sets Γ.*

Remark 2.8.

- Rate functions being lower semi-continuous, their level sets are closed, while good rate functions even enjoy compact level sets. As a consequence, for good rate functions, the infimum over (non-empty) closed sets is achieved. For good rate functions,

$$\inf_F I = \lim_{\delta\searrow0} \inf_{F^{(\delta)}} I \tag{2.38}$$

  holds for all closed sets F, where $F^{(\delta)}$ denotes the δ-neighbourhood of F.

- Recall from the previous section that in general, limes inferior and limes superior in (2.37) do not coincide. If they do coincide for a set Γ, we call Γ an *I-continuity set.*
- The rate function is unique in the sense that for a given family of measures and a given scale ε, there exists at most one rate function.
- If $\mathcal{E} = \mathcal{B}(E)$ (Borel sets) and the family of measures $\{\mu_\varepsilon\}_{\varepsilon>0}$ is exponentially tight, then the validity of a weak LDP already implies that a (full) LDP with a good rate function holds. (Recall that $\{\mu_\varepsilon\}_{\varepsilon>0}$ is called *exponentially tight* if for any $\alpha < \infty$ there exists a compact set K_α such that

$$\limsup_{\varepsilon\to0} \varepsilon \log \mu_\varepsilon(K_\alpha^{\mathrm{c}}) < -\alpha, \tag{2.39}$$

  meaning that the complement of K_α "is not seen" on a logarithmic scale.)

Example 2.9. Let us give an example of a family of measures satisfying a weak LDP but not satisfying a full LDP. Simply choose $\mu_n = \delta_n$ on $\mathcal{B}(\mathbb{R})$, that is, μ_n is the Dirac measure concentrated in $n \in \mathbb{N}$. The $\{\mu_n\}_n$ satisfies a weak LDP with scale $\varepsilon = 1/n$ and a good rate function I. Indeed, for a compact set F and sufficiently large n, $\mu_n(F) = 0$, so that for compact sets, the upper bound in (2.37) holds with $I(x) = \infty$ for all x. But for $I(x) \equiv \infty$, the lower bound is trivial.

A full LDP cannot hold, as the choice $F = [1, \infty)$ shows:

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n([1,\infty)) = 0 > -\infty = -\inf_{[1,\infty)} I. \tag{2.40}$$

In the previous section, we have already seen that there are equivalent formulations for the upper and the lower bound in (2.37), which are more convenient to establish.

Lemma 2.10.

- *The upper bound in* (2.37) *is equivalent to the following statement:*
  $\forall \alpha < \infty\ \forall \Gamma \in \mathcal{E}$ *with* $\overline{\Gamma} \subset \Phi(\alpha)^c = \{x \in E \colon I(x) \leqslant \alpha\}^c$,
  $$\limsup_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(\Gamma) \leqslant -\alpha. \tag{2.41}$$
- *The lower bound in* (2.37) *is equivalent to the following statement:*
  $\forall x$ *with* $I(x) < \infty\ \forall \Gamma \in \mathcal{E}$ *with* $x \in \Gamma^\circ$,
  $$\liminf_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(\Gamma) \geqslant -I(x). \tag{2.42}$$

Corollary 2.11. *For any good rate function* I, *for all* x, $I(x)$ *can be calculated using*

$$I(x) = -\lim_{\delta\searrow 0} \liminf_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(B(x,\delta)) = -\lim_{\delta\searrow 0} \limsup_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(B(x,\delta)), \tag{2.43}$$

where $B(x,\delta)$ *denotes the ball of radius* δ *around* x.

PROOF: Fix x. Then on the one hand, (2.42) shows that

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(B(x,\delta)) \geqslant -I(x) \tag{2.44}$$

for all δ. On the other hand, the usual large-deviation upper bound shows

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(B(x,\delta)) \leqslant -\inf_{\overline{B(x,\delta)}} I \leqslant -\inf_{\{x\}^{(2\delta)}} I \to -I(x) \tag{2.45}$$

as $\delta \to 0$, where we used that good rate functions satisfy

$$I(x) = \inf_{\{x\}} I = \lim_{\delta\searrow 0} \inf_{\{x\}^{(2\delta)}} I. \tag{2.46}$$

□

Here a word of warning is due. The validity of (2.43) for all x does *not* imply a LDP. A counterexample can be found in [15, p. 90], for example.

The preceding lemma shows a way to find the rate function, if the limits can be calculated. Another approach which is known to be successful for a large class of problems uses convex analysis:

Definition 2.12. *Let E be a separable Banach space.*[4] *For a probability measure μ on E, the* logarithmic moment-generating function *is the map*

$$E^* \ni \lambda \mapsto \Lambda_\mu(\lambda) := \log \int_E e^{\langle \lambda, x\rangle}\, \mu(dx) \in (-\infty, \infty], \tag{2.47}$$

where E^* is the topological dual of E and $\langle \lambda, x\rangle$ denotes the duality relation.

Assume that

$$\Lambda(\lambda) := \lim_{\varepsilon\to 0} \varepsilon \Lambda_{\mu_\varepsilon}(\lambda/\varepsilon) \tag{2.48}$$

exists in $[-\infty, \infty]$, for all $\lambda \in E^*$. Then Λ is a convex function on E^*, and its *Fenchel–Legendre transform*

$$E \ni x \mapsto \Lambda^*(x) := \sup_{\lambda\in E^*} \{\langle \lambda, x\rangle - \Lambda(\lambda)\} \tag{2.49}$$

is non-negative, lower semi-continuous and convex. For all *compact* $F \subset E$, we have the upper large-deviation bound

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(F) \leqslant -\inf_F \Lambda^*. \tag{2.50}$$

(Recall that for an exponentially tight family of measures, (2.50) will hold for all closed F.)

Thus, Λ^* is a candidate for the rate function. But note that rate functions are not necessarily convex. In that case, Λ^* cannot be the rate function. But if a rate function is good and convex, and some moment condition is satisfied, then the limit (2.48) defining Λ exists, and Λ^* is the rate function. For details, see for instance [13, Section 2.2].

Let us conclude this section with Varadhan's Lemma.

Theorem 2.13 (Varadhan's Lemma). *Let E be a separable metric space and assume that $\{\mu_\varepsilon\}_\varepsilon$ satisfies a LDP with a good rate function I and that the function $\phi : E \to \mathbb{R}$ is continuous. Assume further that the following* tail condition *holds*

$$\lim_{L\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \int_{\{x:\, \phi(x)\geqslant L\}} e^{\phi(x)/\varepsilon}\, \mu_\varepsilon(dx) = -\infty. \tag{2.51}$$

Then

$$\lim_{\varepsilon\to 0} \varepsilon \log \int e^{\phi(x)/\varepsilon}\, \mu_\varepsilon(dx) = \sup_E [\phi - I]. \tag{2.52}$$

In particular, if the following moment condition holds for some $\alpha \in (1, \infty)$,

$$\sup_{0<\varepsilon\leqslant 1} \left(\int e^{\alpha\phi(x)/\varepsilon}\, \mu_\varepsilon(dx) \right)^{\varepsilon} < \infty, \tag{2.53}$$

then the tail condition (2.51) *is satisfied.*

Varadhan's Lemma can be seen as an infinite-dimensional version of the Laplace method. In order to see this, we assume for the moment that $E = \mathbb{R}$ and that the measures μ_ε have densities $d\mu_\varepsilon / dx \sim e^{-I(x)/\varepsilon}$ with respect to Lebesgue measure. Then

$$\int_{\mathbb{R}} e^{\Phi(x)/\varepsilon}\, \mu_\varepsilon(dx) \sim \int_{\mathbb{R}} e^{[\Phi(x)-I(x)]/\varepsilon}\, dx \sim e^{\sup_x[\Phi(x)-I(x)]/\varepsilon} \tag{2.54}$$

by the Laplace method, which is the statement of Varadhan's Lemma.

[4] This condition can be relaxed, see [13, Section 2.2].

2.4 The contraction principle

The contraction principle is a tool to obtain new LDP's from an established one. There is a trivial version, which we state first.

Lemma 2.14 (The contraction principle – trivial version). *Let X, Y be regular Hausdorff spaces, $I : X \to [0, \infty]$ a -/*good *rate function and $f : X \to Y$ a continuous function. Define*

$$I'(y) = \inf\{I(x) \colon x \in X \text{ with } y = f(x)\}. \tag{2.55}$$

Then

- *I' is a good rate function on Y.*
- *If I governs a LDP for $\{\mu_\varepsilon\}_{\varepsilon>0}$, then I' governs a LDP for the induced measures $\{\mu_\varepsilon f^{-1}\}_{\varepsilon>0}$.*

Remark 2.15.

- The contraction principle can in particular be applied for X and Y being the same space but equipped with different topologies.
- If the rate function I is not good, then upper and lower bound as stated in the LDP for $\{\mu_\varepsilon f^{-1}\}_\varepsilon$ still hold, but I' can fail to be lower semi-continuous and thus fail to be a rate function, as the example $X = Y = \mathbb{R}$, $I(x) \equiv 0$ and $f(x) = \mathrm{e}^x$ shows. (The level sets of I are obviously not compact, and $I'(0) = +\infty \not\leqslant 0 = \liminf_{n\to\infty} I'(1/n)$ shows that I' is not lower semi-continuous.)

The general version of the contraction principle allows to replace the continuous function f by an "almost continuous" function, namely by a function which can be approximated sufficiently good by continuous functions. We state the result without a proof as we are not going to use its general form. Instead we will emulate the proof in the particular situation needed in the subsequent section.

Theorem 2.16 (The contraction principle). *Let X a regular Hausdorff space, $I : X \to [0, \infty]$ a* good *rate function, and $f : X \to Y$ a measurable function from X into a separable metric space (Y, d). Assume that f can be approximated by continuous functions $f_n : X \to Y$, satisfying*

$$\limsup_{n\to\infty} \sup_{\{x\in X \colon I(x)\leqslant\alpha\}} d(f_n(x), f(x)) = 0 \qquad \text{for all } \alpha < \infty. \tag{2.56}$$

Then

$$I'(y) = \inf\{I(x) \colon x \in X \text{ with } y = f(x)\} \tag{2.57}$$

is a good rate function.

If I governs a LDP for $\{\mu_\varepsilon\}_{\varepsilon>0}$ and

$$\lim_{n\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon\{x \in X \colon d(f_n(x), f(x)) \geqslant \delta\} = -\infty \qquad \text{for all } \delta > 0, \tag{2.58}$$

then I' governs a LDP for the induced measures $\{\mu_\varepsilon f^{-1}\}_{\varepsilon>0}$.

2.5 Sample-path large deviations for strong solutions of stochastic differential equations: Wentzell–Freidlin theory

Let us return to our original problem. We are interested in the behaviour of the strong solution x^ε of a SDE

$$\mathrm{d}x_t^\varepsilon = b(x_t^\varepsilon)\,\mathrm{d}t + \sqrt{\varepsilon}\sigma(x_t^\varepsilon)\,\mathrm{d}W_t, \qquad x_0^\varepsilon = x, \tag{2.59}$$

in $\mathbb{R}^d$, where we assume that b and σ satisfy the usual Lipschitz and bounded-growth conditions, so that a unique strong solution exists, cf. Theorem 1.11.

We first consider the special case $\sigma(x) \equiv \mathrm{Id}$, which is much simpler than the case of a general σ as the trivial version of the contraction principle can be used to obtain a LDP for x^ε from Schilder's Theorem as follows.

Define $F : \mathcal{C}_0 \to \mathcal{C}$ by $g \mapsto F(g) = f$, where f is the unique solution in $\mathcal{C}$ of the integral equation

$$f(t) = x + \int_0^t b(f(s))\,\mathrm{d}s + g(t). \tag{2.60}$$

Note that $F(\sqrt{\varepsilon}W) = x^\varepsilon$.

The Lipschitz continuity of b implies via Gronwall's lemma that F is continuous. Indeed, for $g_1, g_2 \in \mathcal{C}_0$,

$$\|F(g_1) - F(g_2)\|_{[0,t]} \leqslant L_b \int_0^t \|F(g_1) - F(g_2)\|_{[0,s]}\,\mathrm{d}s + \|g_1 - g_2\|_{[0,T]} \qquad \text{for all } t \in [0,T], \tag{2.61}$$

which yields by Gronwall's lemma

$$\|F(g_1) - F(g_2)\|_{[0,T]} \leqslant \mathrm{e}^{L_b T}\|g_1 - g_2\|_{[0,T]}. \tag{2.62}$$

Thus the continuity of F is proved and we can apply the contraction principle.

By Schilder's theorem, $\{\sqrt{\varepsilon}W\}_{\varepsilon>0}$ satisfies a LDP with the good rate function

$$I^{\mathrm{BM}}(\varphi) = I^{\mathrm{BM}}_{[0,T],0}(\varphi) = \begin{cases} \frac{1}{2}\|\varphi\|^2_{H_1}, & \text{if } \varphi \in H_1, \\ +\infty, & \text{otherwise,} \end{cases} \tag{2.63}$$

and the contraction principle implies that $x^\varepsilon = F(\sqrt{\varepsilon}W)$ satisfies a LDP with the good rate function

$$I(f) = I_{[0,T],x}(f) = \inf\{I^{\mathrm{BM}}_{[0,T],0}(g) \colon g \in \mathcal{C}_0 \text{ with } f = F(g)\}. \tag{2.64}$$

It remains to identify I. First note that if $g \notin H_1$, then $f = F(g) \notin H_1$. Assume now that there exists at least one $g \in H_1$ such that $f = F(g)$. Then f is almost surely differentiable,

$$\dot{f}_t = b(f_t) + \dot{g}_t, \tag{2.65}$$

and $f \in H_1$ follows. Solving (2.65) for $\dot{g}_t$,

$$I(f) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \|\dot{f}_s - b(f_s)\|^2\,\mathrm{d}s, & \text{if } f \in H_1, \\ +\infty, & \text{otherwise,} \end{cases} \tag{2.66}$$

follows.

Let us now return to the case of a general diffusion coefficient σ. We are going to prove the following result.

Theorem 2.17. *Assume that b and σ are Lipschitz continuous and that there exists a constant $M > 0$ such that $\|b(x)\| \leqslant M$ and $\frac{1}{M}\,\mathrm{Id} \leqslant \sigma(x)\sigma(x)^T \leqslant M\,\mathrm{Id}$ hold for all x, the latter in the sense that all eigenvalues of the symmetric matrices*

$$a(x) \coloneqq \sigma(x)\sigma(x)^T \tag{2.67}$$

lie in the interval $[1/M, M]$.

Then $\{\mathbb{P}(x^\varepsilon)^{-1}\}_{\varepsilon>0}$ satisfies a LDP with the good rate function

$$I(f) = I_{[0,T],x}(f) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \|a(f_s)^{-1/2}[\dot{f}_s - b(f_s)]\|^2\,\mathrm{d}s, & \text{if } f - x \in H_1, \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.68}$$

Remark 2.18.

- While we will only prove the result stated above, the conditions on drift and diffusion coefficient can be relaxed. The boundedness assumptions can be replaced by $\|b(x)\| \leqslant M(1+\|x\|^2)^{1/2}$ and $0 < a(x) \leqslant M(1+\|x\|^2)\mathrm{Id}$ for all x. Then the LDP as stated above still holds, and the rate function I is good.
- Even if we drop the lower bound on the eigenvalues of $a(x)$ and only require $a(x) \leqslant M(1+\|x\|^2)\mathrm{Id}$, the LDP remains valid with a good rate function. But note that in this case the identification of the rate function may fail, and we only know that

$$I(f) = \inf\Big\{I^{\mathrm{BM}}(g) : g \in H_1 \text{ with } f_t = x + \int_0^t b(f_s)\,\mathrm{d}s + \int_0^t a(f_s)^{1/2}\dot{g}_s\,\mathrm{d}s,\ t \in [0,T]\Big\}. \tag{2.69}$$

PROOF OF THEOREM 2.17. As we cannot apply the trivial version of the contraction principle directly, we first approximate x^ε by Euler approximations

$$x_t^{\varepsilon,n} = x + \int_0^t b(x_s^{\varepsilon,n})\,\mathrm{d}s + \sqrt{\varepsilon}\int_0^t \sigma(x_{T_n(s)}^{\varepsilon,n})\,\mathrm{d}W_s, \tag{2.70}$$

where

$$T_n(s) = \frac{[ns]}{n}, \tag{2.71}$$

obtained by "freezing" the diffusion coefficient locally. By doing so, we may proceed as before and define *continuous* functions $F_n : \mathcal{C}_0 \to \mathcal{C}$, by setting $F_n(g) = f$ where f is the unique solution of the integral equation

$$f(t) = x + \int_0^t b(f(s))\,\mathrm{d}s + \sum_{k=1}^{[nt]} \sigma(f_{k/n})[g(k/n \wedge t) - g((k-1)/n)]. \tag{2.72}$$

We haven chosen F_n in such a way that $F_n(\sqrt{\varepsilon}W) = x^{\varepsilon,n}$. By Schilder's Theorem and the contraction principle, we know that $\mathbb{P}\{x^{\varepsilon,n}\}_{\varepsilon>0}$ satisfies a LDP with the good rate function

$$I^n(f) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \|a(f_{T_n(s)})^{-1/2}[\dot{f}_s - b(f_s)]\|^2\,\mathrm{d}s, & \text{if } f - x \in H_1, \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.73}$$

Now we have to show that I^n indeed converges towards I, uniformly in the sense that for any set Γ, $\inf_\Gamma I^n$ converges towards the corresponding infimum for I.

Lemma 2.19.
$$\lim_{n\to\infty} \inf_{\Gamma} I^n = \inf_{\Gamma} I \qquad \forall \Gamma. \tag{2.74}$$

Next, we need to show that $x^{n,\varepsilon}$ is a sufficiently good approximation to x^ε, so that the difference between the two processes becomes negligible on the large-deviations scale as $n \to \infty$.

Lemma 2.20.
$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{\|x^{\varepsilon,n} - x^\varepsilon\|_\infty > \delta\} \to -\infty \qquad \text{as } n \to \infty. \tag{2.75}$$

As soon as these results are established, the upper and the lower bound in the LDP can be shown by easy arguments.

Upper bound:
Let F be an arbitrary closed set in $\mathcal{C}_0$. Then $x^\varepsilon \in F$ implies that either the Euler approximation $x^{\varepsilon,n}$ is at least in the closed δ-neighbourhood of F or that the distance between path and its Euler approximation is larger than δ:

$$\mathbb{P}\{x^\varepsilon \in F\} \leqslant \mathbb{P}\big\{x^{\varepsilon,n} \leqslant \overline{F^{(\delta)}}\big\} + \mathbb{P}\{\|x^\varepsilon - x^{\varepsilon,n}\|_\infty > \delta\} \qquad \forall n \in \mathbb{N}\ \forall \delta > 0. \tag{2.76}$$

By the LDP for the Euler approximations and the preceeding lemmas we find that

$$\begin{aligned}
&\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon \in F\} \\
&\quad\leqslant \lim_{n\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log\Big(2\max\Big\{\mathbb{P}\big\{x^{\varepsilon,n} \leqslant \overline{F^{(\delta)}}\big\}, \mathbb{P}\{\|x^\varepsilon - x^{\varepsilon,n}\|_\infty > \delta\}\Big\}\Big) \\
&\quad\leqslant \max\Big\{-\inf_{\overline{F^{(\delta)}}} I, -\infty\Big\} = -\inf_{\overline{F^{(\delta)}}} I
\end{aligned} \tag{2.77}$$

for all $\delta > 0$. Finally, taking the limit $\delta \searrow 0$ completes the proof.

