Large deviations and stochastic resonance

P. Imkeller, C. Hein
Lecture Course SS 2009

November 30, 2015

Contents

Chapter 1

Introduction

The trigger of stochastic resonance will be diffusion exits from domains. The main aim of this chapter is to determine the exit times of randomly perturbed dynamical systems from domains of attraction of their stable fixed points. Think of a potential diffusion, physically the motion of an overdamped particle in a potential landscape. In dimension 1, we may think of a smooth double well potential function $U : \mathbf{R} \to \mathbf{R}$, such that for instance U is smooth, possesses exactly two local minima at ± 1, and a unique local maximum at 0. Further $\lim_{x\to\pm\infty} U(x) = \infty$. Then the differential equation

$$dX_t^{\epsilon,x} = -U'(X_t^{\epsilon,x})dt, \quad X_0^{\epsilon,x} = x,$$

generates a dynamical system which possesses two *stable* fixed points ± 1 and one unstable fixed point 0. If the initial state $x \in]-\infty, 0[$, we have $\lim_{t\to\infty} X_t^{\epsilon,x} = -1$, whereas $\lim_{t\to\infty} X_t^{\epsilon,x} = 1$, if $x \in]0, \infty[$. We call the intervals containing the fixed points *domains of attraction*. The particle traveling in the potential landscape described by the solutions of this differential equation can therefore not exit a domain of attraction, once starting its motion inside.

This feature of the motion changes drastically, once noise is added to the system. Let $(\Omega, \mathcal{F}, P)$ with the coordinate process $W = (W_t)_{t\geq 0}$ be a one-dimensional canonical Wiener space, with the canonical filtration $(\mathcal{F}_t)_{t\geq 0}$. So for $\epsilon > 0$ we can consider the stochastic differential equation

$$dX_t^{\epsilon,x} = -U'(X_t^{\epsilon,x})dt + \sqrt{\epsilon}dW_t, \quad X_0^{\epsilon,x} = x.$$

By adding noise, however small its intensity ϵ may be, transitions between the states ± 1 become possible. We then call the states *meta-stable*, and the domains of attraction lose their original meaning. However, in the *small noise limit*, i.e. as $\epsilon \to 0$, the particle starting in one of them will typically, i.e. with high probability, leave this domain in times depending on ϵ and given by the asymptotic quantities

$$\exp(\frac{2(U(0) - U(-1))}{\epsilon}) \quad \text{if} \quad x \in]-\infty, 0[,$$

and

$$\exp(\frac{2(U(0) - U(1))}{\epsilon}) \quad \text{if} \quad x \in]0, \infty[.$$

To show this, Freidlin and Wentzell applied the theory of *large deviations* for diffusion processes perturbed by small Gaussian noise. We shall develop the essentials of this theory, and then present the law of asymptotic exit times, noted before in papers by chemists and physicists occupied with phenomena of reaction- diffusion. In this context, large deviations concern the asymptotic behavior of the laws $\mu_\epsilon = P \circ X^{\epsilon,x}$, as $\epsilon \to 0$. In this case, μ_ϵ lives on the space of real valued continuous functions defined on $\mathbf{R}_+$, endowed with the topology of uniform convergence on compact subintervals of $\mathbf{R}_+$. In general, we let the measures live on some topological space $\mathbf{X}$ with system of Borel sets $\mathbf{B}$.

Chapter 2

Brownian motion via Fourier series

In this Chapter, we shall present Brownian motion in an approach based on Fourier series with respect to the orthonormal system of Haar functions. This approach will be seen to open an easy and fast route to large deviations principles for Brownian motion, the basic noise process added to deterministic dynamical systems to provide the time homogeneous randomly perturbed dynamical systems that are the main objects of interest for this Chapter. In fact, we shall present a direct proof of Schilder's Theorem which only uses this Fourier series representation and the large deviation principle for one-dimensional Gaussian variables. The basic idea of this approach for large deviations on function spaces is triggered by an observation by Ciesielski according to which smoothness properties of functions in Hölder spaces can be studied via a universal Banach space isomorphism through convergence properties of sequences. We first present Ciesielski's isomorphism.

2.1 The Ciesielski isomorphism of Hölder and sequence spaces

For $0 < \alpha \leq 1$ let $\mathcal{C}_\alpha([0,1])$ be the space of all α-Hölder continuous paths on $[0,1]$ starting in 0. This space is a Banach space endowed with the Hölder norm

$$\|f\|_\alpha = \sup_{0 \leq t < s \leq 1} \frac{|f(t) - f(s)|}{|t-s|^\alpha}.$$

Denote, moreover, by $\mathcal{C}$ the Banach space of all bounded sequences $\eta = (\eta_n)_{n \geq 0}$ endowed with the norm $||\eta||_\infty = \sup_{n \geq 0} |\eta_n|$. We call two Banach spaces *isomorphic* if there exists a one-to-one linear map between the spaces. By means of Fourier decomposition of Hölder continuous functions with the Schauder basis we first prove that $\mathcal{C}_\alpha([0,1])$ and $\mathcal{C}$ are isomorphic. For this purpose we introduce the *Haar functions.* For $t \in [0,1]$ let $\chi_0(t) \equiv 1$ and

$$\chi_{2^k+l}(t) = \begin{cases} \sqrt{2^k} & \text{if } t \in \left[\frac{2l}{2^{k+1}}, \frac{2l+1}{2^{k+1}}\right[, \quad l = 0, 1, \ldots, 2^k - 1, \quad k \geq 0, \\ -\sqrt{2^k} & \text{if } t \in \left[\frac{2l+1}{2^{k+1}}, \frac{2l+2}{2^{k+1}}\right[, \\ 0 & \text{otherwise.} \end{cases}$$

This basis is a complete orthonormal system of $L^2([0,1])$. The *Schauder functions* $(\phi_n)_{n\geq 0}$ are just the primitives of the Haar system, given for $t \in [0,1]$ by

$$\phi_n(t) = \int_0^t \chi_n(s)ds.$$

If $f \in \mathcal{C}([0,1])$ starting in 0 possesses a square integrable density $\dot{f}$ so that for $t \in [0,1]$ we have

$$f(t) = \int_0^t \dot{f}(s)ds,$$

we can write

$$\dot{f} = \sum_{n=0}^{\infty} \langle \chi_n, \dot{f} \rangle \chi_n,$$

and therefore

$$f = \sum_{n=0}^{\infty} \langle \chi_n, \dot{f} \rangle \phi_n.$$

Indeed, due to the fact that for $k \geq 0$ fixed, and $0 \leq l_1, l_2 \leq 2^k - 1, l_1 \neq l_2$ the supports of the functions $\phi_{2^k+l_1}$ and $\phi_{2^k+l_2}$ are disjoint and the functions uniformly bounded by $2^{-\frac{k}{2}-1}$, we may estimate for $K \in \mathbb{N}, q \geq p \geq 2^K$ by means of Cauchy-Schwarz' inequality

$$\begin{aligned} || \sum_{n=p}^{q} \langle \chi_n, \dot{f} \rangle \phi_n || & \qquad (2.1) \\ & \leq \sum_{k=K}^{\infty} || \sum_{l=0}^{2^k-1} \langle \chi_{2^k+l}, \dot{f} \rangle \phi_{2^k+l} || \\ & \leq \sum_{k=K}^{\infty} \sup_{0 \leq l \leq 2^k-1} |\langle \chi_{2^k+l}, \dot{f} \rangle| 2^{-\frac{k}{2}-1} \\ & \leq \sum_{k=K}^{\infty} [\int_0^1 \dot{f}^2(s)ds]^{\frac{1}{2}} 2^{-\frac{k}{2}-1}. \end{aligned}$$

This clearly implies the convergence of the series in the uniform norm. We shall now see by following Ciesielski [4] that this representation may be extended to Hölder spaces. For this purpose denote for $n = 2^k + l$

$$\langle \chi_n, df \rangle = \sqrt{2^k}[2f(\frac{2l+1}{2^{k+1}}) - f(\frac{2l+2}{2^{k+1}}) - f(\frac{2l}{2^{k+1}})].$$

This just gives the integral of the function χ_n with respect to f as an integrator.

Lemma 2.1.1. *Let $0 < \alpha \leq 1$, and let $f \in \mathcal{C}_\alpha([0,1])$. Then*

$$f = \sum_{n=0}^{\infty} \langle \chi_n, df \rangle \phi_n,$$

with convergence with respect to the uniform norm.

Proof. It can be seen easily that f may be approximated in the uniform norm by a sequence $(f_m)_{m\in\mathbb{N}}$ of functions possessing square integrable densities $(\dot{f}_m)_{m\in\mathbb{N}}$, and with α-Hölder norms bounded by the one of f. Take for instance a sequence obtained from f by smoothing with a sequence of smooth approximations of the unit. More precisely, let $\rho : [-1,1] \to \mathbb{R}_+$ be a C^∞ function such that $\int_{-1}^1 \rho(x)dx = 1$. For $m \in \mathbb{N}$, let $\rho_m = m\,\rho(m\cdot)$, and $f_m(t) = \int_{-1}^1 f(t-x)\rho_m(x)dx, t \in [0,1]$ (here we assume f to be trivially extended continuously to $[-1,2]$ by constant branches). Obviously, $(f_m)_{m\in\mathbb{N}}$ converges to f in the uniform norm, and for each m, f_m possesses a square integrable density $\dot{f}_m$. Moreover, for $s,t \in [0,1], m \in \mathbb{N}, \alpha \in]0,1]$

$$\frac{|f_m(t)-f_m(s)|}{|t-s|^\alpha} \leq \int_{-1}^1 \frac{|f(t-x)-f(s-x)|}{|t-s|^\alpha}\phi_m(x)dx,$$

hence $||f_m||_\alpha \leq ||f||_\alpha$.

Since we know from the above discussion that the desired representations hold for f_m for all $m \in \mathbb{N}$, a dominated convergence argument shows that we have to prove

$$\sup_{m\in\mathbb{N}} ||\sum_{n=p}^q \langle\chi_n, df_m)\rangle\phi_n|| \to 0 \quad \text{as} \quad p \leq q \to \infty. \tag{2.2}$$

To do this, we have to modify the estimate (2.1) a bit. In fact, for any $m \in \mathbb{N}$, and $K \in \mathbb{N}, q \geq p \geq 2^K$ we have

$$\begin{aligned} ||\sum_{n=p}^q \langle\chi_n, df_m\rangle\phi_n|| & \qquad\qquad (2.3)\\ \leq & \sum_{k=K}^\infty ||\sum_{l=0}^{2^k-1} \langle\chi_{2^k+l}, df_m\rangle\phi_{2^k+l}|| \\ \leq & \sum_{k=K}^\infty \sup_{0\leq l\leq 2^k-1} |\langle\chi_{2^k+l}, df_m\rangle| 2^{-\frac{k}{2}-1} \\ \leq & \sum_{k=K}^\infty ||f_m||_\alpha 2^{-\alpha k} \\ \leq & \ ||f||_\alpha \sum_{k=K}^\infty 2^{-\alpha k}. \end{aligned}$$

The latter expression obviously converges to 0 as $K \to \infty$. This completes the proof. □

The following Theorem states that $\mathcal{C}_\alpha([0,1])$ and $\mathcal{C}$ are isomorphic.

Theorem 2.1.2. *Let $0 < \alpha < 1$. For $\mathbb{N} \ni n = 2^k + l, k \geq 0, 0 \leq l \leq 2^k - 1$ let*

$$c_n(\alpha) = 2^{k(\alpha-\frac{1}{2})+\alpha-1}, \quad c_0(\alpha) = 1.$$

Define

$$T_\alpha : \mathcal{C}_\alpha([0,1]) \to \mathcal{C}, \quad f \mapsto (c_n(\alpha)\,\langle\chi_n, df\rangle)_{n\geq 0}.$$

Then

$$T_\alpha^{-1} : \mathcal{C} \to \mathcal{C}_\alpha([0,1]), \quad (\eta_n)_{n\geq 0} \mapsto \sum_{n=0}^{\infty} \frac{1}{c_n(\alpha)} \eta_n \ \phi_n.$$

T_α is an isomorphism, and for the operator norms we have the following inequalities

$$||T_\alpha|| = 1, \quad ||T_\alpha^{-1}|| \leq \frac{2}{(2^\alpha - 1)(2^{1-\alpha} - 1)}.$$

Proof. By definition, for $\mathbb{N} \ni n = 2^k + l, k \geq 0, 0 \leq l \leq 2^k - 1$ we have

$$|\langle \chi_n, df\rangle| \leq 2^{-(k+1)\alpha + \frac{k}{2} + 1} ||f||_\alpha = \frac{1}{c_n(\alpha)} ||f||_\alpha. \tag{2.4}$$

Therefore, T_α is well defined, and we have

$$||T_\alpha|| \leq 1.$$

Moreover, for $f(t) = t, 0 \leq t \leq 1$, we have $\langle \chi_0, df\rangle = 1$, while for $n \in \mathbb{N}$ we have $\langle \chi_n, df\rangle = 0$. Hence $||T_\alpha f||_\infty = ||f||_\alpha$. This implies $||T_\alpha|| = 1$. Lemma 2.1.1 shows that T_α is one-to-one and that T_α^{-1} is its inverse.

We next prove the inequality for the operator norm of T_α^{-1}. Let $\eta = (\eta_n)_{n\geq 0} \in \mathcal{C}$ be given, choose $0 \leq s < t \leq 1$, and write $f = T_\alpha^{-1}(\eta)$. Then we have

$$|f(t) - f(s)| \leq ||\eta||_\infty [|t-s| + \sum_{k=0}^{\infty} \sum_{l=0}^{2^k - 1} \frac{1}{c_{2^k+l}(\alpha)} |\phi_{2^k+l}(t) - \phi_{2^k+l}(s)|]. \tag{2.5}$$

Now choose $k_0 \geq 0$ such that

$$2^{-k_0 - 1} < |t-s| \leq 2^{-k_0}.$$

Then for $0 \leq k < k_0$ by definition of the Schauder functions

$$\begin{aligned} &\sum_{l=0}^{2^k-1} \frac{1}{c_{2^k+l}(\alpha)} |\phi_{2^k+l}(t) - \phi_{2^k+l}(s)| \\ &\leq \ 2^{-k(\alpha - \frac{1}{2}) - \alpha + 1} \ 2^{\frac{k}{2}} \ |t-s| \\ &\leq \ 2^{k(1-\alpha) - \alpha + 1 - k_0(1-\alpha)} \ |t-s|^\alpha = (2^{1-\alpha})^{(1+k-k_0)} \ |t-s|^\alpha, \end{aligned} \tag{2.6}$$

while for $k \geq k_0$

$$\begin{aligned} &\sum_{l=0}^{2^k-1} \frac{1}{c_{2^k+l}(\alpha)} |\phi_{2^k+l}(t) - \phi_{2^k+l}(s)| \\ &\leq \ 2^{-k(\alpha - \frac{1}{2}) - \alpha + 1} \ 2^{-\frac{k}{2}} \\ &\leq \ 2^{-k\alpha - \alpha + 1 + (k_0+1)\alpha} \ |t-s|^\alpha = (2^\alpha)^{(k_0 - k)} \ |t-s|^\alpha. \end{aligned} \tag{2.7}$$

Combining (2.5), (2.6) and (2.7), we obtain the estimate

$$\frac{|f(t) - f(s)|}{|t-s|^\alpha} \leq \frac{2}{(2^\alpha - 1)(2^{1-\alpha} - 1)} \, ||\eta||_\infty,$$

and therefore

$$||T_\alpha^{-1}|| \leq \frac{2}{(2^\alpha - 1)(2^{1-\alpha} - 1)}.$$

□

The spaces we are ultimately interested in are those in which almost all sample paths of the Brownian motion are living. We therefore have to extend the isomorphism of Theorem 2.1.2 to the following subspaces of Hölder continuous functions. For $0 < \alpha \leq 1$ let $\mathcal{C}_\alpha^0([0,1])$ be the subspace of $\mathcal{C}([0,1])$ composed of all functions f for which $f(0) = 0$ and

$$\lim_{\delta \to 0} \sup_{0 \leq s < t \leq 1, |t-s| \leq \delta} \frac{|f(t) - f(s)|}{|t-s|^\alpha} = 0.$$

The isomorphism of Theorem 2.1.2 will then be restricted to the subspace $\mathcal{C}_0$ of all sequences $\eta = (\eta_n)_{n \geq 0}$ in $\mathcal{C}$ which converge to 0 as $n \to \infty$ as a target space. The following Theorem holds.

Theorem 2.1.3. *Let $0 < \alpha < 1$. Let $c_n(\alpha), n \geq 0$, be defined as in Theorem 2.1.2. Define*

$$T_{\alpha,0} : \mathcal{C}_\alpha^0([0,1]) \to \mathcal{C}_0, \quad f \mapsto (c_n(\alpha) \langle \chi_n, df \rangle)_{n \geq 0}.$$

Then

$$T_{\alpha,0}^{-1} : \mathcal{C}_0 \to \mathcal{C}_\alpha^0([0,1]), \quad (\eta_n)_{n \geq 0} \mapsto \sum_{n=0}^{\infty} \frac{1}{c_n(\alpha)} \eta_n \ \phi_n.$$

$T_{\alpha,0}$ is an isomorphism, and for the operator norms we have the following inequalities

$$||T_{\alpha,0}|| = 1, \quad ||T_{\alpha,0}^{-1}|| \leq \frac{2}{(2^\alpha - 1)(2^{1-\alpha} - 1)}.$$

Proof. Note first that (2.4) can be strengthened by definition to read

$$|\langle \chi_n, df \rangle| \leq \frac{1}{c_n(\alpha)} \sup_{0 \leq s < t \leq 1, |t-s| \leq 2^{-k-1}} \frac{|f(t) - f(s)|}{|t-s|^\alpha}. \tag{2.8}$$

Hence, by definition of $\mathcal{C}_\alpha^0([0,1])$, we obviously obtain that $T_{\alpha,0}$ is well defined. To prove that also $T_{\alpha,0}^{-1}$ is well defined, we just have a closer inspection of the arguments that led to the operator norm inequality in the proof of Theorem 2.1.2. First, note that for $\eta = (\eta_n)_{n \geq 0} \in \mathcal{C}_0$, with $f = T_{\alpha,0}^{-1}(\eta)$

$$|f(t) - f(s)| \leq [|\eta_0(\alpha)| \, |t-s| + \sum_{k=0}^{\infty} \sum_{l=0}^{2^k - 1} \frac{1}{c_{2^k+l}(\alpha)} |\eta_{2^k+l}| \, |\phi_{2^k+l}(t) - \phi_{2^k+l}(s)|]. \tag{2.9}$$

Now choose again $k_0 \geq 0$ such that

$$2^{-k_0 - 1} < |t-s| \leq 2^{-k_0},$$

and denote $\tau_n = \sup_{k\geq n} |\eta_k|$. Then for $0 \leq k < k_0$

$$\begin{aligned}
&\sum_{l=0}^{2^k-1} \frac{1}{c_{2^k+l}(\alpha)} |\eta_{2^k+l}| \, |\phi_{2^k+l}(t) - \phi_{2^k+l}(s)| \qquad (2.10)\\
&\leq \tau_{2^k} \, 2^{-k(\alpha-\frac{1}{2})-\alpha+1} \, 2^{\frac{k}{2}} \, |t-s| \\
&\leq \tau_{2^k} \, 2^{k(1-\alpha)-\alpha+1-k_0(1-\alpha)} \, |t-s|^\alpha = \tau_{2^k} \, (2^{1-\alpha})^{(1+k-k_0)} \, |t-s|^\alpha,
\end{aligned}$$

while for $k \geq k_0$

$$\begin{aligned}
&\sum_{l=0}^{2^k-1} \frac{1}{c_{2^k+l}(\alpha)} |\eta_{2^k+l}| \, |\phi_{2^k+l}(t) - \phi_{2^k+l}(s)| \qquad (2.11)\\
&\leq \tau_{2^{k_0}} \, 2^{-k(\alpha-\frac{1}{2})-\alpha+1} \, 2^{-\frac{k}{2}} \\
&\leq \tau_{2^{k_0}} \, 2^{-k\alpha-\alpha+1+(k_0+1)\alpha} \, |t-s|^\alpha = \tau_{2^{k_0}} \, (2^\alpha)^{(k_0-k)} \, |t-s|^\alpha.
\end{aligned}$$

Hence (2.9), (2.10) and (2.11) imply

$$\frac{|f(t)-f(s)|}{|t-s|^\alpha} \leq [||\eta||_\infty \, |t-s|^{1-\alpha} + \sum_{k\leq k_0} \tau_{2^k} \, (2^{1-\alpha})^{(1+k-k_0)} + \frac{1}{2^\alpha - 1} \tau_{2^{k_0}}].$$

Now $k_0 \to \infty$ as $|t-s| \to 0$. This, however, entails that $f \in \mathcal{C}^0_\alpha([0,1])$. All arguments used in the proof of Theorem 2.1.2 to show the inequalities for the operator norms are valid here. Just note that the function $f(t) = t, 0 \leq t \leq 1$, is in $\mathcal{C}^0_\alpha([0,1])$. □

2.2 The Schauder representation of Brownian motion

We shall now present an approach of the study of one-dimensional Brownian motion which is close to Wiener's representation of Brownian motion by Fourier series with trigonometric functions as a basis. Our basis will be given by the Haar functions and their primitives. In fact, the trajectories of Brownian motion will be described just as in the preceding section continuous functions were isomorphically described by sequences. Given a Brownian motion X indexed by the unit interval, with the same notation as in the preceding section we write it sample by sample as a series with coefficients $\langle \chi_n, dX \rangle, n \in \mathbb{N}$. Due to the scaling properties and the structure of Haar functions, these random coefficients are i.i.d standard normal random variables. This, in turn, allows us to construct Brownian motion indexed by the unit interval by taking any sequence of i.i.d. standard normal variables $(Z_n)_{n\in\mathbb{N}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and defining the stochastic process

$$W_t = \sum_{n=0}^{\infty} Z_n \phi_n(t), \quad t \in [0,1]. \qquad (2.12)$$

To get information about the quality of convergence of this Fourier series, we need to control the size of the random sequence $(Z_n)_{n\in\mathbb{N}}$ in the following Lemma.

Lemma 2.2.1. *There exists a real valued random variable C such that for $n \geq 2$ we have*

$$|Z_n| \leq C\sqrt{\ln n}. \tag{2.13}$$

Proof. For $x \geq 1, n \geq 2$ we have

$$\mathbb{P}(|Z_n| \geq x) = \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du \leq \sqrt{\frac{2}{\pi}} \int_x^\infty u e^{-\frac{u^2}{2}} du = \sqrt{\frac{2}{\pi}} e^{-\frac{x^2}{2}}.$$

Hence for $\beta > 1$

$$\mathbb{P}(|Z_n| \geq \sqrt{2\beta \ln n}) \leq \sqrt{\frac{2}{\pi}}\, e^{-\beta \ln n} = \sqrt{\frac{2}{\pi}} n^{-\beta}.$$

Therefore, the lemma of Borel-Cantelli yields that $|Z_n| \leq \sqrt{2\beta \ln n}$ for almost all $n \in \mathbb{N}$ with probability 1. Hence the random variable $C = \sup_{n\geq 2} \frac{|Z_n|}{\sqrt{\ln n}}$ is almost surely finite, and yields the desired inequality. □

The preceding Lemma enables us to state that the convergence in (2.12) is absolute and therefore the process continuous. Its law has the characteristics of the law of a Brownian motion, as the following Theorem shows.

Theorem 2.2.2. *The series in (2.12) converges absolutely in the uniform norm to a continuous process W which is a Brownian motion on $[0,1]$.*

Proof. Let us first prove the absolute convergence of the series in the uniform norm. This will evidently imply that W is continuous. Let $k, p, q \in \mathbb{N}$ be such that $q \geq p \geq 2^k$. Then for $t \in [0,1]$ we have with the random variable C of the preceding Lemma

$$\begin{aligned} \sum_{n=p}^{q} |Z_n| \phi_n(t) &\leq C \sum_{n=p}^{q} \sqrt{\ln n}\ \phi_n(t) \\ &\leq C \sum_{j=k}^{\infty} \sum_{l=0}^{2^j-1} \sqrt{j+1}\ \sqrt{\ln 2}\ \phi_{2^j+l}(t) \\ &\leq C \sum_{j=k}^{\infty} \sqrt{j+1}\ 2^{-\frac{j}{2}-1}, \end{aligned}$$

which converges to 0 as q, p tend to ∞, independently of $t \in [0,1]$.

To prove that W is a Gaussian process with $\mathbb{E}(W_t) = 0$ and $\mathrm{cov}(W_t, W_s) = s \wedge t$, for $0 \leq s, t \leq 1$, we first note that the series also converges in square norm. In fact, we have for $t \in [0,1]$, $k, p, q \in \mathbb{N}$ such that $q \geq p \geq 2^k$ by the law properties of $Z_n, n \geq 0$,

$$\mathbb{E}([\sum_{n=p}^{q} Z_n \phi_n(t)]^2) = \sum_{n=p}^{q} \phi_n(t)^2 \leq \sum_{j=k}^{\infty} 2^{-j-2},$$

which converges to 0 as $p, q \to \infty$. Next, let $d \in \mathbb{N}$, $0 \leq t_1 < \cdots < t_d \leq 1$, and $\vartheta = (\vartheta_1, \cdots, \vartheta_d) \in \mathbb{R}^d$ be given. We compute the Fourier transform $\varphi(\vartheta)$ of the vector $(W_{t_1}, \cdots, W_{t_d})$ at ϑ. We have by dominated convergence and the law properties of

$Z_n, n \geq 0$, again

$$\begin{aligned}
\varphi(\vartheta) &= \mathbb{E}(\exp(i\sum_{j=1}^d \vartheta_j W_{t_j})) \\
&= \mathbb{E}(\exp(i\sum_{j=1}^d \vartheta_j \sum_{n=0}^\infty Z_n \phi_n(t_j))) \\
&= \prod_{n=0}^\infty \mathbb{E}(\exp(iZ_n \sum_{j=1}^d \vartheta_j \phi_n(t_j))) \\
&= \prod_{n=0}^\infty \exp(-\frac{1}{2}(\sum_{j=1}^d \vartheta_j \phi_n(t_j))^2) \\
&= \exp(-\frac{1}{2}\sum_{n=0}^\infty(\sum_{j=1}^d \vartheta_j \phi_n(t_j))^2) \\
&= \exp(-\frac{1}{2}\sum_{j,k=1}^d \vartheta_j \vartheta_k \ \sum_{n=0}^\infty \phi_n(t_j)\phi_n(t_k)).
\end{aligned}$$

Now observe that Parseval's equation implies for $1 \leq j, k \leq d$

$$t_j \wedge t_k = \langle 1_{[0,t_j]}, 1_{[0,t_k]} \rangle = \sum_{n=0}^\infty \langle 1_{[0,t_j]}, \chi_n \rangle \, \langle 1_{[0,t_k]}, \chi_n \rangle = \sum_{n=0}^\infty \phi_n(t_j)\phi_n(t_k).$$

Therefore we finally obtain

$$\varphi(\vartheta) = \exp(-\frac{1}{2}\sum_{j,k=1}^d \vartheta_j \vartheta_k \ t_j \wedge t_k).$$

But this means that $(W_{t_1}, \cdots, W_{t_d})$ is Gaussian with expectation vector 0 and covariance matrix C with entries $c_{jk} = t_j \wedge t_k, 1 \leq j, k \leq d$. It is easy to see that these properties imply that the process W possesses independent increments which are Gaussian with mean 0 and variance corresponding to the length of the increment intervals. This, however, characterizes a Brownian motion. □

The usual one-dimensional Brownian motion $W = (W_t)_{t \geq 0}$ can now be introduced by starting with a sequence of independent Brownian motions $(W^n)_{n \in \mathbb{N}}$ indexed by the unit interval and given by the construction discussed above on the basis of Schauder representations. Then we can evidently set

$$W_t = \sum_{k=1}^{[t]} W_1^k + W_{t-[t]}^{[t]+1}, \quad t \geq 0.$$

And finally, the d-dimensional Brownian motion $W = (W^1, \cdots, W^d)$ is just a d-tuple of independent one-dimensional Brownian motions indexed by $\mathbb{R}_+$.

We now use the Schauder representation of Brownian motion to show its Hölder continuity properties.

Theorem 2.2.3. *The Brownian motion $W = (W_t)_{0\le t\le 1}$ is Hölder continuous of order $\alpha < 1/2$. Its trajectories are a.s. nowhere Hölder continuous of order $\alpha > 1/2$. Moreover we have (Lévy's modulus of continuity)*

$$\mathbb{P}\Big(\sup_{0\le s<t\le 1} \frac{|W_t - W_s|}{h(|t-s|)} < \infty\Big) = 1, \tag{2.14}$$

where $h(u) = \sqrt{u\log(1/u)}, u > 0$. In particular, for $\alpha < \frac{1}{2}$, the trajectories of W are $\mathbb{P}$-a.s. contained in the space $\mathcal{C}^0_\alpha([0,1])$.

Proof. Let first $\alpha \in]0,1[$, $(c_n)_{n\ge 0}$ be a sequence of real numbers for which there exists $c \in \mathbb{R}$ such that for $n = 2^k + l, 0 \le l \le 2^k - 1$ we have

$$|c_n| \le c\sqrt{k}.$$

Let

$$f(t) = \sum_{n=0}^{\infty} c_n\,\phi_n(t), \quad t \in [0,1].$$

We shall prove that $\sup_{0\le s<t\le 1} \frac{|f(t)-f(s)|}{h(|t-s|)} < \infty$. Due to Lemma 2.2.1, this will imply the claimed formula. In fact, due to the continuity properties of ϕ_0, we may assume that $c_0 = 0$. Then for $0 \le s < t \le 1$

$$|f(t) - f(s)| \le \sum_{k=0}^{\infty}\sum_{l=0}^{2^k-1} |c_{2^k+l}|\,|\phi_{2^k+l}(t) - \phi_{2^k+l}(s)|. \tag{2.15}$$

Now choose $k_0 \ge 0$ such that

$$2^{-k_0-1} < |t-s| \le 2^{-k_0}.$$

Then for $0 \le k < k_0$

$$\begin{aligned} &\sum_{l=0}^{2^k-1} |c_{2^k+l}|\,|\phi_{2^k+l}(t) - \phi_{2^k+l}(s)| \\ &\le\; c\sqrt{k}\,2^{\frac{k}{2}}\,|t-s| \\ &\le\; c\sqrt{k}\,2^{\frac{k-k_0}{2}}\,|t-s|^{\frac{1}{2}} \\ &\le\; \frac{c}{\sqrt{\ln 2}}\sqrt{\frac{k}{k_0}}\,2^{\frac{k-k_0}{2}}\sqrt{|t-s|\,\ln\frac{1}{|t-s|}}, \end{aligned} \tag{2.16}$$

while for $k \ge k_0$

$$\begin{aligned} &\sum_{l=0}^{2^k-1} |c_{2^k+l}|\,|\phi_{2^k+l}(t) - \phi_{2^k+l}(s)| \\ &\le\; c\sqrt{k}\,2^{-\frac{k}{2}} \\ &\le\; \frac{c}{\sqrt{\ln 2}}\sqrt{\frac{k}{k_0}}\,2^{\frac{k_0+1-k}{2}}\sqrt{|t-s|\,\ln\frac{1}{|t-s|}}. \end{aligned} \tag{2.17}$$

It is easy to see, for instance by estimating $\int_a^\infty \sqrt{x}2^{-x}dx$, and $\int_1^a \sqrt{x}2^x dx$ for $a \geq 1$ using integration by parts that the sum in k of the two estimates can be taken and yields a finite upper bound which does not depend on k_0. Hence (2.15), (2.16) and (2.17) imply

$$\frac{|f(t)-f(s)|}{\sqrt{|t-s|\,\ln\frac{1}{|t-s|}}} \leq c',$$

for some constant c' independent of s and t. This implies the desired inequality, and all claims about Hölder continuity for $\alpha < \frac{1}{2}$.