Lower bound:
To establish the lower bound, we take an arbitrary open subset of $\mathcal{C}_0$ and choose an arbitrary element $x \in G$ as well as a $\delta > 0$ such that the open ball $B(x, 2\delta)$ is contained in G. Then, as before, we split the event we are interested in according to whether the Euler approximation is close to x^ε or not. Thus we find that

$$\mathbb{P}\{x^{\varepsilon,n} \in B(x,\delta)\} \leqslant \mathbb{P}\{x^\varepsilon \in G\} + \mathbb{P}\{\|x^\varepsilon - x^{\varepsilon,n}\|_\infty > \delta\} \qquad \forall n \in \mathbb{N}. \tag{2.78}$$

Again employing the LDP for the Euler approximations and the two preceeding lemmas, we obtain

$$\begin{aligned}
-I(x) &\leqslant -\inf_{B(x,\delta)} I \leqslant \lim_{n\to\infty} \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^{\varepsilon,n} \in B(x,\delta)\} \\
&\leqslant \max\Big\{\liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon \in G\}, \lim_{n\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{\|x^\varepsilon - x^{\varepsilon,n}\|_\infty > \delta\}\Big\} \\
&= \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{x^\varepsilon \in G\}.
\end{aligned} \tag{2.79}$$

As $x \in G$ was chosen arbitrarily, the lower bound follows. □

We omit the proof of Lemma 2.19, as it is not difficult and only sketch the proof of Lemma 2.20.

PROOF OF LEMMA 2.20. For simplicity, we assume that $x = 0$ and $T = 1$. We want to show that the deviation

$$y_t^{\varepsilon,n} := x_t^{\varepsilon,n} - x_t^\varepsilon \tag{2.80}$$

is typically small in the sense that for any $\delta > 0$,

$$\lim_{n\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{\|x^{\varepsilon,n} - x^\varepsilon\|_\infty > \delta\} = -\infty. \tag{2.81}$$

We fix $\delta > 0$, choose an arbitrary $\varrho > 0$ and introduce the stopping times

$$\tau^{\varepsilon,n} := \inf\{t \geqslant 0\colon \|x_t^{\varepsilon,n} - x_{T_n(t)}^{\varepsilon,n}\| \geqslant \varrho\} \wedge 1, \tag{2.82}$$

$$\zeta^{\varepsilon,n} := \inf\{t \geqslant 0\colon \|y_t^{\varepsilon,n}\| > \delta\} \wedge \tau^{\varepsilon,n}. \tag{2.83}$$

We want to show that $\zeta^{\varepsilon,n} = 1$ with high probability, and consider $\tau^{\varepsilon,n}$ as an auxiliary stopping time, in order to deal with the unlikely event that $x^{\varepsilon,n}$ changes by more than ϱ on a short time interval. We estimate the probability we are interested in by splitting the event as follows:

$$\mathbb{P}\{\|y^{\varepsilon,n}\|_\infty > \delta\} \leqslant \mathbb{P}\{\tau^{\varepsilon,n} < 1\} + \mathbb{P}\{\|y^{\varepsilon,n}\|_\infty > \delta, \tau^{\varepsilon,n} \geqslant 1\}. \tag{2.84}$$

The first term on the right-hand side can be estimated using standard estimates. First a crude estimate shows that

$$\begin{aligned}\mathbb{P}\{\tau^{\varepsilon,n} < 1\} &\leqslant \sum_{k=0}^{n-1} \mathbb{P}\Big\{\sup_{k/n\leqslant t<(k+1)/n} \|x_t^{\varepsilon,n} - x_{T_n(t)}^{\varepsilon,n}\| \geqslant \varrho\Big\} \\ &\leqslant n\,\mathbb{P}\Big\{\sup_{0\leqslant t<1/n} \|W_t\| \geqslant (\varrho - M/n)/\sqrt{\varepsilon}M\Big\},\end{aligned} \tag{2.85}$$

where we denote by M the bound on $\|b(x)\|$ and $\|\sigma(x)\|$. Now, a standard estimate shows that

$$\mathbb{P}\{\tau^{\varepsilon,n} < 1\} \leqslant 2nd\,\mathrm{e}^{-n(\varrho-M/n)^2/2d\varepsilon M^2}, \tag{2.86}$$

and

$$\lim_{n\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\{\tau^{\varepsilon,n} < 1\} = -\infty \tag{2.87}$$

follows for any $\varrho > 0$.

The estimation of the second term is more involved and we will only sketch it. First note that

$$\begin{aligned}\mathbb{P}\{\|y^{\varepsilon,n}\|_\infty > \delta, \tau^{\varepsilon,n} \geqslant 1\} &\leqslant \mathbb{E}\bigg\{1_{\{\|y^{\varepsilon,n}\|_\infty>\delta\}}\bigg(\frac{\varrho^2 + \|y_{\zeta^{\varepsilon,n}}^{\varepsilon,n}\|^2}{\varrho^2+\delta^2}\bigg)^{1/\varepsilon}\bigg\} \\ &\leqslant \Big(\frac{1}{\varrho^2+\delta^2}\Big)^{1/\varepsilon}\mathbb{E}\Big\{\big(\varrho^2 + \|y_{\zeta^{\varepsilon,n}}^{\varepsilon,n}\|^2\big)^{1/\varepsilon}\Big\}.\end{aligned} \tag{2.88}$$

Defining $f(y) := (\varrho^2 + \|y\|^2)^{1/\varepsilon}$ and $u(t) := \mathbb{E}\{f(y_{t\wedge\zeta^{\varepsilon,n}}^{\varepsilon,n})\}$, we need to estimate $u(1)$. $u(t)$ can be estimated using Itô's formula for $f(y_t^{\varepsilon,n})$ and then taking expectations. Finally applying Gronwall's lemma,

$$u(t) \leqslant \varrho^{2/\varepsilon}\,\mathrm{e}^{const\,(d)t/\varepsilon} \tag{2.89}$$

follows for all $t \leqslant 1$. Thus,

$$\mathbb{P}\{\|y^{\varepsilon,n}\|_\infty > \delta, \tau^{\varepsilon,n} \geqslant 1\} \leqslant \Big(\frac{1}{\varrho^2+\delta^2}\Big)^{1/\varepsilon} u(1) \leqslant \Big(\frac{1}{\varrho^2+\delta^2}\Big)^{1/\varepsilon} \varrho^{2/\varepsilon}\,\mathrm{e}^{const(d)/\varepsilon}, \tag{2.90}$$

yielding

$$\limsup_{\varrho\to 0}\sup_{n\geqslant 1}\limsup_{\varepsilon\to 0}\varepsilon\log\mathbb{P}\{\|y^{\varepsilon,n}\|_\infty > \delta, \tau^{\varepsilon,n} \geqslant 1\} = -\infty, \tag{2.91}$$

which completes the proof. □

3 Diffusion exit from a domain

3.1 Introduction

In this section we want to study the noise-induced exit from a neighbourhood of an equilibrium point of the corresponding deterministic system. We study mean exit times, the asymptotic behaviour of exit times as well as exit locations. Our presentation follows [12].

We continue to assume the usual conditions which assure the existence and uniqueness of a strong solution of the SDE

$$\mathrm{d}x_t^\varepsilon = b(x_t^\varepsilon)\,\mathrm{d}t + \sqrt{\varepsilon}\sigma(x_t^\varepsilon)\,\mathrm{d}W_t, \qquad x_0^\varepsilon = x, \tag{3.1}$$

in $\mathbb{R}^d$ on the time interval $[0,T]$, i.e., we assume that b and σ are Lipschitz continuous and grow at most linearly for large $\|x\|$. We denote by $\mathbb{P}_x(x^\varepsilon)^{-1}$ the distribution of the sample path x^ε, solving (3.1) with initial condition x.

Throughout the whole section, we assume that D is a bounded (open) domain.

Let us recall that the information on first-exit times and exit locations can in principle be obtained exactly in form of solutions to partial differential equations (PDEs) involving the generator of the diffusion process x^ε:

Theorem 3.1. *Assume that ∂D is smooth and that $a(x) = \sigma(x)\sigma(x)^T \geqslant (1/M)\,\mathrm{Id}$ in the sense that the smallest eigenvalue of $a(x)$ is uniformly bounded away from zero as x varies. Denote by $\mathcal{L}^\varepsilon$,*

$$\mathcal{L}^\varepsilon v(t,x) = \frac{\varepsilon}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}v(t,x) + \langle b(x), \nabla v(t,x)\rangle, \tag{3.2}$$

the infinitesimal generator of the diffusion x^ε. Finally, fix an initial condition $x \in D$ and let

$$\tau^\varepsilon := \inf\{t \geqslant 0 \colon x_t^\varepsilon \notin D\} \tag{3.3}$$

denote the first-exit time from D. Then the following assertions hold.

- $\mathbb{P}_x\{\tau^\varepsilon \leqslant t\}$ *is the unique solution of the PDE*

$$\begin{aligned} \frac{\partial v(t,x)}{\partial t} &= \mathcal{L}^\varepsilon v && \text{for } x \in D,\ t > 0, \\ v(0,x) &= 0, && \text{for } x \in D, \\ v(t,x) &= 1 && \text{for } x \in \partial D,\ t > 0. \end{aligned} \tag{3.4}$$

- $\mathbb{E}_x\{\tau^\varepsilon\}$ *is the unique solution of the PDE*

$$\begin{aligned} \mathcal{L}^\varepsilon u &= -1 && \text{in } D, \\ u &= 0 && \text{on } \partial D. \end{aligned} \tag{3.5}$$

- *For a continuous function f on the boundary of D, $\mathbb{E}_x\{f(x_{\tau^\varepsilon}^\varepsilon\}$ is the unique solution of the PDE*

$$\begin{aligned} \mathcal{L}^\varepsilon w &= 0 && \text{in } D, \\ w &= f && \text{on } \partial D. \end{aligned} \tag{3.6}$$

Remark 3.2.

- Note that the boundary conditions can easily be understood. For instance, $u = 0$ on ∂D just tells us that the mean exit time from D, when starting on the boundary, is zero, i. e., the process exits immediately.
- The assumption of ∂D being smooth can be relaxed. Some regularity of ∂D is needed, such as the "exterior-ball condition", which states that for any point $y \in \partial D$ there exists a ball $B \subset D^c$ such that the closure of D and the closure of the ball intersect and only intersect in y. This excludes that y is such that the diffusion starting in y on the boundary of D does not exit from $\overline{D}$ immediately.
- The assumption on $a(x)$ guarantees that the infinitesimal generator $\mathcal{L}^\varepsilon$ of the diffusion process is uniformly elliptic.

Let us consider a simple one-dimensional example.

Example 3.3. Let $d = 1$ and assume $b(0) = 0$, $b(x) < 0$ for $x > 0$, $b(x) > 0$ for $x < 0$ and $\sigma(x) \equiv 1$. Thus we are considering the overdamped motion of a Brownian particle in the one-dimensional potential $U(x) := -\int_0^x b(y)\,\mathrm{d}y$, which has exactly one potential well, containing 0, and no saddles. The drift will always push the particle in the direction of the bottom of the well. Let $D = (\alpha_1, \alpha_2) \ni 0$. We want to calculate the probability that x^ε leaves D, say at α_1. Thus we solve the (one-dimensional) Dirichlet problem

$$\begin{aligned} \mathcal{L}^\varepsilon w &= 0 && \text{in } D, \\ w &= f && \text{on } \partial D, \end{aligned} \tag{3.7}$$

with

$$f(x) = \begin{cases} 1, & \text{for } x = \alpha_1, \\ 0, & \text{for } x = \alpha_2. \end{cases} \tag{3.8}$$

The solution is given by

$$w(x) = \mathbb{P}_x\{x^\varepsilon_{\tau^\varepsilon} = \alpha_1\} = \mathbb{E}_x f(x^\varepsilon_{\tau^\varepsilon}) = \frac{\int_x^{\alpha_2} \mathrm{e}^{2U(y)/\varepsilon}\,\mathrm{d}y}{\int_{\alpha_1}^{\alpha_2} \mathrm{e}^{2U(y)/\varepsilon}\,\mathrm{d}y}. \tag{3.9}$$

Thus, for $x \in D$,

$$\begin{aligned} \lim_{\varepsilon\to 0} \mathbb{P}_x\{x^\varepsilon_{\tau^\varepsilon} = \alpha_1\} &= 1, && \text{if } U(\alpha_1) < U(\alpha_2), \\ \lim_{\varepsilon\to 0} \mathbb{P}_x\{x^\varepsilon_{\tau^\varepsilon} = \alpha_1\} &= 0, && \text{if } U(\alpha_2) < U(\alpha_1), \end{aligned} \tag{3.10}$$

and an application of Laplace's method also shows

$$\lim_{\varepsilon\to 0} \mathbb{P}_x\{x^\varepsilon_{\tau^\varepsilon} = \alpha_1\} = \frac{\frac{1}{|U'(\alpha_1)|}}{\frac{1}{|U'(\alpha_1)|} + \frac{1}{|U'(\alpha_2)|}}, \qquad \text{if } U(\alpha_1) = U(\alpha_2). \tag{3.11}$$

Thus, in principle, the above equations provide all the information. In practise, in the multidimensional case, the corresponding PDEs cannot always be solved and even a computational approach is not necessarily feasible. Here the LDP from the previous section comes into play as it will provide us at least with the asymptotic behaviour of first-exit times and exit locations, although the results will be less precise than those in the preceding example.

3.2 Quasipotentials

Recall that the family $\{\mathbb{P}_x(x^\varepsilon)^{-1}\}_{\varepsilon>0}$ satisfies a LDP, the rate function of which we denote by $I = I_{[0,T],x}$. Recall from the previous section that if $a(x) = \sigma(x)\sigma(x)^T$ is positive definite, then

$$I(f) = I_{[0,T],x}(f) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \|a(f_s)^{-1/2}[\dot{f}_s - b(f_s)]\|^2 \,\mathrm{d}s, & \text{if } f - x \in H_1, \\ +\infty, & \text{otherwise.} \end{cases} \tag{3.12}$$

As a corollary to the LDP for $\{\mathbb{P}_x(x^\varepsilon)^{-1}\}_{\varepsilon>0}$, we obtain a first result on the diffusion exit from a domain.

Corollary 3.4. *Assume that the smallest eigenvalue of $a(x)$ is uniformly bounded away from zero. Choose an initial condition $x \in D$. Then the first-exit time*

$$\tau^\varepsilon := \inf\{t \geqslant 0 \colon x_t^\varepsilon \notin D\} \tag{3.13}$$

of x^ε from D satisfies

$$\lim_{\varepsilon\to 0} \varepsilon \log \mathbb{P}_x\{\tau^\varepsilon \leqslant t\} = -\inf\{V(x,y;s) \colon s \in [0,t], y \notin D\}, \tag{3.14}$$

where

$$V(x,y;s) := \inf\{I_{[0,s],x}(\varphi) \colon \varphi \in \mathcal{C}([0,s],\mathbb{R}^d), \varphi_0 = x, \varphi_s = y\} \tag{3.15}$$

denotes the cost of forcing a path to connect x and y in time s.

Remark 3.5.

- The statement of the corollary includes in particular the existence of the limit $\varepsilon \to 0$. We skip its proof.
- Observe that the calculation of the asymptotical behaviour thus reduces to a variational problem. $V(x,y;s)$ can be obtained as the solution of a Hamilton–Jacobi equation and the extremals from an Euler equation.

Already the preceding corollary shows the importance of

$$\begin{aligned} V(x,y;t) &= \inf\{I_{[0,t],x}(\varphi) \colon \varphi \in \mathcal{C}([0,t],\mathbb{R}^d), \varphi_0 = x, \varphi_t = y\} \\ &= \inf\Big\{\frac{1}{2}\int_0^t \|u_s\|^2 \,\mathrm{d}s : u \in \mathcal{L}_2([0,t],\mathbb{R}^d) \text{ such that} \\ &\qquad\qquad \varphi_s = x + \int_0^s b(\varphi_s)\,\mathrm{d}s + \int_0^s \sigma(\varphi_s)u_s\,\mathrm{d}s,\ s \in [0,t], \text{ and } \varphi_t = y\Big\}. \end{aligned} \tag{3.16}$$

We define

$$V(x,y) := \inf_{t>0} V(x,y;t), \tag{3.17}$$

denoting the cost of forcing a path starting in x to reach y eventually. In the case of the deterministic dynamics possessing a unique stable equilibrium point at the origin, $y \mapsto V(0,y)$ is called *quasipotential*. If D contains this equilibrium point, the initial condition becomes increasingly unimportant as $\varepsilon \to 0$, as sample paths will generally visit a neighbourhood of the origin before attempting to exit from D. Thus the quasipotential which measures asymptotically the difficulty or cost of leaving D does not depend on the initial condition.

The following lemma relates the quasipotential to the potential in the case the drift coefficient actually derives from a potential.

Lemma 3.6. *Assume $\sigma = \mathrm{Id}$ and that there exist a continuously differentiable potential U on $\overline{D}$ and a transversal term l satisfying the following conditions*

$$\begin{aligned} &U(0) = 0, \\ &U(x) > 0 \qquad \text{for all } x \neq 0, \\ &\nabla U(x) \neq 0 \qquad \text{for all } x \neq 0, \\ &\langle l(x), \nabla U(x) \rangle = 0, \\ &b(x) = -\nabla U(x) + l(x). \end{aligned} \tag{3.18}$$

Then the quasipotential $V(0, y)$ satisfies $V(0, y) = 2U(y)$ for all $y \in \overline{D}$ such that $U(y) \leqslant U_0 := \min_{z \in \partial D} U(z)$.

If U is twice continuously differentiable and l is continuously differentiable, then the rate function I has a unique extremal φ on the set

$$\{\varphi \in \mathcal{C}((-\infty, T], \mathbb{R}^d) \colon \lim_{s \to -\infty} \varphi_s = 0, \varphi_T = x\}; \tag{3.19}$$

this extremal is the solution of the ODE

$$\begin{cases} \dot{\varphi}_s = +\nabla U(\varphi_s) + l(\varphi_s) & s \in (-\infty, T], \\ \varphi_T = x. \end{cases} \tag{3.20}$$

Remark 3.7.

- Note that the relation "the quasipotential is twice the potential" holds only for y such that $U(y) \leqslant U_0$, where U_0 is the lowest value U takes on along the boundary of D. The reason becomes clear if we again think of the quasipotential as representing the "cost of forcing the process" to reach a certain point on the boundary. As the process will leave D where it is most easy, the process will not feel regions where $U(y)$ is larger than U_0 but simply exit near a point $z \in \partial D$ with $U(z) = U_0$. Thus the quasipotential does not know about the potential U in regions where $U(y) > U_0$.

- If $l(y) \equiv 0$, (3.20) reflects that "the best way" for the process to reach some point y is to "climb directly towards the exit point" in the sense that it is cheapest in terms of the rate function to go from 0 to y along the path the deterministic motion would take from y to 0. If the transversal term $l(y) \not\equiv 0$, then the direction of the resulting cycling is retained.

PROOF OF LEMMA 3.6. For the sake of simplicity in the presentation, let us assume that $l(y) \equiv 0$. (The necessary changes in order to incorporate general l are obvious, we only need to take into account that l is vertical to ∇U.)