Let us next fix $\alpha > \frac{1}{2}$. For $c > 0, \epsilon > 0$ let

$$G(\alpha, c, \epsilon) = \{\omega \in \Omega : \ \exists\ s \in [0,1]\ \forall\ t \in [0,1],\ |s-t| \leq \epsilon : \ |W_t(\omega) - W_s(\omega)| \leq c|s-t|^\alpha\}.$$

We will show that $\mathbb{P}(G(\alpha, c, \epsilon)) = 0$ for all $c, \epsilon > 0$, and thus that W is a. s. nowhere Hölder continuous of order α. To this end, for all $m, n \in \mathbb{N}, m \leq n$, and $0 \leq k < n$ let

$$X_{m,k} = \max\{|W_{\frac{j}{n}} - W_{\frac{j+1}{n}}| : k \leq j < m+k\}.$$

Let $\omega \in G(\alpha, c, \epsilon)$. Choose $n \in \mathbb{N}$ so that $\frac{m}{n} \leq \epsilon$. Let $s \in [0,1]$ be given such that for all $t \in [0,1]$ satisfying $|s-t| \leq \epsilon$ we have $|W_t(\omega) - W_s(\omega)| \leq c|s-t|^\alpha$. Choose $0 \leq k \leq n-m$ such that $\frac{k}{n} \leq s < \frac{k+m}{n}$. Then for $k \leq j < k+m$

$$\begin{aligned} |W_{\frac{j}{n}}(\omega) - W_{\frac{j+1}{n}}(\omega)| &\leq |W_{\frac{j}{n}}(\omega) - W_s(\omega)| + |W_s(\omega) - W_{\frac{j+1}{n}}(\omega)| \\ &\leq c|\frac{j}{n} - s|^\alpha + c|s - \frac{j+1}{n}|^\alpha \leq 2c(\frac{m}{n})^\alpha. \end{aligned}$$

This proves that

$$G(\alpha, c, \epsilon) \subset \{\min_{0 \leq k \leq n-m} X_{m,k} \leq 2c(\frac{m}{n})^\alpha\}. \tag{2.18}$$

Let us now estimate the probability of the latter set. Indeed, we have using independence and stationarity of the laws of the increments of W, and its scaling properties

$$\begin{aligned} \mathbb{P}(\min_{0 \leq k \leq n-m} X_{m,k} \leq 2c(\frac{m}{n})^\alpha) &\leq n\mathbb{P}(X_{m,1} \leq 2c(\frac{m}{n})^\alpha) \\ &\leq n\mathbb{P}(|W_{\frac{1}{n}}| \leq 2c(\frac{m}{n})^\alpha)^m \\ &= n\mathbb{P}(|W_1| \leq 2c\sqrt{n}(\frac{m}{n})^\alpha)^m \\ &\leq n[\sqrt{\frac{2}{\pi}}2c\sqrt{n}(\frac{m}{n})^\alpha]^m = n^{1+(\frac{1}{2}-\alpha)m}[\sqrt{\frac{2}{\pi}}2cm^\alpha]^m. \end{aligned}$$

Now choose m so that $1 + (\frac{1}{2} - \alpha)m < 0$. Then let $n \to \infty$ to obtain that

$$\mathbb{P}(G(\alpha, c, \epsilon)) = 0,$$

as desired. □

Chapter 3

The large deviation principle

In this course we shall mainly be concerned with the calculation of large deviation rates for diffusion processes X^ϵ derived from dynamical systems perturbed by additive Brownian noise of small intensity ϵ. The rates will be calculated in two steps: first we shall establish the large deviations principle for small Brownian motion $\sqrt{\epsilon}W$. In a second step diffusions will be considered as continuous maps of Brownian motion, and large deviations principles transferred via the contraction principle. In this Chapter, we shall prepare these steps by discussing the general framework of large deviations theory.

3.1 Concept and basic properties

To state the large deviation principle, and investigate its basic properties, let $(\mu_\epsilon)_{\epsilon>0}$ be a family of probability measures on a topological (Hausdorff) space $(\mathbf{X}, \mathbf{B})$ ($\mathbf{B}$ is the Borel σ-algebra). Think of μ_ϵ as the law of $\sqrt{\epsilon}W$ or X^ϵ, $\epsilon > 0$. And think of the topological space $\mathbf{X}$ as $\mathcal{C}([0,1])$ or a Hölder space $\mathcal{C}^0_\alpha([0,1])$, in which the functions vanish at 0. The principle concerns the limiting behavior of exponential rates of $(\mu_\epsilon)_{\epsilon>0}$ as $\epsilon \to 0$ in terms of a *rate function.*

Definition 3.1.1. *A* rate function *is a lower semicontinuous function $I : \mathbf{X} \to [0, \infty]$, i.e. for all $\alpha \in [0, \infty[$, the level sets*

$$\Psi_I(\alpha) = \{x \in \mathbf{X} : I(x) \le \alpha\}$$

are closed. I is called good rate function*, if all level sets are compact.*

Remark
If the topology of $\mathbf{X}$ has a countable basis, lower semicontinuity of I is equivalent to the property

$$\liminf_{n\to\infty} I(x_n) \ge I(x)$$

for all sequences $(x_n)_{n\in\mathbb{N}} \subset \mathbf{X}$ converging to $x \in \mathbf{X}$.

Definition 3.1.2. *Let I be a rate function. A family of probability measures $(\mu_\epsilon)_{\epsilon>0}$ on* $(\mathbf{X}, \mathbf{B})$ satisfies the large deviation principle (LDP) with rate function I *if for all* $\Gamma \in \mathbf{B}$

we have

$$-\inf_{x\in\Gamma^o} I(x) \le \liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(\Gamma) \le \limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(\Gamma) \le -\inf_{x\in\overline{\Gamma}} I(x).$$

Here Γ^o *resp.* $\overline{\Gamma}$ *denote the open kernel resp. the closed hull of* Γ.

The following equivalent characterization is evident, but often more practical to prove.

Remark 1

$(\mu_\epsilon)_{\epsilon>0}$ satisfies a LDP with rate function I iff the following conditions are satisfied.

(a) For every $\alpha < \infty$ and every $\Gamma \in \mathbf{B}$ such that $\inf_{x\in\overline{\Gamma}} I(y) \ge \alpha$ we have

$$\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(\Gamma) \le -\alpha.$$

(b) For $x \in \mathbf{X}$ with $I(x) < \infty$ and any $\Gamma \in \mathbf{B}$ with $x \in \Gamma^o$ we have

$$\liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(\Gamma) \ge -I(x).$$

Remark 2

Let $(\mu_\epsilon)_{\epsilon>0}$ be a family of probability measures, I a rate function. Then the LDP is equivalent to the following statements:

(a) for any closed set $F \subset \mathbf{X}$ we have

$$\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(F) \le -\inf_{x\in F} I(x),$$

(b) for any open set $G \subset \mathbf{X}$ we have

$$\liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(G) \ge -\inf_{x\in G} I(x).$$

Proof. 1. The LDP evidently implies (a) and (b).

2. Assume that (a) and (b) are satisfied, and let $\Gamma \in \mathbf{B}$. By (a) we have

$$\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(\Gamma) \le \limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(\overline{\Gamma}) \le -\inf_{x\in\overline{\Gamma}} I(x).$$

By (b) we have

$$\liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(\Gamma) \ge \liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(\Gamma^o) \ge -\inf_{x\in\Gamma^o} I(x).$$

Combining the two inequalities gives the defining property. □

Definition 3.1.3. *We say that* $(\mu_\epsilon)_{\epsilon>0}$ *satisfies a* weak LDP *with rate function* I, *if (a) for compact instead of closed sets and (b) of the preceding remark are satisfied.*

In practise, one often has the validity of the weak LDP. To conclude from this the validity of the LDP, some appropriate tightness condition is needed.

Definition 3.1.4. *A family $(\mu_\epsilon)_{\epsilon>0}$ is said to be* exponentially tight *if for every $\alpha < \infty$ there exists a compact set $K_\alpha \subset \mathbf{X}$ such that*

$$\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(K_\alpha^c) < -\alpha.$$

We show how exponential tightness can be used to deduce a LD principle.

Lemma 3.1.1. *Let $(\mu_\epsilon)_{\epsilon>0}$ be an exponentially tight family of probability measures. Then we have:*

*(a) The condition (*lower bound*)*

$$\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(K) \le -\inf_{x\in K} I(x), \quad K \subset \mathbf{X} \textit{ compact}$$

implies the lower bound for closed sets $F \subset \mathbf{X}$.

*(b) The condition (*upper bound*)*

$$\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(G) \ge -\inf_{x\in G} I(x), \quad G \subset \mathbf{X} \textit{ open}$$

implies that I is a good rate function.

Proof. 1. Let $F \subset \mathbf{X}$ be closed, and $\alpha < \infty$ such that $\inf_{x\in F} I(x) \ge \alpha$. Choose K_α according to the definition of exponential tightness. Then for any $\epsilon > 0$

$$\mu_\epsilon(F) \le \mu_\epsilon(F \cap K_\alpha) + \mu_\epsilon(K_\alpha^c).$$

Now, for $a, b : \mathbb{R}_+ \to \mathbb{R}_+$, due to $\ln(a(\epsilon)+b(\epsilon)) \le \ln(2a(\epsilon)) \vee \ln(2b(\epsilon)) = \ln(a(\epsilon)) \vee \ln(b(\epsilon)) + \ln 2$ we have $\lim_{\epsilon\to 0} \epsilon \ln(a(\epsilon)+b(\epsilon)) \le \lim_{\epsilon\to 0} \epsilon \ln(a(\epsilon)) \vee \ln(b(\epsilon))$ and therefore by hypothesis applied to $F \cap K_\alpha$

$$\begin{aligned}
\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(F) &\le \limsup_{\epsilon\to 0} \epsilon \log[\mu_\epsilon(F \cap K_\alpha) + \mu_\epsilon(K_\alpha^c)] \\
&\le \limsup_{\epsilon\to 0} \epsilon \log[\mu_\epsilon(F \cap K_\alpha)] \vee \limsup_{\epsilon\to 0} \epsilon \log[\mu_\epsilon(K_\alpha^c)] \\
&= [\limsup_{\epsilon\to 0} \epsilon \log[\mu_\epsilon(F \cap K_\alpha)] \\
&\le -\inf_{x\in F\cap K_\alpha} I(x) \le -\inf_{x\in F} I(x).
\end{aligned}$$

2. For $\alpha < \infty$ let K_α be chosen according to the definition of exponential tightness. We have to show that $\Psi_I(\alpha)$ is compact. Apply the lower bound to the open set K_α^c. Then we have

$$-\inf_{x\in K_\alpha^c} I(x) \le \liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(K_\alpha^c) < -\alpha,$$

i.e.

$$\inf_{x\in K_\alpha^c} I(x) > \alpha,$$

which means that $I(x) \le \alpha$ implies $x \in K_\alpha$. Hence $\Psi_I(\alpha) \subset K_\alpha$ is compact. □

3.2 Construction of LDP from exponential rates of elementary sets

Large deviations principles state exponential rates for all open and closed sets of a topological space. Suppose that originally the exponential rates are only known for some simple sets for instance belonging to a basis of the topology. We shall now give a sufficient criterion under which from those rates one can obtain an LDP. In fact, we start with discussing a weak LDP.

Theorem 3.2.1. *Let $\mathcal{G}_0$ be a collection of open sets in the topology of $(\mathbf{X}, \mathbf{B})$ such that for each open set G and each $y \in G$ there is $G_0 \in \mathcal{G}_0$ such that $y \in G_0 \subset G$, I a rate function, $(\mu_\epsilon)_{\epsilon>0}$ a family of probability measures. Assume that for every $G \in \mathcal{G}_0$ we have*

$$-\inf_{x\in G} I(x) = \lim_{\epsilon\to 0} \epsilon \ln \mu_\epsilon(G).$$

Then $(\mu_\epsilon)_{\epsilon>0}$ satisfies a weak LDP with rate function I.

Proof. Let us first establish the lower bound. In fact, let G be an open set. Choose $x \in G$, and a basis set G_0 such that $x \in G_0 \subset G$. Then evidently

$$\liminf_{\epsilon\to 0} \epsilon \ln \mu_\epsilon(G) \geq \liminf_{\epsilon\to 0} \epsilon \ln \mu_\epsilon(G_0) = -\inf_{y\in G_0} I(y) \geq -I(x).$$

Now the lower bound follows readily by taking the sup of $-I(x), x \in G$, on the right hand side, the left hand side not depending on x.

For the upper bound, fix a compact subset K of $\mathbf{X}$. For $\delta > 0$ denote

$$I^\delta(x) = (I(x) - \delta) \wedge \frac{1}{\delta}, \quad x \in \mathbf{X}.$$

For any $x \in K$, use the lower semicontinuity of I, more precisely that $\{y \in \mathbf{X} : I(y) > I^\delta(x)\}$ is open to choose a set $G_x \in \mathcal{G}_0$ such that

$$-I^\delta(x) \geq \limsup_{\epsilon\to 0} \epsilon \ln \mu_\epsilon(G_x).$$

Use compactness of K to extract from the open cover $K \subset \cup_{x\in K} G_x$ a finite subcover $K \subset \cup_{i=1}^n G_{x_i}$. Then with an argument as in the proof of Lemma 3.1.1 we obtain

$$\limsup_{\epsilon\to 0} \epsilon \ln \mu_\epsilon(K) \leq \max_{1\leq i\leq n} \limsup_{\epsilon\to 0} \epsilon \ln \mu_\epsilon(G_{x_i}) \leq -\min_{1\leq i\leq n} I^\delta(x_i) \leq -\inf_{x\in K} I^\delta(x).$$

Now let $\delta \to 0$, to complete the proof. □

Corollary 3.2.1. *Let $\mathcal{G}_0$ be a collection of open sets in the topology of $(\mathbf{X}, \mathbf{B})$ such that for each open set G and each $y \in G$ there is $G_0 \in \mathcal{G}_0$ such that $y \in G_0 \subset G$, I a rate function, $(\mu_\epsilon)_{\epsilon>0}$ an exponentially tight family of probability measures. Assume that for every $G \in \mathcal{G}_0$ we have*

$$-\inf_{x\in G} I(x) = \lim_{\epsilon\to 0} \epsilon \ln \mu_\epsilon(G).$$

Then $(\mu_\epsilon)_{\epsilon>0}$ satisfies an LDP with good rate function I.

Proof. Apply Lemma 3.1.1 to Theorem 3.2.1. □

3.3 Transformations of LDP

Assume we have established an LDP for a family of probability measures $(\mu_\epsilon)_{\epsilon>0}$ on a topological space $(\mathbf{X}, \mathbf{B})$, and $f : \mathbf{X} \to \mathbf{Y}$ is a continuous map to a topological (Hausdorff) space $(\mathbf{Y}, \mathbf{C})$. Then we will show that the family $(\nu_\epsilon = \mu_\epsilon \circ f^{-1} : \epsilon > 0)$ also satisfies an LDP.

Theorem 3.3.1 (contraction principle)**.** *Let $(\mathbf{X}, \mathbf{B}), (\mathbf{Y}, \mathbf{C})$ be topological spaces, and $f : \mathbf{X} \to \mathbf{Y}$ a continuous mapping. Let $I : \mathbf{X} \to [0, \infty]$ be a good rate function.*

(a) For $y \in \mathbf{Y}$ let

$$I'(y) = \inf\{I(x) : x \in \mathbf{X}, y = f(x)\}.$$

Then I' is a good rate function on $\mathbf{Y}$.