Fix an arbitrary $x \in \overline{D}$ such that $U(x) \leqslant U_0$. If a path φ_s does not leave $\overline{D}$ during a time interval $[T_1, T_2]$, then the mean-value theorem shows

$$U(\varphi_{T_2}) - U(\varphi_{T_1}) = \int_{T_1}^{T_2} \langle \nabla U(\varphi_s), \dot{\varphi}_s \rangle \, \mathrm{d}s. \tag{3.21}$$

Thus

$$\begin{aligned} I_{[T_1,T_2],0}(\varphi) &= \frac{1}{2}\int_{T_1}^{T_2} \|\dot{\varphi}_s + \nabla U(\varphi_s)\|^2 \,\mathrm{d}s \\ &= \frac{1}{2}\int_{T_1}^{T_2} \|\dot{\varphi}_s - \nabla U(\varphi_s)\|^2 \,\mathrm{d}s + 2\int_{T_1}^{T_2} \langle \nabla U(\varphi_s), \dot{\varphi}_s \rangle \,\mathrm{d}s \\ &\geqslant [U(\varphi_{T_2}) - U(\varphi_{T_1})]. \end{aligned} \tag{3.22}$$

Choosing a φ such that φ remains in $\overline{D}$ during the time interval $[T_1, T_2]$ and satisfies $\varphi_{T_1} = 0$ as well as $\varphi_{T_2} = x$, shows that

$$I_{[T_1,T_2],0}(\varphi) \geqslant 2U(x). \tag{3.23}$$

Now, if φ does leave $\overline{D}$ during $[T_1, T_2]$, then there exists a $\widetilde{T} \in [T_1, T_2]$ such that $U(\varphi_{\widetilde{T}}) = U_0$ and

$$I_{[T_1,T_2],0}(\varphi) \geqslant I_{[T_1,\widetilde{T}],0}(\varphi) \geqslant 2U(\varphi_{\widetilde{T}}) = 2U_0 \geqslant 2U(x) \tag{3.24}$$

establishes the lower bound for general φ.

In order for equality to hold in (3.22), we need φ to satisfy $\|\dot{\varphi}_s - \nabla U(\varphi_s)\| = 0$ for all s. Note that $\nabla U(0) = 0$ excludes the possibility of a solution to (3.20) going from 0 to x in finite time, but all solutions to (3.20) satisfy $\lim_{s\to-\infty} \varphi_s = 0$. Therefore, for such a φ, we obtain

$$I_{(-\infty,T],0}(\varphi) = 2\int_{-\infty}^{T} \langle \nabla U(\varphi_s), \dot{\varphi}_s \rangle \,\mathrm{d}s = 2[U(x) - U(0)] = 2U(x). \tag{3.25}$$

Here we used that we have already seen that an optimal path cannot leave $\overline{D}$ before reaching x. By the definition of the quasipotential, $V(0, x) = U(x)$ follows.

For twice continuously differentiable U, the solution to (3.20) is unique and thus the uniqueness of the extremal follows. □

3.3 Diffusion exit: Classical results

Let us know turn to the main result of this section. We start be stating our assumptions in addition to those guaranteeing existence and uniqueness of a strong solution to the SDE (3.1). Our presentation follows [12].

Let D be a bounded domain with a smooth boundary ∂D, and assume that the following assumptions hold.

Assumption 3.8. The deterministic dynamical system $\dot{x}_t = b(x_t)$ has a unique stable equilibrium point $x^* = 0$ in D, and if $x_0 \in D$, then $x_t \in D$ for all $t > 0$ and $\lim_{t\to\infty} x_t = x^* = 0$.

Assumption 3.9. If $x_0 \in \partial D$, then $\lim_{t\to\infty} x_t = x^* = 0$.

Assumption 3.10.

$$\overline{V} := \inf_{z\in\partial D} V(0,z) < \infty. \tag{3.26}$$

Assumption 3.11. There exist a constant $K > 0$ and a maximal radius $\varrho_0 > 0$ such that for all $\varrho \leqslant \varrho_0$ and all x_0, y satisfying $\|x_0 - z\| + \|z - y\| \leqslant \varrho$ for some $z \in \partial D \cup \{x^*\}$, there exist a "control" $u \in \mathcal{L}_2$, $\|u\|_\infty < K$, and a time $T(\varrho)$ such that the path ϕ_t, defined by

$$\phi_t = x_0 + \int_0^t b(\phi_s)\,\mathrm{d}s + \int_0^t \sigma(\phi_s)u_s\,\mathrm{d}s, \tag{3.27}$$

satisfies $\phi_{T(\varrho)} = y$, where $T(\varrho) \to 0$ as $\varrho \to 0$. (Note that x_0 and y are not assumed to lie in D.)

Let us comment on these assumptions.

Remark 3.12.

- Assumptions 3.8 and 3.9 are assumptions on the deterministic dynamics only and state that $\overline{D}$ is positively invariant and that x^* is asymptotically stable: When the deterministic process starts in D or on its boundary, then the process cannot leave $\overline{D}$ and will approach the stable equilibrium point x^*. Note that we are only interested in the situation described in Assumption 3.8, because if the deterministic dynamics leaves D in finite time, then we already know that the random process, which stays close to its deterministic counterpart on time scales of order 1, will with overwhelming probability also leave D within a time of order 1. Thus we study the case of the deterministic dynamics not leaving D. Naturally noise will nevertheless cause the random process to leave D occasionally. Let us remark that exit time and location will not be determined by the deterministic dynamics.

- Assumption 3.9 excludes a characteristic boundary[5], i. e., the boundary between different domains of attraction for the deterministic dynamics is excluded from D. This means, for example, if the drift coefficient derives from a potential, then D cannot contain a saddle between potential wells, so that the results we are going to prove below can a priori not be used to study transitions between wells. Fortunately, this assumption can be relaxed, at the price of much more involved proofs.

[5] ∂D is called characteristic, if $\langle b(z), n(z)\rangle = 0$ for all $z \in \partial D$, where n denotes the outer normal vector to ∂D.

- Assumption 3.10 assures that the random process has a chance to reach the boundary on the scale we are considering. Were $\overline{V} = \infty$, all possible exit points would be equally unlikely.

- Finally, Assumption 3.11 is a controllability condition and states that there exists a *bounded* control such that the controlled process connects the initial condition x_0 and y within time $T(\varrho)$. In particular, we require that the closer x_0 and y are either to the boundary of D or to the equilibrium point x^* and to each other, the less time the controlled process needs to connect the two points.

  While Assumption 3.10 assures that the boundary of D is accessible on the large-deviation scale, Assumption 3.11 guarantees that moving away from x^* or crossing the boundary it is not "too expensive" in terms of the quasipotential. This as expressed in Lemma 3.13 below.

  It is not difficult to show that Assumption 3.11 is trivially satisfied if $a(x) = \sigma(x)\sigma(x)^T$ is positive definite for $x = x^* = 0$ and uniformly positive definite for $x \in \partial D$.

We will use the following consequence of Assumption 3.11, which can be considered as a continuity property for the quasipotential near $x^* = 0$ and near the boundary of D.

Lemma 3.13. *For all $\delta > 0$ there exists a $\varrho > 0$ such that*

$$\sup_{x,y \in B(0,\varrho)} \inf_{t \in [0,1]} V(x, y; t) < \delta \tag{3.28}$$

and

$$\sup_{x,y : \inf_{z \in \partial D}(\|x-z\|+\|z-y\|) \leqslant \varrho} \inf_{t \in [0,1]} V(x, y; t) < \delta. \tag{3.29}$$

The main result of this section is the following.

Theorem 3.14. *Suppose that Assumptions 3.8 to 3.11 are satisfied. Then the following assertions hold true.*
First-exit time: *For all initial conditions $x \in D$ and all $\delta > 0$,*

$$\lim_{\varepsilon \to 0} \mathbb{P}_x\big\{ \mathrm{e}^{(\overline{V}-\delta)/\varepsilon} < \tau^\varepsilon < \mathrm{e}^{(\overline{V}+\delta)/\varepsilon} \big\} = 1 \tag{3.30}$$

and

$$\lim_{\varepsilon \to 0} \varepsilon \log \mathbb{E}_x\{\tau^\varepsilon\} = \overline{V}. \tag{3.31}$$

First-exit location: *For any closed subset $N \subset \partial D$ satisfying $\inf_{z \in N} V(0,z) > \overline{V}$, and all initial conditions $x \in D$,*

$$\lim_{\varepsilon \to 0} \mathbb{P}_x\big\{ x^\varepsilon_{\tau^\varepsilon} \in N \big\} = 0. \tag{3.32}$$

If $V(0,\cdot)$ has a unique minimum z^* on ∂D, then for all initial conditions $x \in D$ and all $\delta > 0$,

$$\lim_{\varepsilon \to 0} \mathbb{P}_x\big\{ |x^\varepsilon_{\tau^\varepsilon} - z^*| < \delta \big\} = 1. \tag{3.33}$$

Remark 3.15.

- The theorem shows that the asymptotic behaviour of the mean exit time is the one predicted by physicists in the case of a drift coefficient deriving from a potential U. Namely, Arrhenius' law [1] states that the logarithm of the mean exit time behaves like twice the hight to be surmounted divided by the noise intensity, that is $2U_0/\varepsilon$, where $U_0 = \min_{z\in\partial D} U(z)$. In the light of Lemma 3.6, this is precisely (3.31). In addition, we find that not only the mean of the first-exit time asymptotically equals $e^{\overline{V}/\varepsilon}$, but that the first-exit time is actually concentrated around its mean as made precise by (3.30).

- The results on the exit location show what we expected. The diffusion process favours an exit near boundary points where $V(0,\cdot)$ is minimal. More precisely, only neighbourhoods of points where $V(0,\cdot)$ attains its minimum play a role. In the limit $\varepsilon \to 0$ of vanishing noise intensity, these neighbourhoods can be chosen arbitrarily small. If $V(0,\cdot)$ has multiple minima on ∂D as in Example 3.3, the weights in (3.33) corresponding to the different minima cannot be obtained by large-deviation techniques.

- As already mentioned, Assumption 3.9 can be relaxed and ∂D being a characteristic boundary can be allowed for, see [11].

The proof of the theorem relies on a series of lemmas which we are going to state now. Let us remark once and for all that we will always assume that the radii of balls are chosen small enough for the balls to be contained in D.

Lemma 3.16 (Lower bound on the probability of an exit when starting near x^*).

$$\forall\eta > 0\ \exists\varrho_0\ \forall\varrho\in(0,\varrho_0)\ \exists T_0 \qquad \liminf_{\varepsilon\to 0} \varepsilon\log \inf_{x\in B(0,\varrho)} \mathbb{P}_x\{\tau^\varepsilon \leqslant T_0\} > -(\overline{V}+\eta). \tag{3.34}$$

We introduce another stopping time

$$\sigma_\varrho := \inf\{t \geqslant 0 \colon x_t^\varepsilon \in B(0,\varrho)\cup\partial D\}, \tag{3.35}$$

which describes the first hitting of either the boundary of D or of a small neighbourhood of x^*.

Lemma 3.17 (x^ε cannot remain in D arbitrarily long without approaching x^*).

$$\forall\varrho > 0 \qquad \lim_{t\to\infty} \limsup_{\varepsilon\to 0} \varepsilon\log\sup_{x\in D} \mathbb{P}_x\{\sigma_\varrho > t\} = -\infty. \tag{3.36}$$

Lemma 3.18 (Probability of leaving D before hitting a neighbourhood of x^*). *For any closed $N\subset\partial D$,*

$$\lim_{\varrho\to 0} \limsup_{\varepsilon\to 0} \varepsilon\log \sup_{y\in\mathcal{S}(0,2\varrho)} \mathbb{P}_y\{x^\varepsilon_{\sigma_\varrho}\in N\} \leqslant -\inf_{z\in N} V(0,z), \tag{3.37}$$

where $\mathcal{S}(0,2\varrho)$ denotes the sphere of radius 2ϱ, centred in $x^* = 0$.

Lemma 3.19 (Probability of returning to a neighbourhood of x^* before leaving D).

$$\forall \varrho > 0 \; \forall x \in D \qquad \lim_{\varepsilon \to 0} \mathbb{P}_x\{x^\varepsilon_{\sigma_\varrho} \in B(0,\varrho)\} = 1. \tag{3.38}$$

Lemma 3.20 (Bound on the distance covered by x^ε during short time).

$$\forall \varrho > 0 \; \forall c > 0 \; \exists T = T(c,\varrho) < \infty \qquad \limsup_{\varepsilon \to 0} \varepsilon \log \sup_{x \in D} \mathbb{P}_x\Big\{ \sup_{t \in [0,T]} \|x^\varepsilon_t - x\| \geqslant \varrho \Big\} \leqslant -c. \tag{3.39}$$

The basic idea of the proof of Theorem 3.14 is the following: If our initial condition is close to x^*, Lemma 3.16 provides a *lower* bound on the probability that x^ε leaves D within finite time, and we use Lemma 3.17 to estimate the time needed for excursions away from x^*. We will iterate this argument below, thereby obtaining a lower bound on the first-exit time. The proof of the corresponding upper bound is slightly more involved, as we will need to keep track of return times to the sphere $\mathcal{S}(0,2\varrho)$ and to the ball $B(0,\varrho)$. The results on the exit location can be obtained by similar arguments.

PROOF OF THEOREM 3.14. We will only prove the results on the first-exit time τ^ε. Fix an initial condition $x \in D$ and an arbitrary $\delta > 0$. Without loss of generality, we may assume that $\overline{V} > 0$.

Lower bound:
Let $\eta = \delta/8$ and choose an $\varrho > 0$. If $x \in B(0,\varrho)$, Lemma 3.16 guarantees the existence of a time $T_0 = T_0(\eta,\varrho)$ and an $\varepsilon_0 > 0$ such that

$$\mathbb{P}_x\{\tau^\varepsilon \leqslant T_0\} > e^{-(\overline{V}+2\eta)/\varepsilon} \tag{3.40}$$

for all $\varepsilon \leqslant \varepsilon_0$, uniformly in x. By Lemma 3.17, we find a $T_1 = T_1(\eta,\varrho)$ satisfying

$$\mathbb{P}_x\{\sigma_\varrho > T_1\} < e^{-\eta/\varepsilon}, \tag{3.41}$$

for all $\varepsilon \leqslant \varepsilon_0$, and holding uniformly in $x \in D$.

We set $T := T_0 + T_1$. Then, for $\varepsilon \leqslant \varepsilon_0$, x^ε has probability

$$q := \inf_{x \in D} \mathbb{P}_x\{\tau^\varepsilon \leqslant T\} \geqslant \inf_{x \in D} \mathbb{P}_x\{\sigma_\varrho \leqslant T_1\} \inf_{x \in B(0,\varrho)} \mathbb{P}_x\{\tau^\varepsilon \leqslant T_0\} \geqslant e^{-(\overline{V}+4\eta)/\varepsilon} \tag{3.42}$$

of having left D before time T.

Iterating with the help of the strong Markov property shows

$$\sup_{x \in D} \mathbb{P}_x\{\tau^\varepsilon > kT\} \leqslant (1-q)^k, \tag{3.43}$$

which in turn yields

$$\sup_{x \in D} \mathbb{E}_x\{\tau^\varepsilon\} \leqslant \sup_{x \in D} T \sum_{k=0}^{\infty} \mathbb{P}_x\{\tau^\varepsilon > kT\} \leqslant \frac{T}{q} \leqslant T\, \mathrm{e}^{(\overline{V}+\delta/2)/\varepsilon} \tag{3.44}$$

by (3.42). This shows the upper bound on the mean first-exit time. By Tchebychev's inequality, we obtain in addition that

$$\sup_{x \in D} \mathbb{P}_x\big\{\tau^\varepsilon \geqslant \mathrm{e}^{(\overline{V}+\delta)/\varepsilon}\big\} \leqslant \mathrm{e}^{-(\overline{V}+\delta)/\varepsilon} \sup_{x \in D} \mathbb{E}_x\{\tau^\varepsilon\} = T\, \mathrm{e}^{-\delta/2\varepsilon}, \tag{3.45}$$

which proves the first part of (3.30).

Upper bound:
As already mentioned, the proof of the upper bound requires a little bit more care. We start be introducing a sequence of alternating stopping times.

$$\begin{aligned} \theta_0 &:= 0, \\ \tau_m &:= \inf\{t \geqslant \theta_m \colon x_t^\varepsilon \in B(0,\varrho) \cup \partial D\}, \\ \theta_{m+1} &:= \inf\{t \geqslant \tau_m \colon x_t^\varepsilon \in \mathcal{S}(0,2\varrho)\}, \end{aligned} \tag{3.46}$$

for all $m \geqslant 0$, where we set $\theta_m = \infty$ if $\tau_m = \tau^\varepsilon$, i.e., if x^ε has left D. These stopping times keep track of the time when x^ε first reaches either the neighbourhood $B(0,\varrho)$ of x^* or the boundary of D, and as long as x^ε does not leave D, alternating visits to the sphere $\mathcal{S}(0,2\varrho)$ and to $B(0,\varrho)$ are recorded, so that we can use these stopping times to measure the excursions away from x^*.

In order to obtain an upper bound for $\mathbb{P}_x\{\tau^\varepsilon \leqslant kT_2\}$, we first choose a suitable T_2. By Lemma 3.20, we can choose T_2 so that the probability that time T_2 is sufficient to go from the sphere of radius ϱ to the sphere of radius 2ϱ is small. Namely, we bound the probability of going from the smaller to the larger sphere by the probability of covering distance ϱ anywhere, and require T_2 to be chosen such that

$$\sup_{x\in D} \mathbb{P}_x\{\theta_m - \tau_{m-1} \leqslant T_2\} \leqslant \sup_{x\in D} \mathbb{P}_x\Bigl\{\sup_{t\in[0,T_2]} \|x^\varepsilon - x\|_\infty > \varrho\Bigr\} \leqslant \mathrm{e}^{-(\overline{V}-\delta/2)/\varepsilon} \tag{3.47}$$

is satisfied for all sufficiently small ε.

Assume $\tau^\varepsilon \leqslant kT_2$. Then, if x^ε leaves D, there must be an index m such that $\tau^\varepsilon = \tau_m$. We want an estimate for τ_m. The value of τ_m will naturally depend on the length of the previous excursions. If all excursions between subsequent visits to $B(0,\varrho)$ have been at least of length T_2, then there cannot have been many, so that $m \leqslant k$ must hold. Thus we find that

$$\mathbb{P}_x\{\tau^\varepsilon \leqslant kT_2\} \leqslant \mathbb{P}_x\{\exists m \leqslant k \colon \tau^\varepsilon = \tau_m\} + \mathbb{P}_x\{\exists m \leqslant k \colon \theta_m - \tau_{m-1} \leqslant T_2\} \tag{3.48}$$

By the definition of T_2, the second term satisfies

$$\mathbb{P}_x\{\exists m \leqslant k \colon \theta_m - \tau_{m-1} \leqslant T_2\} \leqslant k\,\mathrm{e}^{-(\overline{V}-\delta/2)/\varepsilon}, \tag{3.49}$$

so that we are left with bounding the first one. For $m \geqslant 1$, we can restart upon hitting the sphere of radius 2ϱ and check whether we reach ∂D before $B(0,\varrho)$, while the case $m=0$ can be dealt with by Lemma 3.19:

$$\mathbb{P}_x\{\exists m \leqslant k \colon \tau^\varepsilon = \tau_m\} \leqslant \mathbb{P}_x\{\tau^\varepsilon = \tau_0\} + \sum_{m=1}^{k} \sup_{y\in\mathcal{S}(0,2\varrho)} \mathbb{P}_y\{x_{\sigma_\varrho}^\varepsilon \in \partial D\}. \tag{3.50}$$

By Lemma 3.18, applied for $N = \partial D$, fixing a small enough ϱ and choosing ε also small enough, allows to estimate

$$\sup_{y\in\mathcal{S}(0,2\varrho)} \mathbb{P}_y\{x_{\sigma_\varrho}^\varepsilon \in \partial D\} \leqslant \mathrm{e}^{-(\overline{V}-\delta/2)/\varepsilon}. \tag{3.51}$$

Finally, Lemma 3.19 shows that for small enough ε,

$$\mathbb{P}_x\{\tau^\varepsilon = \tau_0\} = \mathbb{P}_x\{x_{\sigma_\varrho}^\varepsilon \notin B(0,\varrho)\} \to 0 \qquad \text{as } \varepsilon \to 0. \tag{3.52}$$

Choosing $k = [e^{(\overline{V}-\delta)/\varepsilon}/T_2] + 1$ shows

$$\lim_{\varepsilon\to 0} \mathbb{P}_x\{\tau^\varepsilon \leqslant e^{(\overline{V}-\delta)/\varepsilon}\} = 0, \tag{3.53}$$

cf. (3.30). It would remain to establish also the other bound on the mean first-exit time, via Tchebychev's inequality, but in order to do so, we would first need to obtain the rate of convergence in (3.52), and we will skip that part of the proof here as we will skip the proof of the results on the exit location. The latter are obtained by splitting the event of interest according to $\tau^\varepsilon \geqslant \tau_m$, $m \in \mathbb{N}$. □

We will not give the complete proofs of all five lemmas. After proving Lemma 3.16 we will only comment on the proofs of the other four lemmas.