(b) Suppose $(\mu_\epsilon)_{\epsilon>0}$ satisfies an LDP with rate function I, and $\nu_\epsilon = \mu_\epsilon \circ f^{-1}, \epsilon > 0$. Then $(\nu_\epsilon)_{\epsilon>0}$ satisfies an LDP with respect to the rate function I'.

Proof. (a) **We have to show:** For $\alpha < \infty$ we have

$$\Psi_{I'}(\alpha) = \{y \in Y : I'(y) \leq \alpha\} \text{ is compact.}$$

In fact, we have by continuity of f and definition $f^{-1}(\{y\}) \cap \Psi_I(\alpha + \epsilon) \neq \emptyset$ for any $\epsilon > 0$ and thus by compactness

$$\Psi_{I'}(\alpha) \subset f(\Psi_I(\alpha)),$$

while the opposite inclusion is trivial. Since f is continuous and $\Psi_I(\alpha)$ compact, the compactness of $\Psi_{I'}(\alpha)$ follows.

(b) Let $H \subset \mathbf{Y}$ be open. Then $f^{-1}(H) \subset \mathbf{X}$ is open, and we have

$$\begin{aligned}\liminf_{\epsilon\to 0} \epsilon \log \nu_\epsilon(H) &= \liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(f^{-1}(H)) \\ &\geq -\inf_{x\in f^{-1}(H)} I(x) = -\inf_{y\in H} \inf_{x\in f^{-1}(y)} I(x) = -\inf_{y\in H} I'(y).\end{aligned}$$

An analogous statement holds for closed sets. □

Can we get a similar transfer of large deviation principles in the other direction? This question becomes relevant for example in the context of different topologies on the same topological space. Under which additional assumptions can the LDP be transferred from the coarser to a finer topology?

Theorem 3.3.2 (inverse contraction principle)**.** *Let $(\mathbf{X}, \mathbf{B}), (\mathbf{Y}, \mathbf{C})$ be topological spaces, and $g : \mathbf{Y} \to \mathbf{X}$ a continuous bijection. Let $(\nu_\epsilon)_{\epsilon>0}$ be a family of probability measures on $(\mathbf{Y}, \mathbf{C})$ which is exponentially tight. Let $\mu_\epsilon = \nu_\epsilon \circ g^{-1}, \epsilon > 0$. If $(\mu_\epsilon)_{\epsilon>0}$ satisfies a LDP with rate function I, then $(\nu_\epsilon)_{\epsilon>0}$ satisfies an LDP with rate function $I' = I \circ g$.*

Proof. 1. **We prove:** I' is a rate function.

In fact, let $\alpha < \infty$ be given. Then

$$\Psi_{I'}(\alpha) = \{y \in \mathbf{Y} : I'(y) \leq \alpha\} = \{y \in \mathbf{Y} : I(g((y)) \leq \alpha\} = g^{-1}(\Psi_I(\alpha)).$$

Since $\Psi_I(\alpha)$ is closed and g continuous, $\Psi_{I'}(\alpha)$ is closed. Hence I' is a rate function.

2. To prove a LDP for $(\nu_\epsilon)_{\epsilon>0}$, according to the exponential tightness of the family and Lemma 3.1.1 all we have to establish is the lower bound and the upper bound for compact sets.

(a) **We show:**

$$\limsup_{\epsilon\to 0} \epsilon \log \nu_\epsilon(K) \leq -\inf_{y\in K} I'(y), \quad K \subset \mathbf{Y} \text{ compact.}$$

Let $K \subset \mathbf{Y}$ be compact. Since g is a continuous bijection, we have

$$\begin{aligned} \limsup_{\epsilon\to 0} \epsilon \log \nu_\epsilon(K) &= \limsup_{\epsilon\to 0} \epsilon \log \nu_\epsilon \circ g^{-1}(g((K))) \\ &= \limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(g(K)) \leq -\inf_{x\in g(K)} I(x) \\ &= -\inf_{y\in K} I'(y). \end{aligned}$$

(b) **We show:** For $y \in \mathbf{Y}$, and G open such that $y \in G$ we have

$$\liminf_{\epsilon\to 0} \epsilon \log \nu_\epsilon(G) \geq -I'(y).$$

Once this is established, the lower bound follows readily by taking the sup of $-I'(y), y \in G$, on the right hand side, the left hand side not depending on y.

So let $y \in \mathbf{Y}$, and G open with $y \in G$. Let $\alpha = I'(y) = I(g(y))$. Choose $K_\alpha \subset \mathbf{Y}$ compact such that

$$\limsup_{\epsilon\to 0} \epsilon \nu_\epsilon(K_\alpha^c) < -\alpha.$$

Since by the continuous bijection property of g we have $g(K_\alpha)^c \subset \mathbf{X}$ open, we may infer from the LDP for $(\mu_\epsilon)_{\epsilon>0}$

$$-\inf_{x\in g(K_\alpha^c)} I(x) \leq \limsup_{\epsilon\to 0} \epsilon \nu_\epsilon(K_\alpha^c) < -\alpha.$$

Hence in particular $g(y) \in g(K_\alpha)$. Moreover, g is a homeomorphism on K_α. Now the identity is a continuous map from $g(K_\alpha)$ to $\mathbf{X}$ with the induced topology. Hence by Theorem 3.3.1 the family $(\mu_\epsilon)_{\epsilon>0}$, restricted to $g(K_\alpha)$ satisfies an LDP. And so

$$\liminf_{\epsilon\to 0} \epsilon \log \nu_\epsilon(G \cap K_\alpha) = \liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(g(G \cap K_\alpha) \geq -I'(y).$$

Hence

$$\begin{aligned} \liminf_{\epsilon\to 0} \epsilon \log \nu_\epsilon(G) &\geq \liminf_{\epsilon\to 0} \epsilon \log \nu_\epsilon(G \cap K_\alpha) \\ &= \geq -I'(y). \end{aligned}$$

This completes the proof. □

As a corollary, we note that exponential tightness helps to transfer an LDP from a coarser to a finer topology.

Corollary 3.3.1. *Let* $\mathbf{X}$ *be a set with two topologies* τ_1 *and* τ_2 *such that* $\tau_2 \subset \tau_1$*. Denote the Borel sets with respect to* τ_1 *by* $\mathbf{B}$*, those with respect to* τ_2 *by* $\mathbf{C}$*. Let* $(\mu_\epsilon)_{\epsilon>0}$ *be an exponentially tight family of probability measures on* $(\mathbf{X}, \mathbf{C})$*. If* $(\mu_\epsilon)_{\epsilon>0}$ *satisfies an LDP with respect to* τ_2*, it satisfies an LDP with respect to* τ_1*.*

Proof. Let $g : \mathbf{X} \to \mathbf{X}$ be the identity. g is a continuous bijection from the finer to the coarser topology. Now apply the Theorem. □

We next consider the situation in which two families of measures are given which describe the laws of families of processes. If the processes are asymptotically close in the sense of the following definition, we can prove that large deviations principles are transferred from one family to the other.

Definition 3.3.1. *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space,* $(\mathbf{Y}, \mathbf{B})$ *a metric measure space with metric* d*. For any* $\epsilon > 0$ *let* $Z_\epsilon, \tilde{Z}_\epsilon$*, be* $\mathbf{Y}$*-valued random variables with joint law* $P_\epsilon = P \circ (Z_\epsilon, \tilde{Z}_\epsilon)^{-1}$*, and laws* $\mu_\epsilon = P \circ Z_\epsilon^{-1}, \tilde{\mu}_\epsilon = P \circ \tilde{Z}_\epsilon^{-1}$*. We call* $(Z_\epsilon)_{\epsilon>0}$ *and* $(\tilde{Z}_\epsilon)_{\epsilon>0}$ exponentially equivalent *if for every* $\delta > 0$*, setting*

$$\Gamma_\delta = \{(y, z) \in \mathbf{Y} \times \mathbf{Y} : d(y, z) > \delta\},$$

and supposing that $d(Z_\epsilon, \tilde{Z}_\epsilon)$ *be measurable, we have*

$$\limsup_{\epsilon \to 0} \epsilon \log P_\epsilon(\Gamma_\delta) = \limsup_{\epsilon \to 0} \epsilon \log P(d(Z_\epsilon, \tilde{Z}_\epsilon) > \delta) = -\infty.$$

In this case the families $(\mu_\epsilon)_{\epsilon>0}$ *and* $(\tilde{\mu}_\epsilon)_{\epsilon>0}$ *are also called* exponentially equivalent*.*

Let us now show that exponentially equivalent families are not distinguished by the LDP.

Theorem 3.3.3. *Let* $(\mathbf{Y}, \mathbf{B})$ *be a metric measurable space with metric* d*,* $(\mu_\epsilon)_{\epsilon>0}$ *a family of probability measures which satisfies an LDP with good rate function* I *on* $(\mathbf{Y}, \mathbf{B})$*. Let* $(\tilde{\mu}_\epsilon)_{\epsilon>0}$ *be exponentially equivalent to* $(\mu_\epsilon)_{\epsilon>0}$*. Then also* $(\tilde{\mu}_\epsilon)_{\epsilon>0}$ *satisfies an LDP with good rate function* I*.*

Proof. The proof will be divided into 4 steps.

1. **We show:** for any $y \in \mathbf{Y}$

$$I(y) = -\inf_{\delta>0} \limsup_{\epsilon \to 0} \epsilon \log \tilde{\mu}_\epsilon(K_\delta(y)) = -\inf_{\delta>0} \liminf_{\epsilon \to 0} \epsilon \log \tilde{\mu}_\epsilon(K_\delta(y)).$$

Let $\delta > 0, y \in \mathbf{Y}$ be given. For any $\epsilon > 0$ we have

$$\mu_\epsilon(K_\delta(y)) \leq \tilde{\mu}_\epsilon(K_{2\delta}(y)) + P_\epsilon(\Gamma_\delta).$$

The lower bounds in the LDP for μ_ϵ further reveal

$$\begin{aligned} -\inf_{z \in K_\delta(y)} I(z) &\leq \liminf_{\epsilon \to 0} \epsilon \log \mu_\epsilon(K_\delta(y)) \\ &\leq \liminf_{\epsilon \to 0} \epsilon \log[\tilde{\mu}_\epsilon(K_{2\delta}(y)) + P_\epsilon(\Gamma_\delta)] \\ &\leq \max\{\liminf_{\epsilon \to 0} \epsilon \log \tilde{\mu}_\epsilon(K_{2\delta}(y)), \liminf_{\epsilon \to 0} \epsilon \log P_\epsilon(\Gamma_\delta)\}. \end{aligned}$$

Since $(\mu_\epsilon)_{\epsilon>0}$ is exponentially equivalent to $(\tilde{\mu}_\epsilon)_{\epsilon>0}$, we further obtain

$$-\inf_{z\in K_\delta(y)} I(z) \leq \liminf_{\epsilon\to 0} \epsilon \log \tilde{\mu}_\epsilon(K_{2\delta}(y)).$$

Repeating the arguments leading to this estimate with the roles of $\tilde{Z}_\epsilon$ and Z_ϵ reversed gives

$$-\inf_{z\in \overline{K_{2\delta}(y)}} I(z) \geq \limsup_{\epsilon\to 0} \epsilon \log \tilde{\mu}_\epsilon(K_\delta(y)).$$

Now, noting that $\overline{K_{2\delta}(y)} \subset K_{3\delta}(y)$, we may take the $\inf_{\delta>0}$ on both sides of the preceding inequalities, and recall the lower semicontinuity of I, to get for $\rho > 0$

$$I(y) \geq \sup_{\delta>0} \inf_{z\in K_\delta(y)} I(z) \geq I(y) - \rho,$$

and hence

$$-I(y) \leq \inf_{\delta>0} \liminf_{\epsilon\to 0} \epsilon \log \tilde{\mu}_\epsilon(K_\delta(y)) \leq \inf_{\delta>0} \limsup_{\epsilon\to 0} \epsilon \log \tilde{\mu}_\epsilon(K_\delta(y)) \leq -I(y).$$

This proves the claim.

2. **We show:** for $y \in \mathbf{Y}$, and $G \subset \mathbf{Y}$ open with $y \in G$ we have

$$-I(y) \leq \liminf_{\epsilon\to 0} \epsilon \log \tilde{\mu}_\epsilon(G),$$

remarking that this statement implies the LDP lower bound for $(\tilde{\mu}_\epsilon)_{\epsilon>0}$.

To see this, use part 1. and the fact that due to $y \in G$ there exists $\delta > 0$ such that $K_\delta(y) \subset G$ to deduce

$$-I(y) = \inf_{\delta>0} \limsup_{\epsilon\to 0} \epsilon \log \tilde{\mu}_\epsilon(K_\delta(y)) \leq \liminf_{\epsilon\to 0} \epsilon \log \tilde{\mu}_\epsilon(G).$$

This is the claimed estimate.

3. Fix $F \subset \mathbf{Y}$ closed, and $\delta > 0$. Let $F^\delta = \{z \in \mathbf{Y} : d(z, F) \leq \delta\}$. **We show:**

$$\limsup_{\epsilon\to 0} \epsilon \log \tilde{\mu}_\epsilon(F) \leq -\inf_{y\in F^\delta} I(y).$$

To see this, note that for $\epsilon > 0$ we have

$$\tilde{\mu}_\epsilon(F) \leq \mu_\epsilon(F^\delta) + P_\epsilon(\Gamma_\delta).$$

Now apply the upper bound of the LDP for μ_ϵ to get

$$\begin{aligned}
\limsup_{\epsilon\to 0} \epsilon \log \tilde{\mu}_\epsilon(F) &\leq \limsup_{\epsilon\to 0} \epsilon \log[\mu_\epsilon(F^\delta) + P_\epsilon(\Gamma_\delta)] \\
&\leq \max\{\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(F^\delta), \limsup_{\epsilon\to 0} \epsilon \log P_\epsilon(\Gamma_\delta)\} \\
&\leq \max\{-\inf_{y\in F^\delta} I(y), \limsup_{\epsilon\to 0} \epsilon \log P_\epsilon(\Gamma_\delta)\}.
\end{aligned}$$

Since $(\mu_\epsilon)_{\epsilon>0}$ is exponentially equivalent to $(\tilde{\mu}_\epsilon)_{\epsilon>0}$, we obtain

$$\limsup_{\epsilon\to 0} \epsilon \log \tilde{\mu}_\epsilon(F) \leq -\inf_{y\in F^\delta} I(y).$$

4. **We prove:** for $F \subset \mathbf{Y}$ closed

$$\inf_{y\in F} I(y) = \lim_{\delta\to 0} \inf_{y\in F^\delta} I(y).$$

Note that, together with 3., this implies the upper bound for $F \subset \mathbf{Y}$ closed.

Let $\eta > 0$. **We need to show:**

$$\gamma = \lim_{\delta\to 0} \inf_{y\in F^\delta} I(y) \geq \inf_{y\in F} I(y) - \eta.$$

Assume, avoiding a trivial case, that $\gamma < \infty$. Let $\alpha = \eta + \gamma$. Then for any $\delta > 0$ we have $\inf_{y\in F^\delta} I(y) \leq \alpha$, hence by definition

$$F^\delta \cap \Psi_I(\alpha) \neq \emptyset.$$

These sets being compact since I is a good rate function, we obtain

$$F \cap \Psi_I(\alpha) = \cap_{\delta>0}[F^\delta \cap \Psi_I(\alpha)] \neq \emptyset.$$

This translates into

$$\inf_{y\in F} I(y) \leq \alpha,$$

and hence into the claimed inequality. $\square$

Chapter 4

Large deviations for Brownian motion

In this Chapter, we shall establish a large deviation principle for d-dimensional Brownian motion. This will serve in the subsequent section to derive a large deviation principle for diffusions driven by additive noise, via the contraction principle. The large deviation principle for Brownian motion is usually referred to as Schilder's theorem. Our method to prove this Theorem takes its motivation from the Fourier series representation of Brownian motion discussed in Chapter 2. In fact, Ciesielski's isomorphism underlying this representation will enable us to reduce the argument for Schilder's Theorem to the large deviation principle for one-dimensional Gaussian random variables.

4.1 Large deviations for one-dimensional Gaussian random variables

The large deviation rate for a one-dimensional Gaussian unit random variable can be directly calculated. Consider a random variable Z with standard normal law, and let μ_ϵ be the law of $\sqrt{\epsilon}Z$. Then the following statement holds.

Theorem 4.1.1. *Let*

$$I(x) = \frac{x^2}{2}, \quad x \in \mathbb{R}.$$

Then for any open set $G \subset \mathbb{R}$ and any closed set $F \subset \mathbb{R}$ we have

$$-\inf_{x \in G} I(x) \leq \liminf_{\epsilon \to 0} \epsilon \log \mu_\epsilon(G),$$

$$-\inf_{x \in F} I(x) \geq \limsup_{\epsilon \to 0} \epsilon \log \mu_\epsilon(F).$$

Proof. We argue for a closed set $F \subset \mathbb{R}$. Let $a = \inf\{|x| : x \in F\}$. Note that the case $a = 0$ is trivial. We may therefore assume that $a > 0$. By symmetry we may further assume that there exists $b \geq a$ such that $F \subset]-\infty, -b] \cup [a, \infty[$. Hence for $\epsilon > 0$

$$\mu_\epsilon(F) \leq \mu_\epsilon([a, \infty[) + \mu_\epsilon(]-\infty, -b]) \leq \frac{2}{\sqrt{2\pi}} \int_{\frac{a}{\sqrt{\epsilon}}}^{\infty} \exp(-\frac{x^2}{2}) dx.$$

For $u > 1$ we have

$$\int_u^\infty \exp(-\frac{x^2}{2})dx \le \int_u^\infty x\exp(-\frac{x^2}{2})dx = \exp(-\frac{1}{2}u^2).$$

Hence for $\epsilon < a^2$

$$\epsilon \ln \mu_\epsilon(F) \le \epsilon[\ln(\frac{2}{\sqrt{2\pi}}) - \frac{a^2}{2\epsilon}],$$

and therefore

$$\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(F) \le -\frac{a^2}{2} = -\inf_{x\in F} I(x).$$

For open sets we need a different inequality. In fact, integration by parts gives for $u > 1$

$$\int_u^\infty \exp(-\frac{x^2}{2})dx = \frac{1}{u}\exp(-\frac{1}{2}u^2) - \int_u^\infty \frac{1}{x^2}\exp(-\frac{x^2}{2})dx,$$

hence

$$\frac{1}{u}\exp(-\frac{1}{2}u^2) \le (1+\frac{1}{u^2})\int_u^\infty \exp(-\frac{x^2}{2})dx$$

and

$$\frac{u}{1+u^2}\exp(-\frac{1}{2}u^2) \le \int_u^\infty \exp(-\frac{x^2}{2})dx.$$

Now let $G \subset \mathbb{R}$ be open, $y \in G$. By symmetry, we may assume that $y > 0$. Let, moreover, $a, b > 0$ such that $y \in]a, b[\subset G$. Then, for ϵ small enough we have

$$\begin{aligned}\mu_\epsilon(G) &\ge \mu_\epsilon(]a,\infty[) - \mu_\epsilon([b,\infty[) \ge \frac{1}{\sqrt{2\pi}}[\frac{\frac{a}{\sqrt{\epsilon}}}{1+\frac{a^2}{\epsilon}}\exp(-\frac{a^2}{2\epsilon}) - \exp(-\frac{b^2}{2\epsilon})] \\ &\ge \frac{1}{\sqrt{2\pi}}\frac{\frac{a}{2\sqrt{\epsilon}}}{1+\frac{a^2}{\epsilon}}\exp(-\frac{a^2}{2\epsilon}).\end{aligned}$$

Therefore

$$\liminf_{\epsilon\to 0} \epsilon \ln \mu_\epsilon(G) \ge -\frac{a^2}{2} \ge -\frac{y^2}{2} = -I(y).$$

This implies the lower bound.

□

4.2 Large deviations for one-dimensional Brownian motion in Hölder space

In this section we use the Fourier series decomposition of one-dimensional Brownian motion in Hölder space $\mathcal{C}^0_\alpha([0,1])$ of order $\alpha < \frac{1}{2}$, and Ciesielski's isomorphism mapping this space to a sequence space, to derive a large deviations principle for Brownian motion with respect to the topology of Hölder space. This remarkable approach was presented in Baldi and

Roynette [3]. Let W be a one-dimensional Brownian motion indexed by $[0,1]$, described by

$$W = \sum_{n=0}^{\infty} Z_n \, \phi_n,$$

with a sequence $(Z_n)_{n\geq 0}$ of i.i.d standard normal variables, and the Schauder functions $(\phi_n)_{n\geq 0}$, as described in section 2. Recall the Haar functions $(\chi_n)_{n\geq 0}$ and the sequences $(c_n(\alpha))_{n\geq 0}$ appearing in Ciesielski's isomorphism in Theorem 2.1.2 for $0<\alpha<1$, given by

$$c_n(\alpha) = 2^{k(\alpha-\frac{1}{2})+\alpha-1}, \quad c_0(\alpha) = 1, \tag{4.1}$$

if $n = 2^k + l$ for $0 \leq l < 2^k$. We investigate the asymptotic behavior of the family of probability measures $(\mu_\epsilon)_{\epsilon>0}$, where μ_ϵ is the law of $\sqrt{\epsilon}W$, $\epsilon > 0$. We remark that according to Theorem 2.2.3 for any $\epsilon > 0, 0 < \alpha < \frac{1}{2}$ we have

$$\mu_\epsilon(\mathcal{C}^0_\alpha([0,1])) = 1. \tag{4.2}$$

Note also that the separability of $\mathcal{C}_0$ is translated into separability of $\mathcal{C}^0_\alpha([0,1])$ by Ciesielski's isomorphism of Theorem 2.1.3, while Theorem 2.1.2 yields that $\mathcal{C}_\alpha([0,1])$ is not separable.

The large deviation rates for Brownian motion will crucially depend on the following function space, the *Cameron-Martin space* of absolutely continuous functions.

Definition 4.2.1. *Let*

$$\begin{aligned} \mathcal{H}_1 &= \Big\{ f : [0,1] \to \mathbb{R},\ f(0) = 0,\ f \text{ abs. cont. with density } \dot{f} \in L^2([0,1]) \Big\} \\ &= \Big\{ \int_0^t \dot{f}(s)ds,\ \dot{f} \in L^2([0,1]) \Big\}. \end{aligned} \tag{4.3}$$

By means of (4.3) we can define the rate function for Brownian motion.

Definition 4.2.2. *Let*

$$I(f) = \begin{cases} \frac{1}{2}\int_0^1 (\dot{f})^2(u)du, & \text{if } f \in \mathcal{H}_1, \\ \infty, & \text{otherwise.} \end{cases} \tag{4.4}$$

The following Theorem can be considered a version of Schilder's Theorem with respect to a finer topology, proved for sets of a basis of the Hölder space topology first. For $\delta > 0, \xi \in \mathcal{C}$ denote $B^\infty_\delta(\xi)$ the ball of radius δ in the topology of $\mathcal{C}$. We consider the basic collection of sets $T^{-1}_{\alpha,0}(B^\infty_\delta(T_{\alpha,0}(\psi)))$ for $\psi \in C^\alpha_0([0,1])$.

Theorem 4.2.1. *Let $0 < \alpha < 1/2$, $\delta > 0$ and $\psi \in (\mathcal{C}^0_\alpha([0,1]), \|\cdot\|_\alpha)$. Then with the rate function I defined by (4.4)*

$$\lim_{\epsilon\to 0} \epsilon \ln \mu_\epsilon(T^{-1}_{\alpha,0}(B^\infty_\delta(T_{\alpha,0}(\psi)))) = -\inf_{f \in T^{-1}_{\alpha,0}(B^\infty_\delta(T_{\alpha,0}((\psi)))} I(f), \tag{4.5}$$

$$\lim_{\epsilon\to 0} \epsilon \ln \mu_\epsilon(\overline{T^{-1}_{\alpha,0}(B^\infty_\delta(T_{\alpha,0}((\psi))))}) = -\inf_{f \in \overline{T^{-1}_{\alpha,0}(B^\infty_\delta(T_{\alpha,0}(\psi)))}} I(f). \tag{4.6}$$

Proof. We give the arguments for (4.5). The proof of (4.6) is almost identical.

1. We use the Schauder representation of Brownian motion W and the function ψ given by

$$W = \sum_{n\geq 0} Z_n \phi_n \quad \text{and} \quad \psi = \sum_{n\geq 0} \frac{\xi_n}{c_n(\alpha)}\, \phi_n. \tag{4.7}$$

We recall the inverse of Ciesielski's isomorphism

$$T_{\alpha,0}^{-1} : \mathcal{C}_0 \to \mathcal{C}_\alpha^0([0,1]), \quad (\eta_n)_{n\geq 0} \mapsto \sum_{n=0}^{\infty} \frac{1}{c_n(\alpha)} \eta_n\ \phi_n,$$

and remark that the sequence $(\xi_n)_{n\geq 0}$ in the representation of ψ just satisfies $T_{\alpha,0}(\psi) = (\xi_n)_{n\geq 0}$, while $T_{\alpha,0}(\sqrt{\epsilon} W) = (\sqrt{\epsilon} c_n(\alpha) Z_n)_{n\geq 0}$. We therefore have

$$\sqrt{\varepsilon} W \in T_{\alpha,0}^{-1}(B_\delta^\infty(\xi) \Longleftrightarrow \sup_{n\geq 0} |\sqrt{\varepsilon} c_n(\alpha) Z_n - \xi_n| < \delta.$$

Hence

$$(\sqrt{\varepsilon} W)^{-1}[T_{\alpha,0}^{-1}(B_\delta^\infty(\xi)] = \bigcap_{n\geq 0} \Big\{ \sqrt{\varepsilon} c_n(\alpha) Z_n \in]\xi_n - \delta, \xi_n + \delta[\Big\}.$$

Since $(Z_n)_{n\geq 0}$ is a family of independent random variables, we obtain for $\epsilon > 0$

$$\mu_\epsilon(T_{\alpha,0}^{-1}(B_\delta^\infty(\xi))) = \prod_{n\geq 0} \mathbb{P}\Big(\sqrt{\varepsilon} c_n(\alpha) Z_n \in]\xi_n - \delta, \xi_n + \delta[\Big) = \prod_{n\geq 0} P_n(\varepsilon). \tag{4.8}$$

We split the sequence of probabilities $(P_n(\varepsilon))_{n\geq 0}$ into four different parts to be treated separately:

$$\begin{aligned}
\Lambda_1 &= \Big\{ n \geq 0 :\ 0 \notin [\xi_n - \delta, \xi_n + \delta] \Big\}, \\
\Lambda_2 &= \Big\{ n \geq 0 :\ \xi_n = \pm\delta \Big\}, \\
\Lambda_3 &= \Big\{ n \geq 0 : [\xi_n - \delta, \xi_n + \delta] \supset \Big[-\tfrac{\delta}{2}, \tfrac{\delta}{2} \Big] \Big\}, \\
\Lambda_4 &= (\Lambda_3)^c \setminus (\Lambda_1 \cup \Lambda_2).
\end{aligned}$$

Let us recall that $(\xi_n)_{n\geq 0} \in \mathcal{C}_0$, so Λ_3 contains almost all $n \geq 0$, and hence $\Lambda_1 \cup \Lambda_2 \cup \Lambda_4 = (\Lambda_3)^c$ is finite.

2. Let us first discuss the contribution of Λ_3. Since $(Z_n)_{n\geq 0}$ are standard normal variables, we have

$$\prod_{n\in\Lambda_3} P_n(\varepsilon) \geq \prod_{n\in\Lambda_3} \mathbb{P}\Big(Z_n \in \Big[-\frac{\delta}{2c_n(\alpha)\sqrt{\varepsilon}}, \frac{\delta}{2c_n(\alpha)\sqrt{\varepsilon}} \Big] \Big) = \prod_{n\in\Lambda_3} \Big(1 - \sqrt{\frac{2}{\pi}} \int_{\delta/(2c_n(\alpha)\sqrt{\varepsilon})}^\infty e^{-u^2/2} du \Big).$$

Now according to (4.1) and our choice of α, $c_n(\alpha) \leq 1, \lim_{n\to\infty} c_n(\alpha) = 0$. Therefore, for $\epsilon > 0$ such that $\epsilon < \delta^2$ and all $n \geq 0$ we may estimate (see proof of Theorem 4.1.1)

$$\int_{\delta/(2c_n(\alpha)\sqrt{\varepsilon})}^\infty \exp(-\frac{x^2}{2}) dx \leq \exp(-\frac{\delta^2}{8c_n(\alpha)^2 \epsilon}).$$

In order to prove that $\prod_{n\in\Lambda_3} P_n(\varepsilon)$ converges to 1 as $\epsilon \to 0$, by the elementary inequality $\ln(\frac{1}{1-x}) = \ln(1+\frac{x}{1-x}) \le \frac{x}{1-x}$ for $x \in]0,1[$ it suffices to prove that $\sum_{n\ge 0}\exp(-\frac{\delta^2}{8c_n(\alpha)^2\varepsilon})$ converges to 0 as $\epsilon \to 0$. This is in fact the case due to (4.1). We deduce

$$\lim_{\varepsilon\to 0}\prod_{n\in\Lambda_3} P_n(\varepsilon) = 1. \tag{4.9}$$

3. Next, we estimate the contribution of Λ_4. Indeed, $|\Lambda_4| < \infty$ and by definition $[\xi_n - \delta, \xi_n + \delta]$ contains a small neighborhood of the origin for any $n \in \Lambda_4$. We obtain

$$\lim_{\varepsilon\to 0}\prod_{n\in\Lambda_4} P_n(\varepsilon) = 1. \tag{4.10}$$

4. Since $|\Lambda_2| < \infty$, its definition immediately gives

$$\lim_{\varepsilon\to 0}\prod_{n\in\Lambda_2} P_n(\varepsilon) = 2^{-|\Lambda_2|}. \tag{4.11}$$

5. Let us finally estimate the contribution of Λ_1. We define

$$\overline{\xi}_n = \begin{cases} \xi_n - \delta, & \text{if } \xi_n > \delta, \\ -(\xi_n + \delta), & \text{if } \xi_n < -\delta. \end{cases}$$

Since for $n \in \Lambda_1$ Z_n has a standard normal law, Theorem 4.1.1 implies

$$\lim_{\epsilon\to 0}\epsilon \ln P_n(\varepsilon) = -\frac{\overline{\xi}_n^2}{2c_n(\alpha)^2}.$$

Since $|\Lambda_1| < \infty$, we therefore have

$$\lim_{\epsilon\to 0}\epsilon \ln \prod_{n\in\Lambda_1} P_n(\varepsilon) = -\sum_{n\in\Lambda_1}\frac{\overline{\xi}_n^2}{2c_n(\alpha)^2}. \tag{4.12}$$

6. Using (4.9), (4.10), (4.11) and (4.12), we can deduce (4.5) if we are able to compare

$$\sum_{n\in\Lambda_1}\frac{\overline{\xi}_n^2}{2c_n(\alpha)^2} \quad \text{with} \quad \inf_{f\in T_{\alpha,0}^{-1}(B_\delta^\infty(\xi))} I(f).$$

By Theorem 2.1.3 any function $f \in \mathcal{C}_\alpha^0([0,1]) \cap \mathcal{H}_1$ has the Schauder representation

$$f = \sum_{n\ge 0}\frac{\eta_n}{c_n(\alpha)}\,\phi_n, \quad \text{with } (\eta_n)_{n\ge 0} \in \mathcal{C}_0.$$

The derivative satisfies $\dot f = \sum_{n\ge 0}\frac{\eta_n}{c_n(\alpha)}\,\chi_n$, and since $(\chi_n)_{n\ge 1}$ is an orthonormal system in $L^2([0,1])$, we obtain

$$\frac{1}{2}\int_0^1 \dot f(s)^2\,ds = \sum_{n\ge 0}\frac{\eta_n^2}{2c_n(\alpha)^2}.$$

So the statement of the Theorem is an immediate consequence of the equality

$$\inf_{f\in T_{\alpha,0}^{-1}(B_\delta^\infty(\xi))\cap\mathcal{H}_1} \frac{1}{2}\int_0^1 \dot{f}(s)^2\,ds = \inf\Big\{\sum_{n\geq 0}\frac{\eta_n^2}{2c_n(\alpha)^2},\ \text{with } \eta_n\in]\xi_n-\delta,\xi_n+\delta[\Big\} = \sum_{n\in\Lambda_1}\frac{\bar{\xi}_n^2}{2c_n(\alpha)^2}.$$

□

Theorem 4.2.1 will allow us to derive an LDP for Brownian motion once we have established exponential tightness of the family $(\mu_\epsilon)_{\epsilon>0}$.