PROOF OF LEMMA 3.16. Fix $\eta, \varrho > 0$ and recall that we want to find a T_0 such that we have at least a certain probability that x^ε leaves D up to time T_0. We want this estimate to be uniform in x from a neighbourhood of x^*:

$$\liminf_{\varepsilon\to 0} \varepsilon \log \inf_{x\in B(0,\varrho)} \mathbb{P}_x\{\tau^\varepsilon \leqslant T_0\} > -(\overline{V}+\eta). \tag{3.54}$$

The basic idea is to construct a deterministic exiting path ϕ^x with $I(\phi^x) \leqslant (2/3)\eta$, leading from x first to x^*, then further on to the boundary of D, and finally we extend this path to a prescribed length.

- First note that by Assumption 3.11 (Lemma 3.13), for small enough ϱ we can find a continuous path ψ^x, leading from x to $x^* = 0$ in time $t_x \leqslant 1$, satisfying

$$I_{[0,t_x],x}(\psi^x) \leqslant \eta/3. \tag{3.55}$$

(Note that for $\sigma = \mathrm{Id}$, we can simply choose $\psi^x(s) = x - sx/\|x\|$ and $t_x = \|x\|$.)

- Next, by the definition of $\overline{V}$ and the general assumption of $\overline{V}$ being finite, we know that there exists a continuous path ϕ, leading from 0 to some $z \in \partial D$ (of course independent of x), with

$$I_{[t_x,t_x+t_0],0}(\psi) \leqslant \overline{V} + \eta/3. \tag{3.56}$$

- Again using Assumption 3.11, we can can choose a point $y \notin \overline{D}$ at a distance, say ϱ, from $z \in \partial D$ and find a corresponding continuous path ϕ^z, leading from z to y in time $t_z \leqslant 1$, with

$$I_{[t_x+t_0,t_x+t_0+t_z],z}(\psi^z) \leqslant \eta/3. \tag{3.57}$$

- Finally, we denote by ϕ^y the path starting in y and following the deterministic dynamics $\dot{\phi}^y_s = b(\phi^y_s)$ during the time interval $[t_x + t_0 + t_z, t_0 + 2]$. Note that

$$I_{[t_x+t_0+t_z,t_0+2],y}(\psi^y) = 0 \tag{3.58}$$

and that this path may return to D (and will do so, unless the time interval during which it is defined is very short).

- Now we concatenate these paths, obtaining a continuous path $\phi^x : [0, t_0 + 2] \to \mathbb{R}^d$, leading from x to some unspecified point, but leaving D and satisfying

$$I_{[0,t_0+2],x}(\phi^x) \leqslant \overline{V} + \eta. \tag{3.59}$$

It remains to compare x^ε to the so constructed exit path. To do so, we define a set of functions

$$\Psi := \bigcup_{x\in B(0,\varrho)} \{\psi \in \mathcal{C} \colon \|\psi - \phi^x\|_\infty < \varrho/2\}, \tag{3.60}$$

which consists of a $\varrho/2$-neighbourhood of the union of all exit paths. Note that these exit paths all agree once they have reached $x^* = 0$. Any path in Ψ must leave D, as all exit paths reach y which has distance ϱ from D, so that any paths in Ψ has at least distance $\varrho/2$ from D at some point. We use this observation to estimate the probability that x^ε leaves D before time $T_0 := t_0 + 2$: If $x^\varepsilon \in \Psi$, then we also know that $\tau^\varepsilon \leqslant T_0$. Thus we can obtain the desired lower bound from the LDP[6] for x^ε:

$$\begin{aligned}
\liminf_{\varepsilon\to 0} \varepsilon \log \inf_{x\in B(0,\varrho)} \mathbb{P}_x\{\tau^\varepsilon \leqslant T_0\} &\geqslant \liminf_{\varepsilon\to 0} \varepsilon \log \inf_{x\in B(0,\varrho)} \mathbb{P}_x\{x^\varepsilon \in \Psi\} \\
&\geqslant - \sup_{x\in B(0,\varrho)} \inf_{\psi\in\Psi} I_{[0,T_0],x}(\psi) \\
&\geqslant - \sup_{x\in B(0,\varrho)} I_{[0,T_0],x}(\psi^x) \geqslant -(\overline{V} + \eta).
\end{aligned} \tag{3.61}$$

Thus the proof of the lemma is complete. □

Let us briefly mention the key ingredients to the proofs of the other lemmas. The proof of Lemmas 3.17 and 3.18 also rely on the LDP for x^ε, while Lemma 3.19 is based on comparison with a deterministic path and Gronwall's lemma. Finally, Lemma 3.20 can be shown with the help of standard martingale arguments.

[6]Note that we use uniformity of the LDP in the initial condition—which we have not proved in the previous section. We refer to the literature.

3.4 Refined results

The classical results presented in Theorem 3.14 provide a rigorous version of Arrhenius' law in physics, extended to drift coefficients not necessarily deriving from a potential and allowing for a rather general class of diffusion coefficients. Since the work of Eyring and Kramers [14, 20], also the prefactor in the asymptotic behaviour of the mean exit time is known in the case of the drift coefficient deriving from a potential. A rigorous proof in the multi-dimensional case was apparently not known until the recent work by Bovier, Eckhoff, Gayrard and Klein [9, 10]. Their proof relies heavily on potential theory and the notion of capacities. Here, we will only summarize the main results.

We consider a (sufficiently smooth) potential $U : \mathbb{R}^d \to \mathbb{R}$ and study the overdamped motion of a Brownian particle in this potential, i. e.,

$$\mathrm{d}x_t^\varepsilon = -\nabla U(x_t^\varepsilon)\,\mathrm{d}t + \sqrt{\varepsilon}\,\mathrm{d}W_t. \tag{3.62}$$

For a given set D, we denote by

$$\tau_D = \inf\{t \geqslant 0 \colon x_t^\varepsilon \in D\} \tag{3.63}$$

the *first-hitting time*, i. e., the time when the diffusion x^ε first hits D.

We are interested in the mean passage time from a minimum x of U to a neighbourhood of those minima of U which lie lower. Let us first assume that U has precisely two minima x and y and that $U(x) > U(y)$ and that there is a unique lowest saddle point between the corresponding wells. As we cannot expect x^ε actually to hit the point y, we consider a (sufficiently small) neighbourhood of y. (D should neither contain x nor the saddle between x and y.) The classical Eyring–Kramers formula states that

$$\mathbb{E}_x \tau_D \sim \frac{2\pi}{|\lambda_1^*(z^*)|}\sqrt{\frac{|\det \nabla^2 U(z^*)|}{\det \nabla^2 U(x)}}\, \mathrm{e}^{2[U(z^*)-U(x)]/\varepsilon}, \tag{3.64}$$

where $\nabla^2 U(z^*)$ denotes the Hessian of U at the saddle z^*, $\nabla^2 U(x)$ denotes the Hessian of U at the minimum x in which we start, and, finally, $\lambda_1^*(z^*)$ denotes the unique negative eigenvalue of the Hessian at the saddle. Note that $2[U(z^*) - U(x)]$ is precisely twice the height to be surmounted in order to get from the minimum x to the saddle z^* and that this equals the value of the quasipotential $V(0, z^*) = \overline{V}$ as introduced above. We remark that the curvature at x and at the saddle play a role, but not the curvature at the other minimum y as it is sufficient to pass over the saddle. Once the saddle is surmounted, dropping into the well is essentially "for free".

From now on, we will use make the following assumptions:

Assumption 3.21.

- $U : \mathbb{R}^d \to \mathbb{R}$ is three times continuously differentiable and $-\nabla U$ satisfies the usual bounded-growth condition for the SDE (3.62).
- The potential U has a finite set of minima $\mathcal{M} = \{x_1, \dots, x_n\}$.

Remark 3.22. These assumptions can be relaxed, for instance U may depend on ε in a controlled way, and we may as well study the SDE (3.62) in a regular domain.

To formulate the generalization of the Eyring–Kramers formula (3.64), to the case of more than two minima, we need some notations which we are going to introduce now.

Notations 3.23.

- Fix a minimum $x_i \in \mathcal{M}$. Then we denote by $\mathcal{M}_i = \{y_1, \dots, y_k\} \subset \mathcal{M} \setminus \{x_i\}$ all minima of U which "lie lower than x_i", i.e., satisfy $U(y_j) \leqslant U(x_i)$ for $j = 1, \dots, k$.

- A point $z^* = z^*(A, B)$ is called a *saddle* between the sets A and B if

$$U(z^*) = U(z^*(A, B)) = \inf_{\varphi \in \mathcal{C}: \varphi_0 \in A, \varphi_1 \in B} \sup_{t \in [0,1]} U(\varphi_t). \tag{3.65}$$

  Thus a saddle is the highest point one has necessarily to cross in order to go from set A to set B. We consider all possible path. Each path has at least one "highest point", and among all paths we choose one such that its highest point is as low as possible. Of course, the path is not uniquely determined as we may always linger around somewhere for a while. In addition, in general the saddle itself is also not unique. There might be several paths leading over different highest points which happen to be of the same height. And a single path may lead over more than one saddle of the same height, looking like a camel's back. We denote by $\mathcal{S}(A, B)$ the set of all saddles between A and B and remark in passing that U is constant on $\mathcal{S}(A, B)$, so that we may write $U(z^*)$ without specifying $z^* \in \mathcal{S}(A, B)$.

We will make the following assumption.

Assumption 3.24 (Non-degeneracy assumption (for given sets A and B)).

- The Hessian of U is non-degenerate at all saddle points between A and B.

- Along any optimal path φ contributing to the definition of a saddle between A and B, $t \mapsto U(\varphi_t)$ has a unique maximum.

The assumption of a unique maximum of the map $t \mapsto U(\varphi_t)$ means that an optimal path has only one "highest point", thus the camel's-back picture is excluded. In principle such a situation could be dealt with, but the presentation would become more complicated.

Let us now state the main results from [9, 10]. The theorem below gives a general form of the Eyring–Kramers formula with an *multiplicative* error term.

Theorem 3.25. *Fix a minimum $x_i \in \mathcal{M}$, and assume that the (closed) set D has a smooth boundary and satisfies the following conditions:*

- *D contains neighbourhoods of all "lower lying minima":*

$$\bigcup_{j=1}^{k} B(y_j, \varepsilon) \subset D. \tag{3.66}$$

- *D does not contain neighbourhoods of the other minima, nor does D contain a neighbourhood of x_i:*

$$\bigcup_{j: x_j \in \mathcal{M} \setminus \mathcal{M}_j} B(x_j, \varepsilon) \cap D = \emptyset. \tag{3.67}$$

- *All saddles between x_i and the set of lower-lying minima $\mathcal{M}_i$ are bounded away from D: There exists a $\delta > 0$, independent of ε, such that*

$$\operatorname{dist}(\mathcal{S}(\{x_i\}, \mathcal{M}_i), D) > \delta. \tag{3.68}$$

Suppose the non-degeneracy assumption holds for all saddles $\mathcal{S}(\{x_i\}, D) = \{z_1^*, \dots, z_k^*\}$ between x_i and D, and in addition, the Hessian of U at the minimum x_i is also non-degenerate. Then the mean first-hitting time τ_D satisfies

$$\mathbb{E}_{x_i}\tau_D = \frac{2\pi\, \mathrm{e}^{2[U(z^*)-U(x_i)]/\varepsilon}\big(1+\mathcal{O}(\sqrt{\varepsilon}|\log\varepsilon|)\big)}{\sqrt{\det\nabla^2 U(x_i)}\ \sum_{j=1}^k \frac{|\lambda_1^*(z_j^*)|}{\sqrt{|\det\nabla^2 U(z_j^*)|}}}. \tag{3.69}$$

Remark 3.26. The assumption of ∂D being smooth can be replaced a regularity assumption.

This result is augmented by [10, Theorem 1.1] which relates the exponentially small eigenvalues of the generator of the diffusion x^ε to the hitting times of nested neighbourhoods describing how the neighbourhoods of deeper and deeper minima of U are reached. Without giving the precise assumptions, we can roughly summarize this by saying that the kth of the n exponentially small eigenvalues of the generator satisfies

$$\lambda_k = \frac{1}{\mathbb{E}_{x_k}\tau_{\bigcup_{i=1}^{k-1} B(x_i,\varepsilon)}}\big(1+\mathcal{O}(\mathrm{e}^{-const/\varepsilon})\big), \tag{3.70}$$

where we assume that the minima x_i are ordered in such a way that the barrier height to be surmounted in order to go from x_k to a neighbourhood of any of the x_j, $j < k$, is lower than the barrier heights involved in passing from any x_j with $j < k$ to any other well corresponding to x_l for some $l \leqslant k$.

Finally, [10, Theorem 1.3] shows that the first hitting times are asymptotically exponentially distributed. Under suitable assumptions,

$$\begin{aligned}\mathbb{P}_{x_i}\{\tau_D > t\,\mathbb{E}_{x_i}\tau_D\} = {}& \mathrm{e}^{-t(1+\mathcal{O}(\mathrm{e}^{-const/\varepsilon}))}\big(1+\mathcal{O}(\mathrm{e}^{-const/\varepsilon})\big)\\ &+\sum_{j>i}\mathcal{O}(\mathrm{e}^{-const/\varepsilon})\,\mathrm{e}^{-t\lambda_j\mathbb{E}_{x_i}\tau_D} + \mathcal{O}(1)\,\mathrm{e}^{-t\mathcal{O}(\varepsilon^{d-1})\mathbb{E}_{x_i}\tau_D}\,.\end{aligned} \tag{3.71}$$

4 Singularly perturbed random dynamical systems

4.1 Introduction

In the last section we finally want to discuss what happens when the potential is not static or, more generally, when the drift and diffusion coefficients are time-dependent. The method presented below can be applied to multidimensional slow–fast systems (see [8]) and we do not need to assume that the drift coefficient derives from a potential, but in order to ease the presentation, we will restrict ourselves to one-dimensional systems, where we always may assume that the drift coefficient derives from a potential, and to constant diffusion coefficients.

We study SDEs of the form

$$\mathrm{d}x_s = f(x_s, \varepsilon s)\,\mathrm{d}s + \sigma\,\mathrm{d}W_s, \qquad f(x, \varepsilon s) = -\frac{\partial}{\partial x} V(x, \varepsilon s), \tag{4.1}$$

in $\mathbb{R}$, where $x \mapsto V(x, \varepsilon s)$ is the potential at time s, and $\varepsilon \ll 1$ is the speed at which the potential varies. We think of a potential $\widetilde{V}(x, \lambda)$, depending on a parameter λ which changes slowly as time evolves, i. e., $\lambda = \lambda(\varepsilon s)$, and $V(x, \varepsilon s) = \widetilde{V}(x, \lambda(\varepsilon s))$.

The SDE describes the overdamped motion of a Brownian particle in the potential V, the potential changing its shape slowly in time. For the potential to change noticeably, we'll need to wait a time of order $1/\varepsilon$. We think of the speed ε being "moderate". By that we mean that the time $T_{\text{forcing}} = 1/\varepsilon$ at which the change in the potential becomes noticeable, is larger than the "minimal relaxation time" $T_{\text{relax}} = 1/2a$, where a denotes the maximal curvature of the potential during the time interval under consideration. When we study transitions from one potential well to another, ε will typically be smaller than the "maximal Kramers time" $T_{\text{Kramers}} = \mathrm{e}^{2H/\sigma^2}$, where H denotes the maximal value of the barrier height during the time interval under consideration. So we are mainly interested in the case

$$T_{\text{relax}} \ll T_{\text{forcing}} \ll T_{\text{Kramers}}, \tag{4.2}$$

i. e.,

$$\mathrm{e}^{-2H/\sigma^2} \ll \varepsilon \ll 2a, \tag{4.3}$$

where a and H depend on the concrete problem we are interested in.

It will prove useful to make a time change in Equation (4.1) by introducing the so-called *slow time* $t = \varepsilon s$. By doing so, we obtain the equivalent SDE

$$\mathrm{d}x_t = \frac{1}{\varepsilon} f(x_t, t)\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \tag{4.4}$$

which we will study below. Note that the scaling factor $1/\sqrt{\varepsilon}$ in the noise term is a direct consequence of the Brownian scaling $W_{ct} = \sqrt{c}W_t$ (in distribution), for all $c > 0$.

4.2 The deterministic case

As a first step, we need to understand the behaviour of the solutions of the deterministic equation corresponding to (4.4), which is a special case of a singularly perturbed ODE:

$$\varepsilon \dot{x}_t^{\det} = f(x_t^{\det}, t), \qquad x_0^{\det} = x_0. \tag{4.5}$$

We want to study the case where the motion starts near the bottom of one of the wells of the potential V, and for the moment, we will assume that this well remains well-separated from any other well V might have, while V changes slowly in time. Thus, no (almost-)bifurcations may occur. Let as formulate our assumptions on f which guarantee that we are in such a situation.

Assumption 4.1. There exist an interval $I = [0, T]$ or $I = [0, \infty)$ and a constant $d > 0$ such that the following properties hold:

- there exists a function $x^* : I \to \mathbb{R}$, called *equilibrium branch*, such that

$$f(x^*(t), t) = 0 \qquad \forall t \in I; \tag{4.6}$$

- f is twice continuously differentiable with respect to x and t for $|x - x^*(t)| \leqslant d$ and $t \in I$, with uniformly bounded derivatives. In particular, there exists a constant $M > 0$ such that $|\partial_{xx} f(x, t)| \leqslant 2M$ in that domain;
- the linearization of f at $x^*(t)$, defined as

$$a^*(t) = \partial_x f(x^*(t), t), \tag{4.7}$$

  is negative and bounded away from zero, that is, there exists a constant $a_0 > 0$ such that

$$a^*(t) \leqslant -a_0 \quad \forall t \in I. \tag{4.8}$$

The assumption that $a^*(t)$ is negative assures that the equilibrium branch is *stable*.