Theorem 4.2.2. *Let $0<\alpha<\frac{1}{2}$. Then $(\mu_\varepsilon)_{\epsilon>0}$ is exponentially tight on the topological space $(\mathcal{C}_\alpha^0([0,1]),\|\cdot\|_\alpha)$. More precisely, for $\delta>0$ and $0<\beta<\alpha<\frac{1}{2}$, $\overline{T_{\alpha,0}^{-1}(B_\delta^\alpha(0))}$ is compact in $\mathcal{C}_\beta^0([0,1])$, and we have*

$$\limsup_{\epsilon\to 0}\epsilon\ln\mu_\varepsilon(\overline{T_{\alpha,0}^{-1}(B_\delta^\alpha(0))}^c)\leq -\frac{\delta^2}{2}. \tag{4.13}$$

Proof. Recall that

$$(\overline{B_\delta^\infty(0)}) = \prod_{n=0}^\infty[-\delta,\delta].$$

Moreover, for $0<\beta<\alpha<\frac{1}{2}$ we have

$$T_{\beta,0}(T_{\alpha,0}^{-1}(\overline{B_\delta^\infty(0)})) = \prod_{n=0}^\infty[-\frac{c_n(\beta)}{c_n(\alpha)}\delta,\frac{c_n(\beta)}{c_n(\alpha)}\delta].$$

The space $\prod_{n=0}^\infty[-\frac{c_n(\beta)}{c_n(\alpha)}\delta,\frac{c_n(\beta)}{c_n(\alpha)}\delta]$ can easily be shown to be compact, for instance by establishing completeness and total boundedness, both of which are obviously satisfied. Since $T_{\beta,0}$ is an isomorphism, we obtain that $\overline{T_{\alpha,0}^{-1}(B_\delta^\infty(0))}$ is compact in $C_\beta^0([0,1])$.

Using the decomposition (4.7), we define for $\delta,\epsilon>0, n\geq 0$

$$A_n = \Big\{Z_n\notin\Big[-\frac{\delta}{\sqrt{\varepsilon}c_n(\alpha)},\frac{\delta}{\sqrt{\varepsilon}c_n(\alpha)}\Big]\Big\}.$$

Ciesielski's isomorphism of Theorem 2.1.3 allows to write

$$\mu_\varepsilon(\overline{B_\delta^\alpha(0)}^c) = \mathbb{P}(\exists n\geq 0 \text{ s.t. } \sqrt{\varepsilon}c_n(\alpha)Z_n\notin[-\delta,\delta]) = \mathbb{P}\Big(\bigcup_{n\geq 0}A_n\Big)\leq\sum_{n\geq 0}\mathbb{P}(A_n).$$

Now for $n\geq 0$ by Theorem 4.1.1

$$\limsup_{\epsilon\to 0}\epsilon\ln\mathbb{P}(A_n) = -\frac{\delta^2}{2c_n(\alpha)^2}.$$

Observe that $c_n(\alpha)$ strictly decreases to 0, starting at $c_0(\alpha)=1$. Hence by monotone convergence using the ideas of the proof of Lemma 3.1.1

$$\limsup_{\epsilon\to 0}\epsilon\ln\sum_{n=0}^\infty\mathbb{P}(A_n) = \limsup_{\epsilon\to 0}\epsilon\sup_{n\geq 0}\ln\mathbb{P}(A_n) = -\frac{\delta^2}{2}.$$

This implies (4.13) and the proof is complete. □

We are ready to state the main result of this section, which is a version of Schilder's Theorem with respect to a finer topology.

Theorem 4.2.3. (Baldi-Roynette) *Let $0 < \alpha < \frac{1}{2}$. For $\epsilon > 0$ let μ_ε be the law of $\sqrt{\epsilon}W$ on the topological space $(\mathcal{C}^0_\alpha([0,1]), ||\cdot||_\alpha)$. Then $(\mu_\varepsilon)_{\epsilon>0}$ satisfies a large deviations principle with the following good rate function*

$$I(f) = \begin{cases} \frac{1}{2}\int_0^1 (\dot{f})^2(u)du, & \text{if } f \in \mathcal{H}_1, \\ \infty & \text{otherwise.} \end{cases} \tag{4.14}$$

Proof. Combine Theorem 4.2.1 and Theorem 4.2.2 in Corollary 3.2.1. □

4.3 Large deviations for Brownian motion with respect to uniform norm

To obtain the classical result of Schilder's Theorem from the large deviation principle with respect to the finer topologies in Theorem 4.2.3, we finally have to apply the contraction principle in the form of Theorem 3.3.1. It is then straightforward to extend the results to multi-dimensional Brownian motions indexed by $\mathbb{R}_+$.

Theorem 4.3.1. (Schilder) *For $\epsilon > 0$ let μ_ε be the law of $\sqrt{\epsilon}W$ on the topological space $(\mathcal{C}([0,1]), ||\cdot||)$. Then $(\mu_\varepsilon)_{\epsilon>0}$ satisfies a large deviations principle with the following good rate function*

$$I(f) = \begin{cases} \frac{1}{2}\int_0^1 (\dot{f})^2(u)du, & \text{if } f \in \mathcal{H}_1, \\ \infty & \text{otherwise.} \end{cases} \tag{4.15}$$

Proof. According to Theorem 4.2.3, $(\mu_\varepsilon)_{\epsilon>0}$ satisfies an LDP on the space $(\mathcal{C}^0_\alpha([0,1]), \|\cdot\|_\alpha)$ with rate function I. Since the Hölder topology is finer than the uniform topology, Theorem 3.3.1 applied to the identity map implies that $(\mu_\varepsilon)_{\epsilon>0}$ satisfies an LDP on the space $(\mathcal{C}^0_\alpha([0,1]), \|\cdot\|_\infty)$ with the same rate function I. Finally we observe that the LDP is preserved under the identity map from $(\mathcal{C}^0_\alpha([0,1]), \|\cdot\|_\infty)$ to $(\mathcal{C}([0,1]), \|\cdot\|_\infty)$, again as a consequence of the contraction principle of Theorem 3.3.1. □

It remains to extend Schilder's Theorem to multi-dimensional Brownian motions.

Let therefore $W = (W^1, \cdots, W^d)$ be a d-dimensional Brownian motion indexed by $[0,1]$. The Cameron-Martin space to be used in this context has to be d-dimensional as well, but still be denoted by the same symbol. Write $|\cdot|$ for the Euclidean norm in $\mathbb{R}^d$.

Definition 4.3.1. *Let*

$$\begin{aligned} \mathcal{H}_1 &= \Big\{ f : [0,1] \to \mathbb{R}^d,\ \ f \text{ abs. cont. with density } \dot{f}, |\dot{f}| \in L^2([0,1]) \Big\} \\ &= \Big\{ \int_0^t \dot{f}(s)ds,\ \ |\dot{f}| \in L^2([0,1]) \Big\}. \end{aligned} \tag{4.16}$$

The d-dimensional version of Schilder's Theorem follows easily from the independence of the components of W. We continue denoting by the same symbol $\mathcal{C}([0,T])$ the space of continuous functions on $[0,T]$ with values in $\mathbb{R}^d$, endowed with the uniform norm $||\cdot||$.

Theorem 4.3.2. (Schilder) *For $\epsilon > 0$ let μ_ε be the law of $\sqrt{\epsilon}W$ on the topological space $(\mathcal{C}([0,1]), ||\cdot||)$. Then $(\mu_\varepsilon)_{\epsilon>0}$ satisfies a large deviations principle with the following good rate function*

$$I(f) = \begin{cases} \frac{1}{2}\int_0^1 |\dot{f}|^2(u)du, & \text{if } f \in \mathcal{H}_1, \\ \infty & \text{otherwise.} \end{cases} \tag{4.17}$$

Proof. For products of open sets in $\mathcal{C}([0,1])$ the result follows directly from Theorem 4.3.1 and the definition of I. This system of open sets forming a basis of the product topology, we may then argue using Corollary 3.2.1 and Theorem 4.2.2, which can be slightly extended to products of compact sets. □

Chapter 5

The Freidlin-Wentzell theory

In this Chapter, we shall extend the LDP for Brownian motion to diffusion processes that are obtained as the strong solutions of stochastic differential equations driven by additive Brownian motion. The basic idea for achieving this is simple. We map trajectories of the Brownian motion to trajectories of the solution of a given stochastic differential equation. Due to the additivity of the noise, this mapping is continuous. Therefore, a strict contraction principle in the sense of Chapter 3 is applicable to transfer the LDP.

5.1 The original theory

Denote $\mathcal{C}_0([0,1])$ the set of all functions $f \in \mathcal{C}([0,1])$ which satisfy $f(0) = 0$ (endowed with the uniform metric). Let $b : \mathbf{R}^d \to \mathbf{R}^d$ be a uniformly Lipschitz continuous function, i.e. there exists a constant B such that

$$|b(x) - b(y)| \leq B|x - y|, \quad x, y \in \mathbf{R}^d.$$

For $\epsilon > 0$ let $X^{\epsilon,x}$ denote the unique strong solution process of the SDE

$$X_t^{\epsilon,x} = x + \int_0^t b(X_s^\epsilon)ds + \sqrt{\epsilon}W_t, \quad 0 \leq t \leq 1. \tag{5.1}$$

We argue that the trajectories of $X^{\epsilon,x}$ are continuous images of the trajectories of $W_\epsilon = \sqrt{\epsilon}W$.

Lemma 5.1.1. *Let $g \in \mathcal{C}_0([0,1]), x \in \mathbf{R}^d$. Then there exists a unique $f \in \mathcal{C}([0,1])$ which satisfies*

$$f(t) = x + \int_0^t b(f(s))ds + g(t), \quad 0 \leq t \leq 1. \tag{5.2}$$

The mapping

$$F : \mathcal{C}_0([0,1]) \to \mathcal{C}([0,1]), g \mapsto f$$

is continuous and one-to-one.

Proof. 1. In order to prove the first claim about the unique solution, **it suffices to show:** For $x \in \mathbf{R}, g \in \mathcal{C}_0([0,1])$, and $T \in [0,1]$ such that $BT < 1$ the equation (5.2) possesses a unique solution on the interval $[0,T]$. Once this is proved, we can repeat the procedure a finite number of times on finitely many adjacent intervals of length bounded by $\delta = BT$, with recursively chosen initial conditions. To see this, define

$$\Gamma : \mathcal{C}([0,T]) \to \mathcal{C}([0,T]), f \mapsto (t \mapsto x + \int_0^t b(f(s))ds + g(t)).$$

Using the global Lipschitz condition for b, we see that F is a contraction. Indeed, for $g \in \mathcal{C}([0,1])$ fixed, $f_1, f_2 \in \mathcal{C}([0,1])$ we have

$$||\Gamma(f_1) - \Gamma(f_2)||_T \leq \delta ||f_1 - f_2||_T.$$

Since $\mathcal{C}([0,T])$ is a Banach space with the norm $||\cdot||_T$, the mapping Γ has a unique fixed point which we define as f.

2. **We show:** F is continuous.

For this purpose, fix $f_i = F(g_i), i = 1,2$, and set $\delta = ||g_1 - g_2||, e(t) = |f_1(t) - f_2(t)|, 0 \leq t \leq 1$. Then for any $0 \leq t \leq 1$ by the Lipschitz continuity of b

$$e(t) \leq B \int_0^t e(s)ds + \delta.$$

Hence by Gronwall's Lemma for any $0 \leq t \leq 1$

$$e(t) \leq \delta e^{Bt}, \text{ and globally } ||f_1 - f_2|| \leq e^B ||g_1 - g_2||.$$

This means that F is even Lipschitz continuous. □

Due to Lemma 5.1.1, the contraction principle is directly applicable and yields a LDP.

Theorem 5.1.2. *For $\epsilon > 0, x \in \mathbf{R}^d$ let $X^{\epsilon,x}$ be a solution of (5.1), and $\mu_\epsilon = \mathbb{P} \circ (X^{\epsilon,x})^{-1}$. Then $(\mu_\epsilon)_{\epsilon>0}$ satisfies an LDP on $\mathcal{C}([0,1])$ with good rate function*

$$I(f) = \begin{cases} \frac{1}{2}\int_0^1 |\dot{f}(t) - b(f(t))|^2 dt, & f - x \in \mathcal{H}_1, \\ \infty, & f - x \notin \mathcal{H}_1. \end{cases}$$

Proof. According to Lemma 5.1.1, $F(W_\epsilon) = X^{\epsilon,x}$, pathwise. Hence Schilder's Theorem 4.3.2 combines with the contraction principle 3.3.1 to yield an LDP for $(\mu_\epsilon)_{\epsilon>0}$ with good rate function

$$\overline{I}(f) = \inf_{\{g \in \mathcal{H}_1 : F(g) = f\}} \frac{1}{2} \int_0^1 |\dot{g}|^2(t) dt.$$

To prove that $\overline{I}(f) = I(f)$ for $f - x \in \mathcal{H}_1$, note that F is one-to-one. Moreover, for $g \in \mathcal{H}_1$ we have that f is a.e. differentiable and

$$\dot{f} = b(f) + \dot{g}.$$

Hence we have for $t \in [0,1]$ by virtue of the global Lipschitz continuity of b

$$|\dot{f}(t)| \leq B \int_0^t |f(s)|ds + |b(0)| + |\dot{g}(t)|,$$

and thus by boundedness of continuous functions $f - x \in \mathcal{H}_1$ as well and we can write

$$\overline{I}(f) = \frac{1}{2}\int_0^1 |\dot{f}(t) - b(f(t))|^2 dt = I(f).$$

□

5.2 An extension of the Freidlin-Wentzell theory

The LDP due to Freidlin and Wentzell presented in the previous section requires global Lipschitz conditions which are typically imposed in standard existence and uniqueness theorems for stochastic differential equations. In the setting of diffusions with a drift of the type of a potential gradient studied in the framework of stochastic resonance besides depending on time the coefficients will not be globally Lipschitz. We therefore need some extensions of the classical LDP result. This extension can be carried out in a general context (see Azencott [1] and two subsequent papers by Priouret [7] and Baldi et al. [2]). We prefer to present a simpler proof which permits to get the desired result in our particular framework.

Let us consider the family $X^{\epsilon,x}$, $x \in \mathbb{R}^d, \epsilon > 0$ of solutions of the SDE

$$X_t^{\epsilon,x} = x + \int_0^t b(X_s^{\epsilon,x})ds + \sqrt{\epsilon}W_t, \qquad t \geq 0. \tag{5.3}$$

Here b is locally Lipschitz continuous and satisfies the following growth condition: there are constants η, $R_0 > 0$ such that

$$\langle x, b(x)\rangle < -\eta|x| \quad \text{for all } |x| \geq R_0. \tag{5.4}$$

As a first consequence of this condition, the existence of a unique strong solution for (5.3) follows (see, for instance [10] Theorem 10.2.2).
Secondly, the growth condition (also called *dissipativity condition*) implies that the diffusion essentially stays inside a big ball $B_R(0)$ of radius R with very high probability: the probability for the diffusion to leave $B_R(0)$ is exponentially small. Two essential conclusions can be drawn from this observation: the law of the diffusion is exponentially tight in the space $C([0,1])^d$, and $X^{\epsilon,x}$ satisfies a LDP with a good rate function.