Note that the equilibrium branch $x^*(t)$ consists of equilibrium points $x^*(t_0)$ of the frozen deterministic system

$$\varepsilon \dot{x}_t^{\mathrm{fr}} = f(x_t^{\mathrm{fr}}, t_0), \qquad t \geqslant 0. \tag{4.9}$$

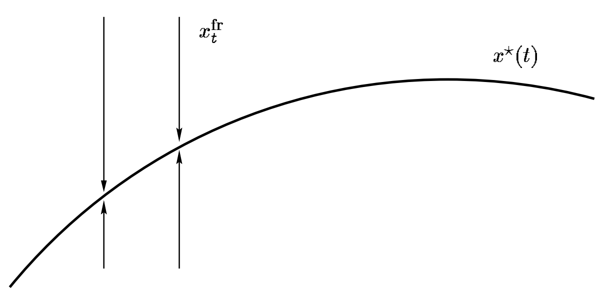

FIGURE 1. Solutions of the frozen deterministic system (4.9) for different values of t_0, presented in the (t_0, x)-plane. Solutions approach their equilibrium point exponentially fast. An equilibrium branch $x^*(t)$ of (4.13) consist of the equilibrium points $x^*(t_0)$ of (4.9).

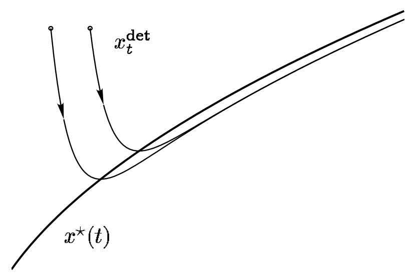

FIGURE 2. Solutions of the slowly time-dependent equation (4.13) represented in the (t, x)-plane. A stable equilibrium branch $x^*(t)$ attracts nearby solutions x_t^{det}. Two solutions with different initial conditions are shown. They converge exponentially fast to each other, as well as to a neighbourhood of order ε of $x^*(t)$.

$x^*(t_0)$ corresponds to the bottom of a well, and $a^*(t_0)$ is the curvature at the bottom. Assuming that $a^*(t_0) \neq 0$ amounts to excluding a flat bottom. When starting the frozen system (4.9) in a neighbourhood of $x^*(t_0)$, it approaches $x^*(t_0)$ exponentially fast (in t/ε), see Figure 1.

In fact, for $x_0 > x^*(t_0)$,

$$\begin{aligned} \varepsilon\frac{\mathrm{d}}{\mathrm{d}t}(x_t^{\text{fr}} - x^*(t_0)) = \varepsilon\dot{x}_t^{\text{fr}} &= a^*(t_0)(x_t^{\text{fr}} - x^*(t_0)) + \mathcal{O}\big((x_t^{\text{fr}} - x^*(t_0))^2\big) \\ &\leqslant \frac{1}{2}a^*(t_0)(x_t^{\text{fr}} - x^*(t_0)), \end{aligned} \tag{4.10}$$

provided $x_t^{\text{fr}} - x^*(t_0) \leqslant const|a^*(t_0)|$. Therefore,

$$x_t^{\text{fr}} - x^*(t_0) \leqslant \mathrm{e}^{-|a^*(t_0)|t/2\varepsilon} \tag{4.11}$$

follows, provided the initial condition x_0 satisfies $x_0 - x^*(t_0) \leqslant const|a^*(t_0)|$.

If $x_0 < x^*(t_0)$, a similar estimate can be obtained, yielding

$$|x_t^{\text{fr}} - x^*(t_0)| \leqslant \mathrm{e}^{-|a^*(t_0)|t/2\varepsilon}, \tag{4.12}$$

whenever $|x_0 - x^*(t_0)| \leqslant const|a^*(t_0)|$.

Let us now discuss the non-frozen, deterministic system

$$\varepsilon\dot{x}_t^{\text{det}} = f(x_t^{\text{det}}, t). \tag{4.13}$$

Condition (4.8) guarantees that the curvature at the bottom of the well is uniformly bounded away from zero at all times, so that there is always a unique "deepest point" in the well. Figure 2 shows solutions of the non-autonomous equation (4.13) for different initial values. The following theorem which is due to Tihonov [27] and Gradšteĭn [17], describes the dynamics of (4.13).

Theorem 4.2. *There are constants $\varepsilon_0, c_0, c_1 > 0$, depending only on f, such that for $0 < \varepsilon \leqslant \varepsilon_0$,*

- (4.13) *admits a particular solution $\widehat{x}_t^{\text{det}}$ such that*

$$|\widehat{x}_t^{\text{det}} - x^*(t)| \leqslant c_1\varepsilon \quad \forall t \in I; \tag{4.14}$$

- *if* $|x_0 - x^*(0)| \leqslant c_0$, *then the solution* $x_t^{\det}$ *of* (4.13) *with initial condition* $x_0^{\det} = x_0$ *satisfies*

$$|x_t^{\det} - \widehat{x}_t^{\det}| \leqslant |x_0 - \widehat{x}_0^{\det}|\,\mathrm{e}^{-a_0 t/2\varepsilon} \qquad \forall t \in I. \tag{4.15}$$

Remark 4.3. The particular solution $\widehat{x}_t^{\det}$ is often called *slow solution* or *adiabatic solution* of Equation (4.13). It is not unique in general, as suggested by (4.15).

PROOF OF THEOREM 4.2. Let $x_t^{\det}$ be any solution of (4.13), and consider the deviation $y_t = x_t^{\det} - x^*(t)$ of $x_t^{\det}$ from $x^*(t)$. Using the first-order Taylor expansion

$$f(x_t^{\det}, t) = f(x^*(t), t) + \partial_x f(x^*(t), t) y_t + b(y_t, t) = 0 + a^*(t) y_t + b(y_t, t), \tag{4.16}$$

where $|b(y,t)| \leqslant M y^2$ for all $t \in I$, $|y| \leqslant d$, we find that

$$\varepsilon \dot{y}_t = f(x_t^{\det}, t) - \varepsilon \dot{x}^*(t) = a^*(t) y_t + b(y_t, t) - \varepsilon \dot{x}^*(t). \tag{4.17}$$

Next we need to show that $|\dot{x}^*(t)|$ is bounded. In fact,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} f(x^*(t), t) = a^*(t) \dot{x}^*(t) + \partial_t f(x^*(t), t) \tag{4.18}$$

yields

$$\dot{x}^*(t) = -\frac{\partial_t f(x^*(t), t)}{a^*(t)}. \tag{4.19}$$

Since we assumed that all derivatives of f are bounded and $a^*(t) \leqslant -a_0$, this shows the existence of a constant $W > 0$ such that $|\dot{x}^*(t)| \leqslant W$.

Therefore, we can estimate

$$\varepsilon \dot{y}_t \leqslant -a_0 y_t + M y_t^2 + \varepsilon W \qquad \text{for } y_t \geqslant 0, \tag{4.20}$$

$$\varepsilon \dot{y}_t \geqslant -a_0 y_t - M y_t^2 - \varepsilon W \qquad \text{for } y_t \leqslant 0. \tag{4.21}$$

Let us consider the case $y_t \geqslant 0$, the other case being similar. We define v_t by

$$\varepsilon \dot{v}_t = -a_0 v_t + M v_t^2 + \varepsilon W =: g(v_t). \tag{4.22}$$

Now, $g(v) = 0$ if and only if

$$v = v_\pm^* := \frac{a_0}{2M} \pm \sqrt{\frac{a_0^2}{4M^2} - \varepsilon \frac{W}{M}}. \tag{4.23}$$

Thus, for small enough ε, $v_\pm^*$ are particular solutions of the equation $\varepsilon \dot{v}_t = g(v_t)$. By the definition of v_t, we know that v_t dominates y_t, provided $v_0 \geqslant y_0$. So we have shown that

$$y_t \leqslant v_-^* = \frac{a_0}{2M} - \frac{a_0}{2M}\sqrt{1 - 4\varepsilon \frac{WM}{a_0^2}} = \varepsilon \frac{W}{a_0} + \mathcal{O}(\varepsilon^2), \tag{4.24}$$

provided $0 \leqslant y_0 \leqslant v_-^*$. Together with the analogous argument for the lower bound we obtain

$$|y_t| \leqslant \varepsilon \frac{W}{a_0} + \mathcal{O}(\varepsilon^2). \tag{4.25}$$

Therefore, choosing any initial condition y_0 satisfying $|y_0| \leqslant \varepsilon W / a_0$ yields a solution y_t satisfying $|y_t| \leqslant c_1 \varepsilon$ for all $t \in I$. This proves (4.14).

Let us now prove that all solutions $x_t^{\det}$, starting in a neighbourhood of $x^*(0)$ approach each other exponentially fast. It suffices to prove (4.15). Let $\widehat{x}_t^{\det}$ be a particular solution satisfying (4.14), and denote by $z_t = x_t^{\det} - \widehat{x}_t^{\det}$ the deviation of $x_t^{\det}$ from that adiabatic solution. Again using a Taylor expansion for f, we find that

$$\varepsilon \dot{z}_t = \hat{a}(t) z_t + \hat{b}(z_t, t), \tag{4.26}$$

where

$$\hat{a}(t) = \partial_x f(\widehat{x}_t^{\det}, t) \leqslant a^*(t) + M(\widehat{x}_t^{\det} - x^*(t)) = a^*(t) + \mathcal{O}(\varepsilon) \leqslant \frac{3}{4} a^*(t) \tag{4.27}$$

for sufficiently small ε. Here we used that $\widehat{x}_t^{\det}$ is a particular solution satisfying (4.14). In addition, we know that $\hat{b}(z,t) \leqslant Mz^2$. Thus, for $0 \leqslant z_t \leqslant a_0/4M$,

$$\varepsilon \dot{z}_t \leqslant -\frac{3}{4} a_0 z_t + M z_t^2 \leqslant -\frac{1}{2} a_0 z_t \tag{4.28}$$

follows. Now, $z_t \leqslant z_0 \, \mathrm{e}^{-a_0 t/2\varepsilon}$ is immediate, provided $0 \leqslant z_0 \leqslant a_0/4M$. Otherwise, a similar argument can be applied, showing that actually $|z_t| \leqslant |z_0| \, \mathrm{e}^{-a_0 t/2\varepsilon}$, whenever $|z_0| \leqslant a_0/4M$. This concludes the proof of Theorem 4.2. □

4.3 Near stable equilibrium branches

Let us return to the random case (4.4), where we shall assume that Assumption 4.1 is satisfied. We remark in passing that we may assume that the usual existence and uniqueness conditions are satisfied, because we are only interested in the local behaviour in a neighbourhood of an equilibrium branch where f is sufficiently smooth by this assumption.

We want to investigate the fluctuations

$$y_t = x_t - x_t^{\det}, \qquad y_0 = 0, \tag{4.29}$$

of the solution x_t of the SDE (4.4) around the corresponding deterministic solution $x_t^{\det}$. We assume that x_t starts at time 0 in some deterministic initial condition x_0, satisfying $|x_0 - x^*(0)| \leqslant c_0$, compare Theorem 4.2. The stochastic process $\{y_t\}_{t\geqslant 0}$ satisfies the SDE

$$\begin{aligned} \mathrm{d}y_t &= \frac{1}{\varepsilon}\big[f(x_t^{\det}+y_t,t) - f(x_t^{\det},t)\big]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t \\ &= \frac{1}{\varepsilon}\big[a(t)y_t + \tilde{b}(y_t,t)\big]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \end{aligned} \tag{4.30}$$

where

$$\begin{aligned} a(t) = \partial_x f(x_t^{\det},t) &= a^*(t) + \mathcal{O}(|x_t^{\det} - x^*(t)|) \\ &= a^*(t) + \mathcal{O}(\varepsilon) + \mathcal{O}\big(|x_0 - x^*(0)|\,\mathrm{e}^{-a_0 t/2\varepsilon}\big) \end{aligned} \tag{4.31}$$

by Theorem 4.2, and $|\tilde{b}(y,t)| \leqslant My^2$ in a neighbourhood of the origin. Note that there are constants $0 < a_- < a_+ < \infty$ such that $-a_+ < a(t) < -a_-$ for all $t \in I$, provided ε and c_0 are small enough.

We will prove that y_t remains in a neighbourhood of the origin with high probability. It is instructive to consider first the linearization of (4.30) around $y = 0$, which has the form

$$\mathrm{d}y_t^0 = \frac{1}{\varepsilon}a(t)y_t^0\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \qquad y_0^0 = 0. \tag{4.32}$$

As we shall see below, y_t^0 gives the main contribution to y_t, so that the Gaussian process obtained by this linearization is indeed a good approximation to y_t. The solution of (4.32) with an *arbitrary* initial condition y_0^0 (independent of to the Brownian motion) is given by

$$y_t^0 = y_0^0\,\mathrm{e}^{\alpha(t)/\varepsilon} + \frac{\sigma}{\sqrt{\varepsilon}}\int_0^t \mathrm{e}^{\alpha(t,s)/\varepsilon}\,\mathrm{d}W_s, \tag{4.33}$$

where we use the notations

$$\alpha(t,s) = \int_s^t a(u)\,\mathrm{d}u \qquad \text{and} \qquad \alpha(t) = \alpha(t,0). \tag{4.34}$$

Note that $\alpha(t,s) \leqslant -a_-(t-s)$ whenever $t \geqslant s$. If y_0^0 has variance $v_0 \geqslant 0$, then y_t^0 has variance

$$v(t) = v_0\,\mathrm{e}^{2\alpha(t)/\varepsilon} + \frac{\sigma^2}{\varepsilon}\int_0^t \mathrm{e}^{2\alpha(t,s)/\varepsilon}\,\mathrm{d}s. \tag{4.35}$$

Since the first term decreases exponentially fast, the initial variance v_0 is "forgotten" as soon as $\mathrm{e}^{2\alpha(t)/\varepsilon}$ is small enough, which happens already for $t \gg \varepsilon$. For $y_0^0 = 0$, (4.33) and the usual tail estimate for Gaussian random variables imply that for any $\delta > 0$,

$$\mathbb{P}_{0,0}\big\{|y_t^0| \geqslant \delta\big\} \leqslant \mathrm{e}^{-\delta^2/2v(t)}, \tag{4.36}$$

and thus the probability of finding y_t^0, at any given time $t \in I$, outside an interval of width much larger than $\sqrt{2v(t)}$ is very small. We summarize this by saying that the *typical spreading* of y_t^0 at time t is $\sqrt{2v(t)}$.

Our aim is now to extend the concentration result (4.36) to the *whole sample path* of the original *nonlinear* process $\{y_t\}_{t\geqslant 0}$. We want to define a "strip", i. e., a space–time set, scaling with $\sqrt{v(t)}$, such that y_t is likely to remain in that strip. We start by investigating the variance $v(t)$. Integration by parts shows that

$$\begin{aligned} v(t) &= v_0 \, \mathrm{e}^{2\alpha(t)/\varepsilon} + \frac{\sigma^2}{\varepsilon} \int_0^t \mathrm{e}^{2\alpha(t,s)/\varepsilon} \, \mathrm{d}s \\ &= v_0 \, \mathrm{e}^{2\alpha(t)/\varepsilon} + \frac{\sigma^2}{\varepsilon} \biggl[\frac{\varepsilon}{2|a(t)|} - \frac{\varepsilon}{2|a(0)|} \, \mathrm{e}^{2\alpha(t)/\varepsilon} + \varepsilon \int_0^t \frac{-a'(s)}{2a(s)^2} \, \mathrm{e}^{2\alpha(t,s)/\varepsilon} \, \mathrm{d}s \biggr]. \end{aligned} \tag{4.37}$$

It remains to bound the integral. First note that

$$\begin{aligned} |a'(s)| &= \frac{\mathrm{d}}{\mathrm{d}s} \partial_x f(x_s^{\det}, s) \\ &= \partial_{sx} f(x_s^{\det}, s) + \partial_{xx} f(x_s^{\det}, s) \frac{\mathrm{d}}{\mathrm{d}s} x_s^{\det} \leqslant const \biggl[1 + \frac{1}{\varepsilon} f(x_s^{\det}, s) \biggr]. \end{aligned} \tag{4.38}$$

To show that $f(x_s^{\det}, s)$ is not too large, we again compare f at $(x_s^{\det}, s)$ with f at the point $(x^\star(s), s)$ on the equilibrium branch. Using a first-order Taylor expansion and Theorem 4.2, we find that

$$\begin{aligned} |f(x_s^{\det}, s)| &\leqslant |f(x^\star(s), s)| + \mathcal{O}(|x_s^{\det} - x^\star(s)|) = \mathcal{O}(|x_s^{\det} - x^\star(s)|) \\ &\leqslant const \, \varepsilon \biggl[1 + \frac{|x_0 - x^\star(0)|}{\varepsilon} \, \mathrm{e}^{-a_0 s/2\varepsilon} \biggr]. \end{aligned} \tag{4.39}$$

Plugging this estimate into (4.37), we find that

$$v(t) = \sigma^2 \biggl[\frac{1}{2|a(t)|} + \biggl(\frac{v_0}{\sigma^2} - \frac{1}{2|a(0)|} \biggr) \mathrm{e}^{2\alpha(t)/\varepsilon} + \mathcal{O}(\varepsilon) \biggr]. \tag{4.40}$$

When starting in $y_0^0 = 0$ at time 0, the process $\{y_t^0\}_{t\geqslant 0}$ has initially variance 0 and it takes a time $\gg \varepsilon$ for the process to relax to metastable equilibrium. When defining the strip, we will pretend that y_t^0 has already relaxed to metastable equilibrium. This amounts to choosing $v_0 = \sigma^2/2|a(0)|$ in (4.35) or (4.40), respectively. Thus we set

$$\zeta(t) = \frac{1}{2|a(0)|} \, \mathrm{e}^{2\alpha(t)/\varepsilon} + \frac{1}{\varepsilon} \int_0^t \mathrm{e}^{2\alpha(t,s)/\varepsilon} \, \mathrm{d}s, \tag{4.41}$$

and define the corresponding strip by

$$\mathcal{B}_{\mathrm{s}}(h) = \bigl\{ (x,t) \in \mathbb{R} \times I \colon |x - x_t^{\det}| < h\sqrt{\zeta(t)} \bigr\}. \tag{4.42}$$

Note that in the case of a static potential, i. e., $a(t) \equiv -a_0$, we find $\zeta(t) \equiv 1/2a_0$, and the strip $\mathcal{B}_{\mathrm{s}}(h)$ is of constant width, namely a $h/\sqrt{2a_0}$-neighbourhood of $x_t^{\det}$.

Let

$$\tau_{\mathcal{B}_{\mathrm{s}}(h)} = \inf\bigl\{ t \geqslant 0 \colon (x_t, t) \notin \mathcal{B}_{\mathrm{s}}(h) \bigr\}, \tag{4.43}$$

denote the first-exit time of $\{x_t\}_{t\geqslant 0}$ from $\mathcal{B}_{\mathrm{s}}(h)$.[7] Before estimating $\tau_{\mathcal{B}_{\mathrm{s}}(h)}$, we shall summarize some properties of the width $\sqrt{\zeta(t)}$ of the strip.

[7] More precisely, $\tau_{\mathcal{B}_{\mathrm{s}}(h)}$ is the first-exit time of $\{(x_t, t)\}_{t\geqslant 0}$ from the space–time set $\mathcal{B}_{\mathrm{s}}(h)$.

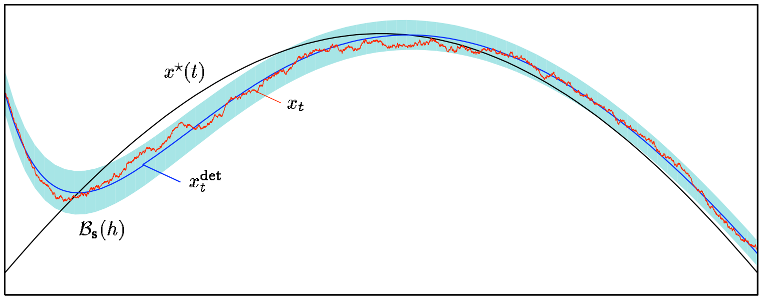

Figure 3. A sample path of the stochastic differential equation (4.4). It is likely to remain in the set $\mathcal{B}_{\mathrm{s}}(h)$ which is centred at x_t^{det}, up to exponentially long times t.

Lemma 4.4. *The function $\zeta(t)$ satisfies the following relations for all $t \in I$.*

$$\zeta(t) = \frac{1}{2|a^\star(t)|} + \mathcal{O}(\varepsilon) + \mathcal{O}\big(|x_0 - x^\star(0)|\,\mathrm{e}^{-a_0 t/2\varepsilon}\big), \tag{4.44}$$

$$\frac{1}{2a_+} \leqslant \zeta(t) \leqslant \frac{1}{2a_-}, \tag{4.45}$$

$$\zeta'(t) \leqslant \frac{1}{\varepsilon}. \tag{4.46}$$

Proof: The proof of (4.44) is a direct consequence of (4.40) and Theorem 4.2, and it remains only to show (4.45) and (4.46). Observe that $\zeta(t)$ is a solution of the linear ordinary differential equation (ODE)

$$\frac{\mathrm{d}\zeta}{\mathrm{d}t} = \frac{1}{\varepsilon}\big(2a(t)\zeta + 1\big), \qquad \zeta(0) = \frac{1}{2|a(0)|}. \tag{4.47}$$

Since $\zeta(t) > 0$ and $a(t) < 0$, we have $\zeta'(t) \leqslant 1/\varepsilon$. We also see that $\zeta'(t) \geqslant 0$ whenever $\zeta(t) \leqslant 1/2a_+$ and $\zeta'(t) \leqslant 0$ whenever $\zeta(t) \geqslant 1/2a_-$. Since $\zeta(0)$ belongs to the interval $[1/2a_+, 1/2a_-]$, $\zeta(t)$ must remain in this interval for all t. □

Notation 4.5. We write $\mathbb{P}_{t_0,x_0}$ to indicate that we consider the solution of (4.4) with initial condition x_0, starting at time t_0. The associated expectation is denoted by $\mathbb{E}_{t_0,x_0}$.

Now we are ready to state and proof our main result for the stable case, see also Figure 3. We will follow the presentation in [8].

Theorem 4.6 (Stable case). *There exist $\varepsilon_0 > 0$, $c_0 > 0$, $h_0 > 0$ and $L > 0$, depending only on f, such that for $0 < \varepsilon \leqslant \varepsilon_0$, $h \leqslant h_0$ and $|x_0 - x^\star(0)| \leqslant c_0$,*

$$\mathbb{P}_{0,x_0}\big\{\tau_{\mathcal{B}_{\mathrm{s}}(h)} < t\big\} \leqslant C(t,\varepsilon)\exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\big[1 - L\varepsilon - Lh\big]\Big\}, \tag{4.48}$$

where

$$C(t,\varepsilon) = \frac{1}{\sqrt{\varepsilon}}\bigg(\frac{|\alpha(t)|}{\varepsilon^2} + 1\bigg). \tag{4.49}$$

Remark 4.7. The result of the preceding theorem remains true when $1/2|a(0)|$ in the definition (4.41) of $\zeta(t)$ is replaced be an arbitrary ζ_0, *provided* $\zeta_0 > 0$. The constant L in the exponent may then depend on ζ_0. Note that $\zeta(t)$ and $v(t)/\sigma^2$ are both solutions of the same ODE $\varepsilon z' = 2a(t)z + 1$, with possibly different initial conditions. If $|x_0 - x^\star(0)| = \mathcal{O}(\varepsilon)$, then $\zeta(t)$ is an adiabatic solution (in the sense of Theorem 4.2) of the differential equation, staying close to the equilibrium branch $z^\star(t) = 1/|2a(t)|$.

Estimate (4.48) has been designed for situations where $\sigma \ll 1$, and is useful for $\sigma \ll h \ll 1$. The exponent is optimal in this case (see Lemma 4.8), but we do not attempt to optimize the prefactor $C(t,\varepsilon)$, which leads only to subexponential corrections.

The t-dependence of the prefactor is to be expected. It is due to the fact that as time increases, the probability of x_t escaping from a neighbourhood of $x_t^{\det}$ also increases. Let us compare (4.48) with the results in the static case, cf. Theorem 3.14. For simplicity, we assume that $V(x,t) \equiv V(x) = a_0 x^2/2$ and that we start in $x_0 = x^\star = 0$. Then $\zeta_s \equiv 1/2a_0$ and (4.48) shows that[8]

$$\mathbb{P}^{0,x_0}\bigl\{\tau_{\mathcal{B}_s(h)} < t\bigr\} \leqslant \frac{1}{\sqrt{\varepsilon}}\Bigl(\frac{a_0 t}{\varepsilon^2}+1\Bigr)\exp\Bigl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\bigl[1-L\varepsilon\bigr]\Bigr\}. \tag{4.50}$$

Thus, the probability that x_s reaches the level $h/\sqrt{2a_0}$ before time t is small as long as

$$t \ll \frac{\varepsilon^2\sqrt{\varepsilon}}{a_0}\exp\Bigl\{\frac{1}{2}\frac{h^2}{\sigma^2}\bigl[1-L\varepsilon\bigr]\Bigr\}. \tag{4.51}$$

The boundary $\pm h\sqrt{\zeta_s}$ of the strip $\mathcal{B}_s(h)$ corresponds to the height

$$H = V(h/\sqrt{2a_0}) - V(0) = \frac{h^2}{4} \tag{4.52}$$

in terms of the potential V. The quasipotential of the LDP is constant on the boundary, and an application of Theorem 3.14 shows that

$$\lim_{\sigma\to 0}\mathbb{P}^{0,0}\bigl\{e^{(h^2-\delta)/2\sigma^2} \leqslant \varepsilon\tau_{\mathcal{B}_s(h)} \leqslant e^{(h^2+\delta)/2\sigma^2}\bigr\} = 1 \qquad \forall\delta>0. \tag{4.53}$$

Here the factor ε is due to the fact that we are working in slow time. In particular, we see that x_t is unlikely to exit from $\mathcal{B}_s(h)$ as long as $t < e^{(h^2-\delta)/2\sigma^2}/\varepsilon$ for some $\delta > 0$. Thus a comparison with (4.51) shows that (4.48) generalizes the lower bound in (4.53) to the case of non-static potentials.

PROOF OF THEOREM 4.6. First of all, check (via Itô's formula) that $y_t = x_t - x_t^{\det}$ satisfies

$$\begin{aligned} y_t &= y_0\, e^{\alpha(t)/\varepsilon} + \frac{1}{\varepsilon}\int_0^t \tilde b(y_s,s)\, e^{\alpha(t,s)/\varepsilon}\,ds + \frac{\sigma}{\sqrt{\varepsilon}}\int_0^t e^{\alpha(t,s)/\varepsilon}\,dW_s \\ &= y_t^0 + \frac{1}{\varepsilon}\int_0^t \tilde b(y_s,s)\, e^{\alpha(t,s)/\varepsilon}\,ds, \end{aligned} \tag{4.54}$$

where $y_0 = 0$, and, therefore, y_t^0 starts in 0 at time 0.

We begin by considering an arbitrary interval $[s,t] \subset I$. Then

$$\mathbb{P}^{0,x_0}\bigl\{\tau_{\mathcal{B}_s(h)} \in [s,t)\bigr\} \leqslant \mathbb{P}^{0,x_0}\bigl\{\exists u\in[s,t) : x_u \notin \mathcal{B}_s(h)\bigr\} = \mathbb{P}^{0,x_0}\Bigl\{\sup_{u\in[s,t)}\frac{|y_u|}{\sqrt{\zeta(u)}} \geqslant h\Bigr\}. \tag{4.55}$$

We would like to estimate the probability on the right-hand side by Doob's submartingale inequality, but neither y_u nor $y_u/\sqrt{\zeta(u)}$ is a (sub-)martingale. The remedy here is to consider only short intervals $[s,t)$, where $y_u/\sqrt{\zeta(u)}$ can be approximated locally by a

[8] In the proof of Theorem 4.6 we shall see that in the case of a linear drift coefficient, the h-dependent error term in the exponent is not present.

martingale. Thus we assume for the moment that t and s are such that $\Delta = |\alpha(t,s)|/\varepsilon$ is small.

For brevity, we shall write τ for $\tau_{\mathcal{B}_s(h)}$.

Reduction to the linearized process:
Let us first convince ourselves that y_t^0 gives the main contribution to y_t whenever $\tau \geqslant t$. In fact, for $\tau \geqslant t$,

$$\begin{aligned} \bigl|y_t - y_t^0\bigr| &= \biggl|\frac{1}{\varepsilon}\int_0^t \tilde{b}(y_u,u)\,\mathrm{e}^{\alpha(t,u)/\varepsilon}\,\mathrm{d}u\biggr| \leqslant M\frac{1}{\varepsilon}\int_0^t y_u^2\,\mathrm{e}^{\alpha(t,u)/\varepsilon}\,\mathrm{d}u \\ &\leqslant Mh^2\frac{1}{\varepsilon}\int_0^t \zeta(u)\,\mathrm{e}^{\alpha(t,u)/\varepsilon}\,\mathrm{d}u \leqslant \frac{M}{2a_-^2}h^2\bigl[1-\mathrm{e}^{\alpha(t)/\varepsilon}\bigr] \leqslant \frac{M}{2a_-^2}h^2, \end{aligned} \tag{4.56}$$

where we used the upper bound on $\zeta(u)$ that is provided by (4.45) and estimated $|a(u)|$ below by a_-. Therefore, again using (4.45), we find that for $\tau \geqslant t$,

$$\frac{\bigl|y_t - y_t^0\bigr|}{\sqrt{\zeta(t)}} \leqslant M\frac{\sqrt{a_+}}{\sqrt{2a_-^2}}h^2, \tag{4.57}$$

which implies that

$$\mathbb{P}_{0,x_0}\bigl\{\tau_{\mathcal{B}_s(h)} \in [s,t)\bigr\} \leqslant \mathbb{P}_{0,x_0}\biggl\{\sup_{u\in[s,t\wedge\tau)}\frac{|y_u^0|}{\sqrt{\zeta(u)}} \geqslant \tilde{h}\biggr\}, \tag{4.58}$$

where $\tilde{h} = h[1 - M(a_+^{1/2}/2a_-^2)h]$.

Estimating the linearization:
The right-hand side of (4.58) can by estimated by

$$\begin{aligned} \mathbb{P}_{0,x_0}&\biggl\{\sup_{u\in[s,t\wedge\tau)}\frac{|y_u^0|}{\sqrt{\zeta(u)}} \geqslant \tilde{h}\biggr\} \\ &\leqslant \mathbb{P}_{0,x_0}\biggl\{\sup_{u\in[s,t)}\frac{1}{\sqrt{\zeta(u)}}\biggl|\frac{\sigma}{\sqrt{\varepsilon}}\int_0^u \mathrm{e}^{\alpha(u,v)/\varepsilon}\,\mathrm{d}W_v\biggr| \geqslant \tilde{h}\biggr\} \\ &\leqslant \mathbb{P}_{0,x_0}\biggl\{\sup_{u\in[s,t)}\biggl|\frac{1}{\sqrt{\varepsilon}}\int_0^u \mathrm{e}^{\alpha(t,v)/\varepsilon}\,\mathrm{d}W_v\biggr| \geqslant \frac{\tilde{h}}{\sigma}\inf_{u\in[s,t)}\sqrt{\zeta(u)}\,\mathrm{e}^{\alpha(t,u)/\varepsilon}\biggr\}. \end{aligned} \tag{4.59}$$

Now, the stochastic integral $Y_u = \int_0^u \mathrm{e}^{\alpha(t,v)/\varepsilon}\,\mathrm{d}W_v$ in the last line is a martingale, so that $\mathrm{e}^{\gamma Y_u^2}$ is a submartingale, for any $\gamma > 0$. Thus we can apply Doob's submartingale inequality (see for instance [19, Theorem 1.3.8]), which yields

$$\mathbb{P}_{0,x_0}\biggl\{\sup_{u\in[s,t)}\biggl|\frac{1}{\sqrt{\varepsilon}}Y_u\biggr| \geqslant H\biggr\} = \mathbb{P}_{0,x_0}\biggl\{\sup_{u\in[s,t)}\mathrm{e}^{\gamma Y_u^2/\varepsilon} \geqslant \mathrm{e}^{\gamma H^2}\biggr\} \leqslant \mathrm{e}^{-\gamma H^2}\,\mathbb{E}_{0,x_0}\bigl\{\mathrm{e}^{\gamma Y_t^2/\varepsilon}\bigr\} \tag{4.60}$$

for any $H > 0$. Since Y_t is a Gaussian random variable, the expectation can be calculated using completion of squares. As Y_t has mean zero and variance $v(t)/\sigma^2$, we find

$$\mathbb{E}_{0,x_0}\bigl\{\mathrm{e}^{\gamma Y_t^2/\varepsilon}\bigr\} = \frac{1}{\sqrt{1-2\gamma v(t)/\sigma^2}}, \tag{4.61}$$

and from (4.40) (with $v_0 = 0$) we know that $v(t)/\sigma^2 \leqslant \zeta(t)$.

It remains to estimate

$$H := \frac{\tilde{h}}{\sigma} \inf_{u \in [s,t)} \sqrt{\zeta(u)}\, \mathrm{e}^{\alpha(t,u)/\varepsilon}. \tag{4.62}$$

As $\alpha(t,u)$ is negative and monotone in u, $\alpha(t,u) \geqslant \alpha(t,s) = -\Delta\varepsilon$ for all $u \in [s,t)$. By (4.46), the width $\sqrt{\zeta(u)}$ of the strip satisfies

$$\zeta(u) \geqslant \zeta(t) - \frac{1}{\varepsilon}(t-u) \geqslant \zeta(t) - \frac{|\alpha(t,s)|}{a_- \varepsilon} \geqslant \zeta(t)\bigl(1 - \mathcal{O}(\Delta)\bigr). \tag{4.63}$$

Now,

$$H \geqslant \frac{\tilde{h}}{\sigma}\sqrt{\zeta(t)}\bigl(1 - \mathcal{O}(\Delta\varepsilon)\bigr)\,\mathrm{e}^{-\Delta} = \frac{h}{\sigma}\sqrt{\zeta(t)}\bigl[1 - \mathcal{O}(\Delta) - \mathcal{O}(h)\bigr] \tag{4.64}$$

follows, and writing $2\zeta(t)\gamma = \hat{\gamma}$, we find

$$\mathbb{P}^{0,x_0}\bigl\{\tau_{\mathcal{B}_\mathrm{s}(h)} \in [s,t)\bigr\} \leqslant \frac{1}{\sqrt{1 - 2\gamma v(t)/\sigma^2}}\,\mathrm{e}^{-\gamma H^2} \leqslant \frac{1}{\sqrt{1-\hat{\gamma}}}\,\mathrm{e}^{-\frac{\hat{\gamma}}{2}\frac{h^2}{\sigma^2}[1-\mathcal{O}(\Delta)-\mathcal{O}(h)]}. \tag{4.65}$$

Somewhat arbitrarily, we opt for $\hat{\gamma} = 1 - \varepsilon$ and choose $\Delta = \varepsilon$. (Note that $\Delta = h$ would also be appropriate, unless h is very small.) Thereby we obtain

$$\mathbb{P}^{0,x_0}\bigl\{\tau_{\mathcal{B}_\mathrm{s}(h)} \in [s,t)\bigr\} \leqslant \frac{1}{\sqrt{\varepsilon}}\,\mathrm{e}^{-\frac{1}{2}\frac{h^2}{\sigma^2}[1-\mathcal{O}(\varepsilon)-\mathcal{O}(h)]}. \tag{4.66}$$

Longer time intervals:
So far, we have only dealt with a short interval $[s,t)$, where "short" is characterized by the requirement $\Delta = |\alpha(t,s)|/\varepsilon = \varepsilon$. To handle arbitrary intervals $[0,t)$, we now introduce a partition $0 = u_0 < u_1 < \dots < u_K = t$, where we require that $\Delta = |\alpha(u_k, u_{k+1})|/\varepsilon = \varepsilon$ for all $k = 0, \dots, K-2$. Therefore,

$$K = \left\lceil \frac{|\alpha(t)|}{\varepsilon^2} \right\rceil \tag{4.67}$$

is the smallest integer equal or larger than $|\alpha(t)|/\varepsilon^2$. Since $|a(u)|$ is bounded above and below, K is of order $1 + t/\varepsilon^2$. The trivial estimate

$$\begin{aligned}\mathbb{P}^{0,x_0}\bigl\{\tau_{\mathcal{B}_\mathrm{s}(h)} \in [0,t)\bigr\} &\leqslant \sum_{k=0}^{K-1} \mathbb{P}^{0,x_0}\bigl\{\tau_{\mathcal{B}_\mathrm{s}(h)} \in [u_k, u_{k+1})\bigr\} \\ &\leqslant \left\lceil \frac{|\alpha(t)|}{\varepsilon^2} \right\rceil \frac{1}{\sqrt{\varepsilon}}\,\mathrm{e}^{-\frac{1}{2}\frac{h^2}{\sigma^2}[1-\mathcal{O}(\varepsilon)-\mathcal{O}(h)]} \\ &\leqslant \frac{1}{\sqrt{\varepsilon}}\left\lceil \frac{|\alpha(t)|}{\varepsilon^2} \right\rceil \mathrm{e}^{-\frac{1}{2}\frac{h^2}{\sigma^2}[1-\mathcal{O}(\varepsilon)-\mathcal{O}(h)]}\end{aligned} \tag{4.68}$$

completes the proof. □

The following lemma augments Theorem 4.6 by providing a lower bound for the probability that x_t leaves the set $\mathcal{B}_\mathrm{s}(h)$ before time t.

Lemma 4.8. *There exist numerical constants* $c, C > 0$ *such that*

$$\mathbb{P}_{0,x_0}\Bigl\{\tau_{\mathcal{B}_s(h)} < t\Bigr\} \geqslant C \exp\Bigl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\bigl[1 + \mathcal{O}(\varepsilon) + \mathcal{O}(h)\bigr]\Bigr\} \tag{4.69}$$

holds for all $0 < \varepsilon \leqslant \varepsilon_0$, $c\sigma \leqslant h \leqslant h_0$, $|x_0 - x^\star(0)| \leqslant c_0$ *and all* t *such that* $|\alpha(t)| \geqslant \varepsilon|\log\varepsilon|$.

This estimate is rather crude and can certainly be improved. In particular, the prefactor should by an increasing function of time. Here we restrict ourselves to the simplest statement which shows that the exponent in (4.48) is optimal.

4.4 Near bifurcation points: Stochastic resonance, hysteresis, and bifurcation delay

Let us finally summarize some results on the case when Assumption 4.1 is violated and the curvature of the potential near its wells does not remain bounded away from zero, so that wells may become flat. We are interested in what will happen when either a well becomes flat and simultaneously the barrier between potential wells becomes low or when a potential changes from one-well to double-well. We only discuss standard examples, as the generalization to more general potentials with possibly more than two wells is immediate (in one-dimensional systems). Our presentation mainly follows [4], and the proofs of the various results can be found in [6, 7, 5].