Let us make precise the exponential tightness of the diffusion paths first. We are interested in the small noise behavior of the exit time from the ball $B_R(0)$, defined by

$$\sigma_R^{\epsilon,x} := \inf\{t \geq 0 : |X_t^{\epsilon,x}| \geq R\}.$$

The following Theorem provides an asymptotic bound for $\sigma_R^{\epsilon,x}$. The arguments of its proof are borrowed from a treatment of self-attracting diffusions by means of large deviations techniques, see [8] and [6].

Theorem 5.2.1. *Let $x \in \mathbb{R}^d, \delta > 0$, and let $r : (0, \delta) \to (0, \infty)$ be a function satisfying*

$$\lim_{\epsilon \to 0} \frac{\epsilon}{r(\epsilon)} = 0.$$

Then there exist $R_1, \epsilon_1 > 0$ and $C > 0$ such that for $R \geq R_1$, $\epsilon < \epsilon_1$

$$\mathbb{P}\Big(\sigma_R^{\epsilon,x} \leq r(\epsilon)\Big) \leq C\eta^2 \, \frac{r(\epsilon)}{\epsilon} \, e^{-\frac{\eta R}{\epsilon}} \qquad \text{for} \quad |x| \leq \frac{R}{2}. \tag{5.5}$$

Remark 5.2.2. *The constants $R_1, \epsilon_1 > 0$ and $C > 0$ are universal in the sense that they do not depend on the particular choice of the drift b, but only on the parameter η and R_0 introduced in the growth condition* (5.4)*, and of course on the function r. Hence the bound is uniform in the class of all diffusions that satisfy* (5.4)*.*

Proof. For convenience of notation, we suppress the superscripts in $X^{\epsilon,x}, \sigma_R^{\epsilon,x}$ etc.
Step 1: First we determine a diffusion process $(Z_t,\ t \geq 0)$ which takes positive values, is easier to handle than X, and dominates it, i.e. such that $|X_t| \leq Z_t$ almost surely, for all $t \geq 0$.
Choose a C^2-function $h : \mathbb{R}^d \to \mathbb{R}$ such that

$$\begin{cases} h(x) = |x|, & \text{if } x \geq R_0, \\ |h(x)| \leq R_0, & \text{if } |x| \leq R_0, \end{cases}$$

where R_0 is the constant given in the growth condition (5.4). By Itô's formula we have for $t \geq 0$

$$h(X_t) = h(x) + \sqrt{\epsilon} \int_0^t \nabla h(X_s)\, dW_s + \int_0^t \langle \nabla h, b\rangle(X_s)\, ds + \frac{\epsilon}{2} \int_0^t \triangle h(X_s)\, ds.$$

For $t \geq 0$ let $\xi_t := \int_0^t |\nabla h(X_s)|^2\, ds$, i.e. ξ is the quadratic variation of the continuous local martingale $M := \int_0^\cdot \nabla h(X_s)\, dW_s$. Since $\nabla h(x) = \frac{x}{|x|}$ for $|x| \geq R_0$, we have $d\xi_t = dt$ on $\{|X_t| \geq R_0\}$.

We now introduce an auxiliary process Z that serves to control $|X|$.
Let $0 < \tilde{\eta} < \eta$. According to Skorokhod's lemma (see Revuz, Yor [9]) there is a unique pair of continuous adapted processes (Z, L) such that L is an increasing process (of finite variation) which increases only at times t for which $Z_t = R_0$, satisfies $Z \geq R_0$, and such that the equation

$$Z := R_0 \vee |x| + \sqrt{\epsilon} M - \tilde{\eta}\xi + L$$

is valid. **We show:** there exists $\epsilon_0 > 0$ such that for any $\epsilon < \epsilon_0$ and $t \geq 0$ we have $|X_t| \leq Z_t$ a.s.
For that purpose, choose $f \in C^2(\mathbb{R})$ such that

$$\begin{cases} f(x) > 0 \text{ and } f'(x) > 0 & \text{for all} \quad x > 0, \\ f(x) = 0 \quad \text{for all} \quad x \leq 0. \end{cases}$$

According to Itô's formula, for $t \geq 0$

$$f(h(X_t) - Z_t) = f(h(x) - R_0 \vee |x|) + \int_0^t f'(h(X_s) - Z_s)\, d(h(X) - Z))_s + \frac{1}{2}\int_0^t f''(h(X_s) - Z_s)\, d\langle h(X) - Z\rangle_s.$$

By definition of h and Z we have

$$\{h(X_t) > Z_t\} = \{h(X_t) > Z_t, |X_t| \leq R_0\} \cup \{h(X_t) > Z_t, |X_t| > R_0\} = \{|X_t| > Z_t\}.$$

Moreover by definition, $h(X) - Z$ is a process of finite variation, so that the bracket term in the preceding Itô formula vanishes. Hence the expression

$$\int_0^t f'\Big(|X_s| - Z_s\Big)\left\{\frac{1}{|X_s|}\langle X_s, b(X_s)\rangle + \frac{\epsilon}{2}\triangle h(X_s) + \tilde{\eta}\right\} ds - \int_0^t f'(|X_s| - Z_s)\, dL_s$$

is an upper bound of $f(h(X_t) - Z_t)$.
Furthermore, $\triangle h(x) = \frac{d-1}{|x|}$ for $|x| \geq R_0$, which by (5.4) implies

$$\frac{1}{|X_s|}\langle X_s, b(X_s)\rangle + \frac{\epsilon}{2}\triangle h(X_s) + \tilde{\eta} \ < \ \frac{\epsilon(d-1)}{2|X_s|} + \tilde{\eta} - \eta \quad \text{on } \{|X_s| > Z_s\}.$$

The latter expression is negative if ϵ is small enough. Summarizing, we can find $\epsilon_0 > 0$ such that $f(|X_t| - Z_t) \leq 0$ for $\epsilon < \epsilon_0$. This implies $|X_t| \leq Z_t$ a.s. by the definition of f, and Step 1 is established.

Step 2: We therefore can bound exit probabilities for X, expressed in terms of σ_R, by those for Z, expressed by an analogously defined stopping time τ_R. We have

$$\mathbb{P}(\sigma_R \leq r(\epsilon)) \leq \mathbb{P}(\tau_R \leq r(\epsilon))$$

for all $\epsilon < \epsilon_0$. **We compute:** $\mathbb{P}(\tau_R \leq r(\epsilon))$ for such ϵ. We have for any $\alpha > 0$

$$\mathbb{P}(\tau_R \leq r(\epsilon)) \leq e^{\alpha r(\epsilon)}\, \mathbb{E}[e^{-\alpha\sigma_R}]. \tag{5.6}$$

In order to find a bound for the right hand side of (5.6), let

$$K := \sup_{|x| \leq R_0} |\nabla h(x)|^2.$$

Then we have $\xi_t \leq Kt$ for all $t \geq 0$. Note that w.l.o.g. h can be chosen so that $K \leq 2R_0$. Now observe that, by Itô's formula, for any $\varphi \in C^2(\mathbb{R})$

$$d\Big(\varphi(Z_t)\, e^{-\frac{\alpha}{K}\xi_t}\Big) = \sqrt{\epsilon}\, \varphi'(Z_t) e^{-\frac{\alpha}{K}\xi_t}\, dM_t + \varphi'(Z_t)\, e^{-\frac{\alpha}{K}\xi_t}\, dL_t + e^{-\frac{\alpha}{K}\xi_t}\Big\{\frac{\epsilon}{2}\varphi''(Z_t) - \tilde{\eta}\varphi'(Z_t) - \frac{\alpha}{K}\varphi(Z_t)\Big\}\, d\xi_t.$$

Now let $R \geq R_0$. If we choose φ such that

$$\begin{cases} \frac{\epsilon}{2}\varphi''(y) - \tilde{\eta}\varphi'(y) - \frac{\alpha}{K}\varphi(y) = 0 & \text{for } y \in [R_0, R], \\ \varphi'(R_0) = 0, \quad \varphi(R) = 1, \end{cases} \tag{5.7}$$

then $\varphi(Z_t)e^{-\frac{\alpha}{K}\xi_t}$ is a local martingale which is bounded up to time τ_R. Hence we are allowed to apply Doob's optional sampling theorem to obtain

$$\varphi(R_0 \vee |x|) = \mathbb{E}[\varphi(Z_{\tau_R})e^{-\frac{\alpha}{K}\xi_{\tau_R}}] = \mathbb{E}[e^{-\frac{\alpha}{K}\xi_{\tau_R}}]. \tag{5.8}$$

But since $\xi_{\tau_R} \leq K\tau_R$, which implies $\mathbb{E}[e^{-\frac{\alpha}{K}\xi_{\tau_R}}] \geq \mathbb{E}[e^{-\alpha\tau_R}]$, and we deduce from (5.6) that

$$\mathbb{P}(\sigma_R \leq r(\epsilon)) \leq e^{\alpha r(\epsilon)}\, \mathbb{E}[e^{-\frac{\alpha}{K}\xi_{\tau_R}}] \leq e^{\alpha r(\epsilon)}\varphi(R_0 \vee |x|). \tag{5.9}$$

Step 3. We estimate φ satisfying (5.7).
Solving the differential equation for φ yields

$$\varphi(x) = \frac{-\lambda^- e^{\lambda^+(x-R_0)} + \lambda^+ e^{\lambda^-(x-R_0)}}{-\lambda^- e^{\lambda^+(R-R_0)} + \lambda^+ e^{\lambda^-(R-R_0)}}, \quad x \in [R_0, R],$$

with $\lambda^\pm = \frac{\tilde{\eta}\pm\sqrt{\tilde{\eta}^2+2\frac{\alpha}{K}\epsilon}}{\epsilon}$. Hence

$$\varphi(x) \leq \frac{(\lambda^+ - \lambda^-)\, e^{\lambda^+(x-R_0)}}{(-\lambda^-)\, e^{\lambda^+(R-R_0)}}, \quad x \in [R_0, R].$$

Taking $\alpha = r(\epsilon)^{-1}$ in (5.9) we obtain

$$\mathbb{P}(\sigma_R \leq r(\epsilon)) \leq e^1\, \varphi(R_0 \vee |x|) \leq \frac{\lambda^+ - \lambda^-}{-\lambda^-} \exp\Big\{1 + \lambda^+(R_0 \vee |x| - R)\Big\}.$$

It is obvious that $\exp\Big\{\lambda^+(R_0 \vee |x| - R)\Big\} \leq \exp\Big\{-\frac{\tilde{\eta}R}{\epsilon}\Big\}$ for $R \geq 2(|x| \vee R_0)$, so it remains to comment on the prefactor. We have

$$\frac{\lambda^+ - \lambda^-}{-\lambda^-} = \frac{2\sqrt{\tilde{\eta}^2 + 2\frac{\alpha}{K}\epsilon}}{\sqrt{\tilde{\eta}^2 + 2\frac{\alpha}{K}\epsilon} - \tilde{\eta}} \leq \frac{4\Big(\tilde{\eta}^2 + \frac{2\epsilon}{Kr(\epsilon)}\Big)}{\frac{2\epsilon}{Kr(\epsilon)}}.$$

Since $\lim_{\epsilon\to 0} \frac{\epsilon}{r(\epsilon)} = 0$, the latter bound behaves like $2\tilde{\eta}^2 K \frac{r(\epsilon)}{\epsilon}$ as $\epsilon \to 0$. Putting all estimates together yields the claimed asymptotic bound with $\tilde{\eta}$ instead of η. Finally, letting $\tilde{\eta} \to \eta$ establishes (5.5). □

Remark 5.2.3. *Theorem 5.2.1 can be easily extended to the context of time inhomogeneous diffusions. Let $X^{\epsilon,x}$ be the solution process of*

$$X_t^{\epsilon,x} = x + \int_0^t b(s, X_s^{\epsilon,x})ds + \sqrt{\epsilon}W_t, \qquad t \geq 0, \quad x \in \mathbb{R}^d,$$

where b is locally Lipschitz continuous with respect to both variables and satisfies the following growth condition: there are constants η, $R_0 > 0$ such that

$$\langle x, b(t,x)\rangle < -\eta|x| \quad \text{for all } t \geq 0,\ |x| \geq R_0. \tag{5.10}$$

Then in the notation of Theorem 5.2.1 the exit time $\sigma_R^{\epsilon,x}$ of $X^{\epsilon,x}$ satisfies

$$\mathbb{P}\Big(\sigma_R^{\epsilon,x} \leq r(\epsilon)\Big) \leq C\eta^2\, \frac{r(\epsilon)}{\epsilon}\, e^{-\frac{\eta R}{\epsilon}} \qquad \text{for} \quad |x| \leq \frac{R}{2}.$$

The result of Theorem 5.2.1 is much sharper than what we really need to obtain large deviations estimates. It shall play a crucial rule in Chapter **??**, where Remark 5.2.3 will be used. The importance of Theorem 5.2.1 for this section is linked to an immediate consequence obtained for $r(\epsilon) = T, \epsilon > 0$.

Corollary 5.2.4. *For all $R \geq R_1$ and $T > 0$ we have*

$$\limsup_{\epsilon \to 0} \epsilon \log \mathbb{P}(\sigma_R^{\epsilon,x} \leq T) \leq -\eta R.$$

For $x \in \mathbb{R}^d$ and $\epsilon > 0$ let μ_ϵ be the law of $X^{\epsilon,x}$ on the function space $C([0,T])^d$ endowed with the uniform norm. Then $(\mu_\epsilon)_{\epsilon>0}$ is exponentially tight.

Proof. For convenience let $T = 1$, choose $x \in \mathbb{R}^d$, and denote $\nu_\epsilon = \mathbb{P} \circ (\sqrt{\epsilon} W)^{-1}, \epsilon > 0$. For $\alpha > 0$ use Theorem 4.2.2 to choose a compact set $K \subset \mathcal{C}([0,1])$ such that

$$\limsup_{\epsilon \to 0} \epsilon \ln \nu_\epsilon(K^c) \leq -\alpha.$$

Next use Theorem 5.2.1 to choose $R > 0$ large enough to ensure $x \in B_{\frac{R}{2}}(0)$ and

$$\limsup_{\epsilon \to 0} \epsilon \ln \mathbb{P}(\sigma_R^{\epsilon,x} \leq 1) \leq -\eta R < -\alpha.$$

Now choose a bounded vector field $b_R : \mathbb{R}^d \to \mathbb{R}^d$ which coincides with b on $B_R(0)$ and which is globally Lipschitz continuous. Let F_R be the continuous map on $\mathcal{C}([0,1])$ corresponding to b_R according to Lemma 5.1.1, and denote

$$Y_t^{\epsilon,x} = x + \int_0^t b_R(Y_s^{\epsilon,x}) ds + \sqrt{\epsilon} W_t, \quad t \geq 0.$$

Then by definition $X^{\epsilon,x} = Y^{\epsilon,x}$ on $\{\sigma_R^{\epsilon,x} > 1\}$. Hence for ϵ small enough with the compact set $F_R(K) \subset \mathcal{C}([0,1])$

$$\begin{aligned} \mu_\epsilon(F_R(K)^c) &\leq \mathbb{P}(X^{\epsilon,x} \notin F_R(K), \sigma_R^{\epsilon,x} > 1) + \mathbb{P}(\sigma_R^{\epsilon,x} \leq 1) \\ &= \mathbb{P}(Y^{\epsilon,x} \notin F_R(K), \sigma_R^{\epsilon,x} > 1) + \mathbb{P}(\sigma_R^{\epsilon,x} \leq 1) \\ &\leq \mathbb{P}(Y^{\epsilon,x} \notin F_R(K)) + \mathbb{P}(\sigma_R^{\epsilon,x} \leq 1) \\ &= \nu_\epsilon(K^c) + \mathbb{P}(\sigma_R^{\epsilon,x} \leq 1). \end{aligned}$$

This implies

$$\limsup_{\epsilon \to 0} \epsilon \ln \mu_\epsilon(F_R(K)^c) \leq \limsup_{\epsilon \to 0} \epsilon \ln \nu_\epsilon(K^c) \vee \limsup_{\epsilon \to 0} \epsilon \ln \mathbb{P}(\sigma_R^{\epsilon,x} \leq 1) \leq -\alpha.$$

This establishes the desired exponential tightness. □

With this at hand, we are in a position to state the large deviations result for the diffusion processes $X^{\epsilon,x}, \epsilon > 0, x \in \mathbb{R}^d$.