We consider the Ginzburg–Landau potential

$$V(x,t) = \frac{1}{4}x^4 - \frac{1}{2}\mu(t)x^2 + \lambda(t)x, \tag{4.70}$$

where $\mu = \mu(t)$ and $\lambda = \lambda(t)$ are parameters. This potential has two wells if $27\lambda^2 < 4\mu^3$ and one well if $27\lambda^2 > 4\mu^3$. Crossing the lines $27\lambda^2 = 4\mu^3$, $\mu > 0$, corresponds to a saddle–node bifurcation, and crossing the point $\lambda = \mu = 0$ corresponds to a pitchfork bifurcation. Equilibrium points are solutions of the equation $x^3 - \mu(t)x - \lambda(t) = 0$; we will denote stable equilibria by $x^*_\pm(t)$, and the saddle, when present, by $x^*_0(t)$.

4.4.1 Stochastic resonance

Let us first consider the case where μ is a positive constant, say $\mu = 1$, and $\lambda(t)$ varies periodically, say $\lambda(t) = -A\cos(2\pi t)$. If $|\lambda| < \lambda_c = 2/(3\sqrt{3})$, then V is a double-well potential. We thus assume that $A < \lambda_c$. The SDE (4.4) then becomes

$$\mathrm{d}x_t = \frac{1}{\varepsilon}[x_t - x_t^3 - A\cos(2\pi t)]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t. \tag{4.71}$$

In the absence of noise, the existence of the potential barrier prevents the solutions from switching between potential wells, see Figure 4. If noise is present, but there is no periodic driving ($A = 0$), solutions will cross the potential barrier at random times, and whose expectation is given by Kramers' time $\varepsilon\,\mathrm{e}^{2H/\sigma^2}$, where H is the height of the barrier ($H = 1/4$ in this case), cf. the classical Eyring–Kramers formula (3.64).

Interesting things happen when both noise and periodic driving $\lambda(t)$ are present. Then the potential barrier will still be crossed at random times, but with a higher probability near the instants of minimal barrier height (i. e., when t is integer or half-integer). This phenomenon produces peaks in the power spectrum of the signal, hence the name *stochastic resonance* (SR).

If the noise intensity is sufficiently large compared to the minimal barrier height, transitions become likely twice per period (sample paths switch back and forth between wells), so that the signal x_t is close, in some sense, to a periodic function (Figure 5). The amplitude of this oscillation may be considerably larger than the amplitude of the forcing $\lambda(t)$, so that the mechanism can be used to amplify weak periodic signals. This phenomenon is also known as *noise-induced synchronization*. Of course, too large noise intensities will spoil the quality of the signal.

The mechanism of stochastic resonance was originally introduced as a possible explanation of the close-to-periodic appearance of the major Ice Ages [3, 2]. Here the

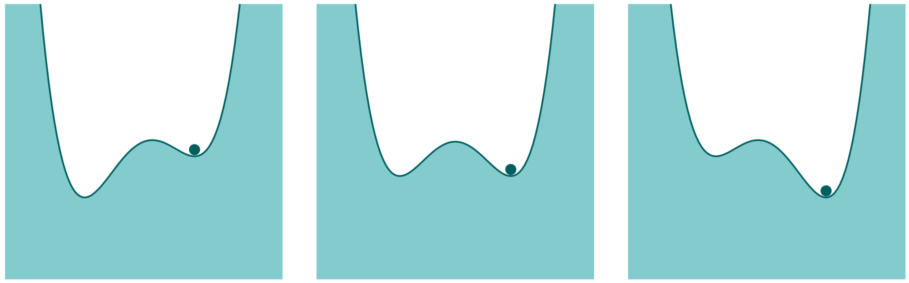

FIGURE 4. The potential $V(x,t) = \frac{1}{4}x^4 - \frac{1}{2}x^2 - A\cos(2\pi t)x$. For $\cos(2\pi t) = 0$, the potential is symmetric (middle), for integer times, the right-hand well approaches the saddle (left), while for half-integer times, the left-hand well approaches the saddle (right). If the amplitude A is smaller than the threshold λ_c, there is always a potential barrier, which an overdamped particle cannot overcome in the deterministic case. Sufficiently strong noise, however, helps the particle to switch from the shallower to the deeper well. This effect is the stronger the lower the barrier is, so that switching typically occurs close to the instants of minimal barrier height.

(quasi-)periodic forcing is caused by variations in the Earth's orbital parameters (Milankovich factors), and the additive noise models the fast unpredictable fluctuations caused by the weather. Meanwhile, SR has been detected in a large number of systems (see for instance [21, 29, 16] for reviews), including ring lasers, electronic devices, and even the sensory system of crayfish and paddlefish [22].

Recall that we assume that $A < \lambda_c$, so that there are always two stable equilibria at $x^*_\pm(t)$ and a saddle at $x^*_0(t)$. We introduce a parameter $a_0 = \lambda_c - A$ which measures the minimal barrier height: At $t = 0$, the barrier height is of order $a_0^{3/2}$ for small a_0, and the distance between $x^*_+(t)$ and the saddle at $x^*_0(t)$ is of order $\sqrt{a_0}$. At $t = 1/2$, the left-hand potential well at $x^*_-(t)$ is likewise close to the saddle. In order for transitions to become possible on a time scale which is not exponentially large, we allow a_0 to become small with ε.

Assume that we start at time $t_0 = -1/4$ in the basin of attraction of the left-hand potential well. Theorem 4.6 shows that transitions are unlikely for $t \ll 0$. Also, for $0 \ll t \ll 1/2$, paths will be concentrated either near $x^*_+(t)$ or near $x^*_-(t)$. This allows us to define the *transition probability* as

$$P_{\text{trans}} = \mathbb{P}_{t_0,x_0}\{x_{t_1} < 0\}, \qquad t_0 = -1/4, \quad t_1 = +1/4. \tag{4.72}$$

The properties of P_{trans} do not depend sensitively of the choices of t_0 and t_1, as long as $-1/2 \ll t_0 \ll 0 \ll t_1 \ll +1/2$. Also the level 0 can be replaced by any level lying between $x^*_-(t)$ and $x^*_+(t)$ for all t.

Theorem 4.9 ([7, Theorems 2.6 and 2.7]). *For the noise intensity, there is a threshold level $\sigma_c = (a_0 \vee \varepsilon)^{3/4}$ with the following properties:*

1. *If $\sigma < \sigma_c$, then*
$$P_{\text{trans}} \leqslant \frac{C}{\varepsilon}\,\mathrm{e}^{-\kappa\sigma_c^2/\sigma^2} \tag{4.73}$$
for some $C, \kappa > 0$. Paths are concentrated in a strip of width $\sigma/(\sqrt{|t|} \vee \sigma_c^{1/3})$ around the deterministic solution tracking $x^*_+(t)$ (Figure 5, upper figure).
2. *If $\sigma > \sigma_c$, then*
$$P_{\text{trans}} \geqslant 1 - C\,\mathrm{e}^{-\kappa\sigma^{4/3}/(\varepsilon|\log\sigma|)} \tag{4.74}$$

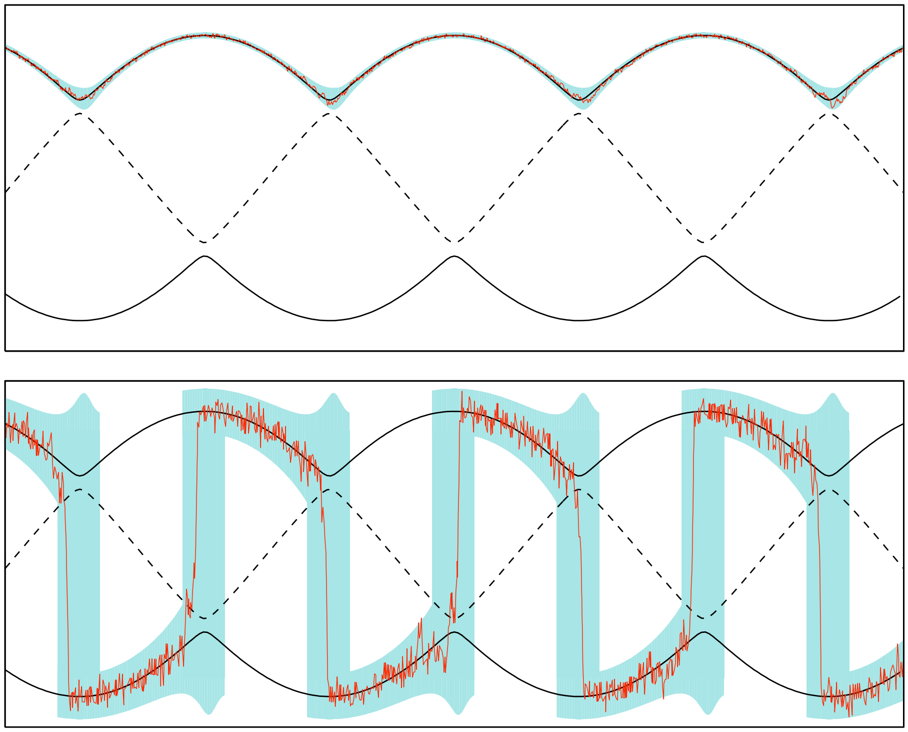

FIGURE 5. Sample paths of the SDE (4.71) for $\varepsilon = a_0 = 0.005$, and $\sigma = 0.02$ (upper picture) and $\sigma = 0.14$ (lower picture). Full curves represent the location of potential wells, the broken curve represents the saddle. For weak noise, the path x_t is likely to stay in the shaded set $\mathcal{B}(h)$, centred at the deterministic solution tracking the right-hand well. The maximal width of $\mathcal{B}(h)$ is of order $h\sigma/(a_0 \vee \varepsilon)^{1/4}$ and is reached at half-integer times. For strong noise, typical paths stay in the shaded set which switches back and forth between the wells at integer and half-integer times. The width of the vertical strips is of order $\sigma^{2/3}$. The "bumps" are due to the fact that one of the wells becomes very flat during the transition window so that paths might also make excursions away from the saddle.

for some $C, \kappa > 0$. Transitions are concentrated in the interval $[-\sigma^{2/3}, \sigma^{2/3}]$. Moreover, for $t \leqslant -\sigma^{2/3}$, paths are concentrated in a strip of width $\sigma/\sqrt{|t|}$ around the deterministic solution tracking $x^\star_+(t)$, while for $t \geqslant \sigma^{2/3}$, they are concentrated in a strip of width $\sigma/\sqrt{t}$ around a deterministic solution tracking $x^\star_-(t)$ (Figure 5, lower figure).

The crossover is quite sharp: For $\sigma \ll \sigma_c$, transitions between potential wells are very unlikely, while for $\sigma \gg \sigma_c$, they are very likely. By "concentrated in a strip of width w", we mean that the probability that a path leaves a strip of width hw decreases like $e^{-\kappa h^2}$ for some $\kappa > 0$. The "typical width w" is our measure of the deviation from the deterministic periodic function, which tracks one potential well in the small-noise case, and switches back and forth between the wells in the large-noise case.

Theorem 4.9 implies in particular that for the periodic signal's amplification to be optimal, the noise intensity σ should exceed the threshold σ_c. Larger noise intensities will increase both spreading of paths (especially just before they cross the potential barrier) and size of transition window, and thus spoil the output's periodicity.

We observe that for $\sigma_c \ll \sigma \ll 1$, paths are concentrated in the right-hand well when $\sin(2\pi t) < 0$, and in the left-hand well when $\sin(2\pi t) > 0$. They switch between wells near integer and half-integer times, when the barrier is lowest. The distribution of x_t is thus shifted in time with respect to the stationary density $\mathrm{e}^{-2V(x,t)/\sigma^2}/N$ of the frozen system, which has most of its mass concentrated in the deeper well and therefore favours the left-hand well whenever $\cos(2\pi t) > 0$, and favours the right-hand well whenever $\cos(2\pi t) < 0$. Since paths may jump from the shallower to the deeper well at a time of order $\sigma^{2/3}$ before the instant of lowest potential barrier, increasing σ reduces the time during which the system is in metastable equilibrium in the shallower potential well. For sufficiently strong noise, the density tracks the instantaneous stationary density, but it lags behind for weaker noise intensity.

Part of our results should appear quite natural. If $a_0 > \varepsilon$, the threshold noise level $\sigma_c = a_0^{3/4}$ behaves like the square root of the minimal barrier height. However, σ_c saturates at $\varepsilon^{3/4}$ for all $a_0 \leqslant \varepsilon$. Hence, even driving amplitudes arbitrarily close to λ_c cannot increase the transition probability. This is a rather subtle dynamical effect, mainly due to the fact that even if the barrier vanishes at $t = 0$, it is lower than $\varepsilon^{3/2}$ during too short a time interval for paths to take advantage. The situation is the same as if there were an "effective potential barrier" of height proportional to σ_c^2. Another remarkable fact is that for $\sigma > \sigma_c$, neither the transition probability nor the width of the transition windows depend on the driving amplitude to leading order.

4.4.2 Hysteresis

We continue to study the SDE (4.71)

$$\mathrm{d}x_t = \frac{1}{\varepsilon}[x_t - x_t^3 - A\cos(2\pi t)]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \tag{4.75}$$

but we also allow for amplitudes $A \geqslant \lambda_c$, so that twice per period the potential barrier may vanish. We want to study the behaviour of x_t not as a function of t but as a function of the associated value of the parameter $\lambda(t) = -A\cos(2\pi t)$. We use $a_0 := \lambda_c - A$ to compare the amplitude to the critical one. Then $a_0 > 0$ corresponds to the situation discussed above. Here, a_0 may change sign.

In the deterministic case, we can give a rough description as follows: For small amplitudes, i.e., $A \leqslant \lambda_c + \mathcal{O}(\varepsilon)$, the deterministic particle remains in the initial well at all

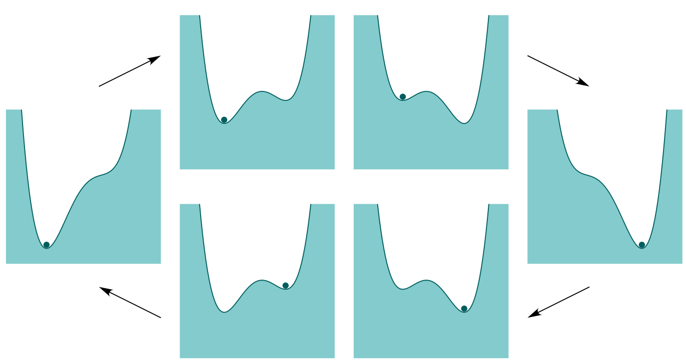

FIGURE 6. The potential $V(x,t) = \frac{1}{4}x^4 - \frac{1}{2}x^2 + \lambda(t)x$, with $\lambda(t) = -A\cos(2\pi t)$, when A exceeds λ_c. In the deterministic case, with $\varepsilon \ll 1$, the overdamped particle jumps to a new well whenever $|\lambda(t)|$ becomes larger than λ_c, leading to hysteresis.

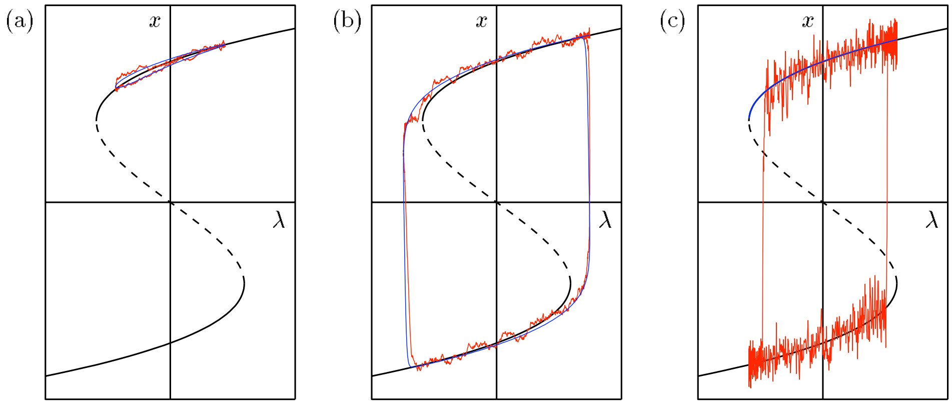

FIGURE 7. Typical random hysteresis "cycles" in the three parameter regimes. (a) Driving amplitude A and noise intensity σ are too small to allow the path to switch potential wells. (b) For large amplitude but weak noise, the path tracks its deterministic counterpart, which switches wells. (c) For sufficiently strong noise, the path can overcome the potential barrier, even before the barrier is lowest or vanishes.

times, while for large amplitudes, i. e., $A \geqslant \lambda_c + \mathcal{O}(\varepsilon)$, the particle switches well twice per period, see Figure 6.

If we add noise to the system, i. e., if we study the SDE (4.75), we find that there are three regimes (see Figure 7), characterized as follows:

- *The small-amplitude regime:* $-a_0 \leqslant$ *const* ε and $\sigma \leqslant (|a_0| \vee \varepsilon)^{3/4}$.
  If the amplitude of the modulation is too small to allow for transitions in the absence of noise and the noise intensity is below threshold, then x_t is unlikely to switch wells. This regime corresponds to the subthreshold behaviour in Theorem 4.9, but extends to the case of a barrier vanishing for a short time.
- *The large-amplitude regime:* $-a_0 \geqslant$ *const* ε and $\sigma \leqslant (\varepsilon\sqrt{|a_0|})^{1/2}$.
  If the amplitude of the modulation is large enough to allow for transitions (even in the absence of noise) and the noise intensity is below threshold, then x_t is close to x_t^{det} and typically changes wells twice per period as x_t^{det} does.
- *The large-noise regime:* Either $-a_0 \leqslant \varepsilon$ and $\sigma \geqslant (|a_0| \vee \varepsilon)^{3/4}$ or $-a_0 \geqslant \varepsilon$ and $\sigma \geqslant (\varepsilon\sqrt{|a_0|})^{1/2}$.
  If the noise intensity is above threshold, then x_t typically switches wells twice per period. The remarkable fact is that this typically happens before the barrier is lowest or vanishes, and the (random) value of λ when the crossing occurs is concentrated around a deterministic value $\widehat{\lambda}$ which satisfies $\widehat{\lambda} = \lambda_c - C\sigma^{4/3}$. To leading order, C does not depend on ε or A. This regime corresponds to the regime above threshold in Theorem 4.9, but also includes the case of a vanishing barrier.

4.4.3 Bifurcation delay

Let us finally discuss the slow passage through a (symmetric) pitchfork bifurcation. We consider again the Ginzburg–Landau potential (4.70), but this time with $\lambda \equiv 0$, and a parameter $\mu(t)$ increasing monotonously through zero. For the sake of simplicity, we shall choose $\mu(t) \equiv t$. As μ changes from negative to positive, the potential transforms from a single-well to a double-well potential, a scenario known as *spontaneous symmetry breaking*.

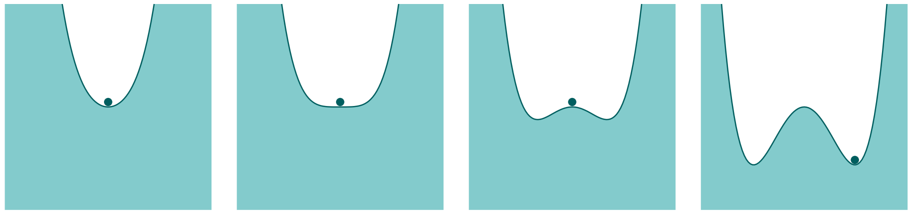

FIGURE 8. The potential $V(x,t) = \frac{1}{4}x^4 - \frac{1}{2}tx^2$ transforms, as μ changes from negative to positive, from a single-well to a double-well potential. In the deterministic case, an overdamped particle stays close to the saddle for a macroscopic time before falling into one of the wells. Noise tends to reduce this delay.

In fact, the symmetry of the potential is not broken, but the symmetry of the state may be, see Figure 8. Solutions tracking initially the potential well at $x = 0$ will choose between one of the new potential wells, but which one of the wells is chosen, and at what time, depends strongly on the noise present in the system.

We consider the nonlinear SDE

$$dx_t = \frac{1}{\varepsilon}\left[tx_t - x_t^3\right] dt + \frac{\sigma}{\sqrt{\varepsilon}} dW_t. \tag{4.76}$$

In the deterministic case $\sigma = 0$, the equation of motion reads

$$\varepsilon \frac{dx_t^{\mathrm{det}}}{dt} = tx_t^{\mathrm{det}} - (x_t^{\mathrm{det}})^3. \tag{4.77}$$

Its solution x_t^{det} with initial condition $x_{t_0}^{\mathrm{det}} = x_0 > 0$ can be written in the form

$$x_t^{\mathrm{det}} = c(x_0,t)\, e^{\alpha(t,t_0)/\varepsilon}, \qquad \alpha(t,t_0) = \int_{t_0}^{t} s\, ds = \frac{1}{2}(t^2 - t_0^2), \tag{4.78}$$

where the function $c(x_0,t)$ is found by substitution into (4.77). Its exact expression is of no importance here, it is sufficient to know that $0 < c(x_0,t) \leqslant x_0$ for all t.

If we start at a time $t_0 < 0$, the solution (4.78) will be attracted exponentially fast by the stable origin. The function $\alpha(t,t_0)$ is negative and decreasing for $t_0 < t < 0$, which implies in particular that x_0^{det} is exponentially small. For $t > 0$, the function $\alpha(t,t_0)$ is increasing, but it remains negative for some time. As a consequence, x_t^{det} remains close to the saddle up to the time $t = \Pi(t_0)$ for which $\alpha(t,t_0)$ reaches 0 again. Shortly after time $\Pi(t_0)$, the solution will jump to the potential well at $+\sqrt{t}$, unless x_0 is exponentially small. $\Pi(t_0)$ is called *bifurcation delay*, and depends only on $\mu(t)$ and t_0. Here, for $\mu(t) \equiv t$, $\alpha(t,t_0) = \frac{1}{2}(t^2 - t_0^2)$ and $\Pi(t_0) = |t_0|$.

The existence of a bifurcation delay may have undesired consequences. Assume for instance that we want to determine the bifurcation diagram of $\dot{x} = \mu x - x^3$ experimentally. Instead of measuring the asymptotic value of x_t for many different values of μ, which is time-consuming (especially near $\mu = 0$ where x_t decays only like $1/\sqrt{t}$), one may be tempted to vary μ slowly during the experiment. This, however, will fail to reveal part of the stable equilibrium branches because of the bifurcation delay. A similar phenomenon exists for the Hopf bifurcation [23, 24].

The delay is due to the fact that x_t approaches the origin exponentially closely. Noise of sufficient intensity will help driving the particle away from the saddle, and should therefore reduce the bifurcation delay. The obvious question is thus: How does the delay depend on the noise intensity?

We already know that $x_t^{\rm det}$ decreases exponentially fast for $t < 0$. Theorem 4.6 shows that paths are concentrated in a neighbourhood of order σ of $x_t^{\rm det}$ on any time interval $[t_0, t_1]$ bounded away from zero, so that we only need to worry about what happens after time t_1, when $x^{\rm det}$ is already exponentially small.

The dispersion of paths will be controlled by

$$\bar{v}(t) = \bar{v}_0 \, \mathrm{e}^{2\alpha(t,t_1)/\varepsilon} + \frac{\sigma^2}{\varepsilon} \int_{t_1}^{t} \mathrm{e}^{2\alpha(t,s)/\varepsilon} \, \mathrm{d}s, \tag{4.79}$$

where $\bar{v}_0$ is a positive constant. One can show that this function grows like $\sigma^2/|t|$ for $t \leqslant -\sqrt{\varepsilon}$, and remains of order $\sigma^2/\sqrt{\varepsilon}$ up to time $\sqrt{\varepsilon}$. Only after time $\sqrt{\varepsilon}$, $\bar{v}(t)$ grows exponentially fast. In analogy with (4.42), we define a strip

$$\mathcal{B}(h) = \bigl\{(x,t)\colon t_1 \leqslant t \leqslant \sqrt{\varepsilon}, |x - x_t^{\rm det}| < h\sqrt{\bar{v}(t)}\bigr\}. \tag{4.80}$$

In order to describe the behaviour for $t \geqslant \sqrt{\varepsilon}$, we further introduce the domain

$$\mathcal{D}(\varrho) = \bigl\{(x,t)\colon \sqrt{\varepsilon} \leqslant t \leqslant T, |x| \leqslant \sqrt{(1-\varrho)t}\bigr\}, \tag{4.81}$$

where ϱ is a parameter in $[0, 2/3)$. Note that $\mathcal{D}(2/3)$ contains those points in space–time where the potential is concave, while $\mathcal{D}(0)$ contains the points located between the bottoms of the wells. Then we have the following description of the behaviour of sample paths.

Theorem 4.10 ([6, Theorems 2.10–2.12])**.**

- *There is a constant $h_0 > 0$ such that for all $h \leqslant h_0\sqrt{\varepsilon}/\sigma$, the first-exit time $\tau_{\mathcal{B}(h)}$ of x_t from $\mathcal{B}(h)$ satisfies*

$$\mathbb{P}_{t_1,x_{t_1}}\bigl\{\tau_{\mathcal{B}(h)} < \sqrt{\varepsilon}\bigr\} \leqslant C_\varepsilon \, \mathrm{e}^{-\kappa h^2}, \tag{4.82}$$

where

$$C_\varepsilon = \frac{|\alpha(\sqrt{\varepsilon}, t_1)| + \mathcal{O}(\varepsilon)}{\varepsilon^2} \qquad \text{and} \qquad \kappa = \frac{1}{2} - \mathcal{O}(\sqrt{\varepsilon}) - \mathcal{O}\Bigl(\frac{\sigma^2 h^2}{\varepsilon}\Bigr). \tag{4.83}$$

- *Assume that $\sigma|\log\sigma|^{3/2} = \mathcal{O}(\sqrt{\varepsilon})$. Then for any $\varrho \in (0, 2/3)$, the first-exit time $\tau_{\mathcal{D}(\varrho)}$ of x_t from $\mathcal{D}(\varrho)$ satisfies*

$$\mathbb{P}_{\sqrt{\varepsilon},x_{\sqrt{\varepsilon}}}\bigl\{\tau_{\mathcal{D}(\varrho)} \geqslant t\bigr\} \leqslant C(t,\varepsilon) \frac{|\log\sigma|}{\sigma} \frac{\mathrm{e}^{-\varrho\alpha(t,\sqrt{\varepsilon})/\varepsilon}}{\sqrt{1 - \mathrm{e}^{-2\varrho\alpha(t,\sqrt{\varepsilon})/\varepsilon}}}, \tag{4.84}$$

where

$$C(t,\varepsilon) = \mathit{const}\; t\Bigl(1 + \frac{\alpha(t,\sqrt{\varepsilon})}{\varepsilon}\Bigr). \tag{4.85}$$

- *Assume x_t leaves $\mathcal{D}(\varrho)$ (with $1/2 < \varrho < 2/3$) through its upper (lower) boundary. Let $x_t^{\mathrm{det},\tau}$ be the deterministic solution starting at time $\tau = \tau_{\mathcal{D}(\varrho)}$ on the upper (lower) boundary of $\mathcal{D}(\varrho)$. Then $x_t^{\mathrm{det},\tau}$ approaches the equilibrium branch at $\sqrt{t}$ (resp., $-\sqrt{t}$) like $\varepsilon/t^{3/2} + \sqrt{\tau}\,\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon}$, where $\eta = 2 - 3\varrho$. Moreover, x_t is likely to stay in a strip centred at $x_t^{\mathrm{det},\tau}$, with width of order $\sigma/\sqrt{t}$, at least up to times of order 1.*

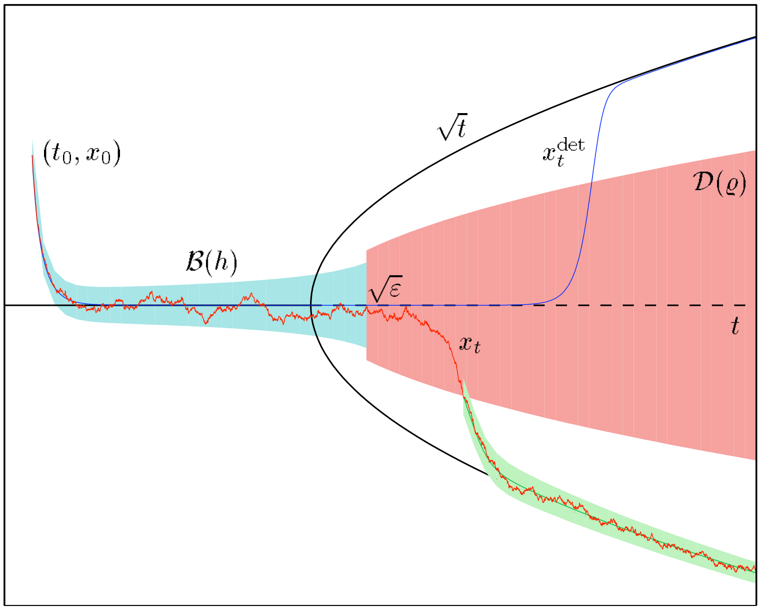

FIGURE 9. A sample path of the SDE (4.76), for $\varepsilon = 0.01$ and $\sigma = 0.015$. A deterministic solution is shown for comparison. Up to time $\sqrt{\varepsilon}$, the path remains in the set $\mathcal{B}(h)$ centred at x_t^{det}, shown here for $h = 3$. It then leaves the set $\mathcal{D}(\varrho)$ (here $\varrho = 2/3$) after a time of order $\sqrt{\varepsilon|\log \sigma|}$, after which it remains in a neighbourhood of the deterministic solution starting at the same time on the boundary of $\mathcal{D}(\varrho)$.

The bound (4.82), which is proved in a similar way as Theorem 4.6, shows that paths are unlikely to leave the strip $\mathcal{B}(h)$ if $1 \ll h \leqslant h_0\sqrt{\varepsilon}/\sigma$. If σ is smaller than $\sqrt{\varepsilon}$, paths remain concentrated in a neighbourhood of the origin up to time $\sqrt{\varepsilon}$, with a typical spreading growing like $\sigma/\sqrt{|t|}$ for $t \leqslant -\sqrt{\varepsilon}$, and remaining of order $\sigma/\varepsilon^{1/4}$ for $|t| \leqslant \sqrt{\varepsilon}$. This is again a dynamical effect: Although there is a saddle at the origin for positive times, its curvature is so small that paths do not have time to escape before $t = \sqrt{\varepsilon}$, see Figure 9.

Relation (4.84) yields an upper bound on the typical time needed to leave $\mathcal{D}(\varrho)$, and thus enter a region where the potential is convex. Since $\alpha(t, \sqrt{\varepsilon})$ grows like $\frac{1}{2}t^2$, the probability not to leave $\mathcal{D}(\varrho)$ before time t becomes small as soon as

$$t \gg \sqrt{\frac{2}{\varrho}\varepsilon|\log \sigma|}. \tag{4.86}$$

The last part of the theorem implies that another time span of the same order is needed for paths to concentrate again, around an adiabatic solution tracking the bottom of the well (at a distance of order $\varepsilon/t^{3/2}$). One can thus say that the typical bifurcation delay time of the dynamical pitchfork bifurcation with noise is of order $\sqrt{\varepsilon|\log \sigma|}$.

As a consequence, we can distinguish three parameter regimes:

I. *Exponentially small noise:* $\sigma \leqslant \mathrm{e}^{-K/\varepsilon}$ for some $K > 0$.
At time $\sqrt{\varepsilon}$, the spreading of paths is still exponentially small. In fact, one can extend Relation (4.82) to all times for which $\alpha(t, \sqrt{\varepsilon}) < K$. If K is larger than $\alpha(\Pi(t_0), 0)$, where $\Pi(t_0)$ is the deterministic bifurcation delay, most paths will track the deterministic solution and follow it into the right-hand potential well.

II. *Moderate noise:* $\mathrm{e}^{-1/\varepsilon^p} \leqslant \sigma \ll \sqrt{\varepsilon}$ for some $p < 1$.
The bifurcation delay lies between $\sqrt{\varepsilon}$ and a constant times $\sqrt{\varepsilon|\log \sigma|} \leqslant \varepsilon^{(1-p)/2}$ with high probability. One can thus speak of a "microscopic" bifurcation delay.

III. *Large noise:* $\sigma \geqslant \sqrt{\varepsilon}$.
The spreading of paths grows like $\sigma/\sqrt{|t|}$ at least up to time $-\sigma$. As t approaches the bifurcation time 0, the bottom of the potential well becomes so flat that the paths are no longer localized near the origin and may switch wells several times before eventually settling for a well. So for large noise intensities, the concept of bifurcation delay should be replaced by *two* variables, namely the first-exit time from a suitably chosen neighbourhood of the saddle and the time when the potential wells become attractive enough to counteract the diffusion.

One can also estimate the probability to reach the right-hand potential well rather than the left-hand potential well. Loosely speaking, if x_s reaches 0 before time t, it has equal probability to choose either potential well. It follows that

$$\mathbb{P}_{t_0,x_0}\{x_t \geqslant 0\} = \frac{1}{2} + \frac{1}{2}\mathbb{P}_{t_0,x_0}\{x_s > 0 \;\forall s \in [t_0,t]\}. \tag{4.87}$$

One can show that for $t = 0$ the second term on the right-hand side is of order

$$\frac{x_0\varepsilon^{1/4}}{\sigma}\,\mathrm{e}^{-|\alpha(0,t_0)|/\varepsilon}, \tag{4.88}$$

and for larger t, this term will be even smaller. Thus in case II, paths will choose one potential well or the other with a probability exponentially close to 1/2.

For our choice of $\mu(t)$, the height of the potential barrier grows without bound. This implies that once x_t has chosen a potential well, its probability *ever* to cross the saddle again is of order $\mathrm{e}^{-const/\sigma^2}$.

The existence of three parameter regimes has some interesting consequences on the experimental determination of a bifurcation diagram. Assume we want to determine the stable equilibrium branches by sweeping the parameter with speed ε. Regime II is the most favourable: In Regime I, part of the stable branches cannot be seen due to the bifurcation delay, while in Regime III, noise will blur the bifurcation diagram. For a given noise intensity σ, the sweeping rate ε should thus satisfy

$$\sigma^2 \ll \varepsilon \ll \bigl(1/|\log\sigma|\bigr)^{1/p} \tag{4.89}$$

in order to produce a good image of the stable equilibria. As long as $\sigma^2 \ll 1/|\log\sigma|^{1/p}$, increasing artificially the noise level allows to work with higher sweeping rates, but of course the image will be more and more blurred.

On the other hand, the relation between noise and delay can be used to measure the intensity of noise present in the system. If a bifurcation delay is observed for a sweeping rate ε_0, repeating the experiment with slower and slower sweeping rates should ultimately suppress the delay. If this happens for $\varepsilon = \varepsilon_1$, then the noise intensity is of order $\mathrm{e}^{-const/\varepsilon_1}$.

References

[1] S. Arrhenius. *J. Phys. Chem.*, 4:226, 1889.

[2] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani. A theory of stochastic resonance in climatic change. *SIAM J. Appl. Math.*, 43(3):565–578, 1983.

[3] R. Benzi, A. Sutera, and A. Vulpiani. The mechanism of stochastic resonance. *J. Phys. A*, 14(11):L453–L457, 1981.

[4] N. Berglund and B. Gentz. Beyond the Fokker–Planck equation: Pathwise control of noisy bistable systems. *J. Phys. A*, 35(9):2057–2091, 2002. DOI 10.1088/0305-4470/35/9/301.

[5] N. Berglund and B. Gentz. The effect of additive noise on dynamical hysteresis. *Nonlinearity*, 15(3):605–632, 2002. DOI 10.1088/0951-7715/15/3/305.

[6] N. Berglund and B. Gentz. Pathwise description of dynamic pitchfork bifurcations with additive noise. *Probab. Theory Related Fields*, 122(3):341–388, 2002. DOI 10.1007/s004400100174.

[7] N. Berglund and B. Gentz. A sample-paths approach to noise-induced synchronization: Stochastic resonance in a double-well potential. *Ann. Appl. Probab.*, 12:1419–1470, 2002.

[8] N. Berglund and B. Gentz. Geometric singular perturbation theory for stochastic differential equations. *J. Differential Equations*, 191:1–54, 2003. DOI 10.1016/S0022-0396(03)00020-2.

[9] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes I. Sharp asymptotics for capacities and exit times. Preprint WIAS-767, 2002.

[10] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes II. Precise asymptotics for small eigenvalues. Preprint WIAS-768, 2002.

[11] M. Day. Large deviations results for the exit problem with characteristic boundary. *J. Math. Anal. Appl.*, 147(1):134–153, 1990.

[12] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics*. Springer-Verlag, New York, second edition, 1998.

[13] J.-D. Deuschel and D. W. Stroock. *Large deviations*. Academic Press, Boston, 1989. Reprinted by the American Mathematical Society, 2001.

[14] H. Eyring. The activated complex in chemical reactions. *Journal of Chemical Physics*, 3:107–115, 1935.

[15] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer-Verlag, New York, second edition, 1998.

[16] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni. Stochastic resonance. *Rev. Mod. Phys.*, 70:223–287, 1998.

[17] I. S. Gradšteĭn. Application of A. M. Lyapunov's theory of stability to the theory of differential equations with small coefficients in the derivatives. *Mat. Sbornik N. S.*, 32(74):263–286, 1953.

[18] K. Ito. On stochastic differential equations. *Mem. Amer. Math. Soc.*, 1951(4), 1951.

[19] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*. Springer-Verlag, New York, 1988.

[20] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.

[21] F. Moss and K. Wiesenfeld. The benefits of background noise. *Scientific American*, 273:50–53, 1995.

[22] A. B. Neiman, D. F. Russell, X. Pei, W. Wojtenek, J. Twitty, E. Simonotto, B. A. Wettring, E. Wagner, L. A. Wilkens, and F. Moss. Stochastic synchronization of electroreceptors in the paddlefish. Technical report, Center for Neurodynamics, University of Missouri-St. Louis, 1999. `http://neurodyn.umsl.edu/~neiman/synchronization/`.

[23] A. I. Neĭshtadt. Persistence of stability loss for dynamical bifurcations I. *Differential Equations*, 23:1385–1391, 1987.

[24] A. I. Neĭshtadt. Persistence of stability loss for dynamical bifurcations II. *Differential Equations*, 24:171–176, 1988.

[25] B. Øksendal. *Stochastic differential equations.* Springer-Verlag, Berlin, 1985.

[26] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[27] A. N. Tihonov. Systems of differential equations containing small parameters in the derivatives. *Mat. Sbornik N. S.*, 31:575–586, 1952.

[28] H. von Weizsäcker and G. Winkler. *Stochastic integrals.* Friedr. Vieweg & Sohn, Braunschweig, 1990.

[29] K. Wiesenfeld and F. Moss. Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs. *Nature*, 373:33–36, 1995.