Theorem 5.2.5. *For $\epsilon > 0$, $x \in \mathbb{R}^d$ let $X^{\epsilon,x}$ be a solution of* (5.3) *with locally Lipschitz drift term b that satisfies the growth condition* (5.10)*, and $\mu_\epsilon = \mathbb{P} \circ (X^{\epsilon,x})^{-1}$. Then $(\mu_\epsilon)_{\epsilon>0}$ satisfies an LDP on $(C([0,1])^d, \|\cdot\|)$ with good rate function*

$$I_x(f) = \begin{cases} \frac{1}{2}\int_0^1 |\dot{f}(t) - b(f(t))|^2 dt, & f - x \in \mathcal{H}_1, \\ \infty, & f - x \notin \mathcal{H}_1. \end{cases}$$

We recall that $\mathcal{H}_1$ is the Cameron-Martin function space and $\|\cdot\|$ is the uniform norm.

Proof. By the exponential tightness of Corollary 5.2.4, it suffices to prove a weak LDP. More precisely, Lemma 3.1.1 explains that it is enough to establish the lower bound for compact sets and the upper bound for open sets. The proof is based on a localization technique.

Step 1. Upper bound:
Let K be compact in $C([0,1])^d$ with respect to the uniform metric. Then there exists $R > 0$ such that $\sup_{f \in K} \|f\| \leq R$. We define a new drift b_R which is globally Lipschitz and equals b on $B_R(0)$: $b_R(x) = b(x)$ for $|x| \leq R$. We again denote by Y the diffusion related to b_R. By definition we have

$$\mathbb{P}(X^{\epsilon,x} \in K) = \mathbb{P}(Y^{\epsilon,x} \in K).$$

Since b_R is globally Lipschitz continuous, Freidlin-Wentzell's theory yields a large deviation principle with rate function I_x^R which agrees with I_x on K, whence:

$$\limsup_{\epsilon \to 0} \epsilon \log \mu_\epsilon(K) = \limsup_{\epsilon \to 0} \epsilon \log \mathbb{P}(Y^{\epsilon,x} \in K) \leq - \inf_{\varphi \in K} I_x^R(\varphi) = - \inf_{\varphi \in K} I_x(\varphi).$$

Step 2. Lower bound: The arguments are similar to the ones above. Let $G \subset C([0,1])^d$ be open, and fix $f \in G$. There exists $0 < \delta < 1$ such that $G_f = \{g \in C([0,1])^d : \|g - f\| < \delta\} \subset G$. By choosing R larger than $\sup_{g \in G_f} \|g\|$ and defining b_R and Y as in Step 1, we obtain

$$\mathbb{P}(X^{\epsilon,x} \in G) \geq \mathbb{P}(X^{\epsilon,x} \in G_f) = \mathbb{P}(Y^{\epsilon,x} \in G_f).$$

The large deviations principle for the diffusion Y gives

$$\liminf_{\epsilon \to 0} \epsilon \log \mu_\epsilon(G_f) = \limsup_{\epsilon \to 0} \epsilon \log \mathbb{P}(Y^{\epsilon,x} \in G_f) \geq - \inf_{\varphi \in G_f} I_x^R(\varphi) = - \inf_{\varphi \in G_f} I_x(\varphi).$$

Due to the arbitrary choice of $f \in G$, we conclude that

$$\liminf_{\epsilon \to 0} \epsilon \log \mu_\epsilon(G) \geq - \inf_{\varphi \in G} I_x(\varphi).$$

The upper bound is satisfied and I_x is a good rate function since $(\mu_\epsilon)_{\epsilon>0}$ is exponentially tight. □

The large deviation principle presented in Theorem 5.2.5 depends on the initial position x of the diffusion process. This dependence can be analyzed and controlled. Indeed, the large deviation principle for the diffusion (5.3) is uniform with respect to its starting point. This fact which will also be very useful in the treatment of the asymptotic properties of the exit time will be established next.

Proposition 5.2.6. (Uniform LDP) *Let $F \subset C([0,1])^d$ be closed, $G \subset C([0,1])^d$ open, and $x \in \mathbb{R}^d$. Then we have*

$$(a)\ \limsup_{\epsilon\to 0, y\to x} \epsilon \log P(X^{\epsilon,y} \in F) \leq -\inf_{f\in F} I_x(f),$$

$$(b)\ \liminf_{\epsilon\to 0, y\to x} \epsilon \log P(X^{\epsilon,y} \in G) \geq -\inf_{f\in G} I_x(f).$$

Proof. According to Theorem 3.3.3 **we have to show:**
for any family $(x_\epsilon)_{\epsilon>0}$ such that $\lim_{\epsilon\to 0} x_\epsilon = x$ the corresponding families of probability measures $\mu_\epsilon = P \circ (X^{\epsilon,x})^{-1}, \epsilon > 0$, and $\tilde{\mu}_\epsilon = P \circ (X^{\epsilon,x_\epsilon})^{-1}, \epsilon > 0$, are exponentially equivalent in $(C([0,1])^d, \|\cdot\|)$.

To see this, fix $\epsilon > 0, x \in \mathbb{R}^d$. Then

$$\|X^{\epsilon,x_\epsilon} - X^{\epsilon,x}\| \leq |x_\epsilon - x| + \int_0^1 |b(X_u^{\epsilon,x_\epsilon}) - b(X_u^{\epsilon,x})|\, du.$$

Let us now use a localization argument. We fix $R > 0$ and define the first exit times of the diffusions $X^{\epsilon,x}$ and X^{ϵ,x_ϵ} of $B_R(0)$ by

$$\sigma_R^{\epsilon,x} = \inf\{t \geq 0 :\ X_t^{\epsilon,x} \notin B_R(0)\},$$

$$\tau_R^{\epsilon,x} = \inf\{t \geq 0 :\ X_t^{\epsilon,x_\epsilon} \notin B_R(0)\},$$

and $\tilde{\sigma}_R^{\epsilon,x} = \sigma_R^{\epsilon,x} \wedge \tau_R^{\epsilon,x}$. Then on the event $\{\tilde{\sigma}_R^{\epsilon,x} > 1\}$, the Lipschitz continuity of b on $B_R(0)$ with Lipschitz constant L_R implies

$$|b(X_u^{\epsilon,x_\epsilon}) - b(X_u^{\epsilon,x})| \leq L_R |X_u^{\epsilon,x_\epsilon} - X_u^{\epsilon,x}|, \quad 0 \leq u \leq 1.$$

Using Gronwall's Lemma (see **??**), we obtain on $\{\tilde{\sigma}_R^{\epsilon,x} > 1\}$

$$\|X^{\epsilon,x_\epsilon} - X^{\epsilon,x}\| \leq |x_\epsilon - x| e^{L_R}.$$

From this we deduce that the two sample paths are close together before one of them exits from $B_R(0)$. Let $\delta > 0$. Then, if we choose ϵ_0 small enough, for all $\epsilon < \epsilon_0$

$$\|X^{\epsilon,x_\epsilon} - X^{\epsilon,x}\| \leq \delta \quad \text{on } \{\tilde{\sigma}_R^{\epsilon,x} > 1\}.$$

Theorem 5.2.1 implies for R large enough so that for $\epsilon < \epsilon_0$ we have $x, x_\epsilon \in B_{\frac{R}{2}}(0)$

$$\begin{aligned}
\limsup_{\epsilon\to 0} \epsilon \log \mathbb{P}(\|X^{\epsilon,x_\epsilon} - X^{\epsilon,x}\| > \delta) &\leq \limsup_{\epsilon\to 0} \epsilon \log \mathbb{P}(\tilde{\sigma}_R^{\epsilon,x} \leq 1) \\
&\leq \limsup_{\epsilon\to 0} \epsilon \log \mathbb{P}(\sigma_R^{\epsilon,x} \leq 1) \vee \limsup_{\epsilon\to 0} \epsilon \log \mathbb{P}(\tau_R^{\epsilon,x} \leq 1) \\
&\leq -\eta R.
\end{aligned}$$

Letting $R \to \infty$ allows to obtain the exponential equivalence between $(\mu_\epsilon)_{\epsilon>0}$ and $(\tilde{\mu}_\epsilon)_{\epsilon>0}$. □

We can draw a conclusion from Proposition 5.2.6 which will turn out to be of practical use later: an LDP which is uniform on compact sets of initial states for the diffusion.

Corollary 5.2.7. *Let $K \subset \mathbb{R}^d$ be compact, $F \subset C([0,1])^d$ closed, $G \subset C([0,1])^d$ open. Then we have*

$$\text{(a)} \ \limsup_{\epsilon \to 0} \sup_{y \in K} \epsilon \log P(X^{\epsilon,y} \in F) \le - \inf_{y \in K, f \in F} I_y(f),$$

$$\text{(b)} \ \liminf_{\epsilon \to 0} \inf_{y \in K} \epsilon \log P(X^{\epsilon,y} \in G) \ge - \sup_{y \in K} \inf_{f \in G} I_y(f).$$

Proof. For similarity of arguments, we only show the upper bound. Let $-I_K$ denote the right hand side of the claimed inequality. For $\delta > 0$ let $I_K^\delta = \min\{I_K - \delta, \frac{1}{\delta}\}$. Now fix $x \in K$. Then by Proposition 5.2.6 and lower semi-continuity of I there exists $\epsilon_x > 0$ such that for any $\epsilon \le \epsilon_x$

$$\epsilon \log \sup_{y \in B_{\epsilon_x}(x)} \mathbb{P}(X^{\epsilon,y} \in F) \le -I_K^\delta.$$

Use compactness of K to choose $x_1, \cdots, x_m \in K$ such that $K \subset \cup_{i=1}^m B_{\epsilon_{x_i}}(x_i)$. Then we have

$$\limsup_{\epsilon \to 0} \epsilon \log \sup_{y \in K} \mathbb{P}(X^{\epsilon,y} \in F) \le \max_{1 \le i \le m} \limsup_{\epsilon \to 0} \epsilon \log \sup_{y \in B_{\epsilon_{x_i}}(x_i)} \mathbb{P}(X^{\epsilon,y} \in F) \le -I_K^\delta.$$

It remains to let $\delta \to 0$. □

Chapter 6

Diffusion exit from a domain

We now return to the problem sketched in the introduction. We shall present the treatment of the exit of a diffusion process from a domain in the simpler case of additive Gaussian noise via the theory by Freidlin and Wentzell, and roughly follow Dembo, Zeitouni [5].

6.1 Properties of the pseudopotential and statement of main result

Consider the SDE

$$X_t^{\epsilon,x} = x + \int_0^t b(X_s^{\epsilon,x})ds + \sqrt{\epsilon}W_t, \quad x \in \mathbf{R}^d, t \geq 0. \tag{6.1}$$

Let $G \subset \mathbb{R}^d$ be a bounded domain. Suppose $b : \mathbf{R}^d \to \mathbf{R}^d$ is a uniformly Lipschitz continuous function, i.e. there exists a constant B such that

$$|b(x) - b(y)| \leq B|x - y|, \quad x, y \in \mathbf{R}^d, \tag{6.2}$$

$$\sup_{x \in \mathbb{R}^d} \{|b(x)|\} \leq B. \tag{6.3}$$

We view $X^{\epsilon,x}$ as a perturbation of the ordinary differential equation

$$\dot{\phi}(t) = b(\phi(t)), \quad t \geq 0,$$

or the dynamical system associated with it. Suppose that the dynamical system has the following properties:

Assumption (A1)
$0 \in G$ is the unique stable equilibrium point, and

$$\phi(0) \in G \text{ implies that for all } t > 0 \text{ we have } \phi(t) \in G, \ \lim_{t\to\infty} \phi(t) = 0.$$

We are interested in the time it takes typically for the trajectories of $X^{\epsilon,x}$ starting at $x \in G$ to leave the domain. More formally, we suppose that the filtration $(\mathcal{F}_t)_{t\geq 0}$ on our probability space $(\Omega, \mathcal{F}, P)$ satisfies the usual conditions, so that the objects of our main interest, the *exit times from G*

$$\tau^{\epsilon,x} = \inf\{t > 0 : X_t^{\epsilon,x} \in \partial G\}, \epsilon > 0, x \in \mathbf{R}^d,$$

are well defined stopping times. Chapter 5 provides explicit formulas for the good rate functions of LDP for the solution processes of (6.1) indexed by $[0,1]$. For diffusions with time duration $t \geq 0$ we continue denoting by $||\cdot||_t$ the uniform norm on $\mathcal{C}([0,t])$ for any $t \geq 0$, and write $||\cdot||$ instead if this is unambiguous. The scaling property of Brownian motion allows a direct extension of the results of Chapter 5 to $[0,t]$ instead of $[0,1]$. In this context the good rate functions are given by

$$I_{x,t}(f) = \begin{cases} \inf_{\{g\in\mathcal{H}_1([0,t]):f=x+\int_0^\cdot b(f(s))ds+g\}} \frac{1}{2}\int_0^t \dot{g}^2(s)ds, & g \in \mathcal{H}_1([0,t]), \\ \infty, & g \notin \mathcal{H}_1([0,t]). \end{cases}$$

Here

$$\begin{aligned}\mathcal{H}_1([0,t]) \quad = \quad & \{g \in \mathcal{C}_0([0,t]) : \text{ there exists } \dot{g} \in L_2([0,t]), \text{ such that} \\ & g = \int_0^\cdot \dot{g}(s)\, ds, \ \int_0^t |\dot{g}|^2(s)ds < \infty\}.\end{aligned}$$

We define the *cost function*

$$V(x,z,t) = \inf_{\{f\in\mathcal{C}([0,t]):f(t)=z\}} I_{x,t}(f), \quad x,z \in \mathbf{R}^d, t > 0.$$

The cost function quantifies the cost for forcing the system to z at time t when starting at x. We further define

$$V(x,z) = \inf_{t>0} V(x,z,t), \ x,z \in \mathbf{R}^d,$$

and call $V(0,z)$ *quasi-potential* of the system. The quasi-potential describes the minimal cost for the system to go to z when starting at the stable equilibrium. To derive the exit time law in terms of the quasi-potential, we shall need the following assumptions.

Assumption (A2)
If $\phi(0) \in \partial G$, then $\lim_{t\to\infty} \phi(t) = 0$.

Assumption (A3)

$$\overline{V} = \inf_{z\in\partial G} V(0,z) < \infty.$$

Remark
(A2) prevents the boundary ∂G to be critical or *characteristic*.

We now prove an important *controllability* property of our diffusion. We shall see that within small neighborhoods of the boundary the smooth system can be controlled from any point x_0 into any other point y_0. Denote $d(x,F) = \inf\{|x-y| : y \in F\}$ for $x \in \mathbf{R}^d$, and a closed subset F in $\mathbf{R}^d$.

Lemma 6.1.1. *We have*

(i) $V(x,y,|x-y|) \leq \frac{1}{2}[1+B]^2|x-y|, \quad x,y \in \mathbb{R}^d.$

(ii) For all $\rho > 0$, $x_0, y_0 \in \mathbf{R}^d, z \in \partial G \cup \{0\}$ *satisfying* $|x_0 - z| + |z - y_0| \leq \rho$ *there exists* $u \in \mathcal{C}([0,\rho])$ *such that* $||u|| \leq 1 + B$ *and* $\phi(\rho) = y_0$ *for the solution* ϕ *of*

$$\phi(t) = x_0 + \int_0^t b(\phi(s))ds + \int_0^t u(s)ds, \quad 0 \leq t \leq \rho.$$

Proof. 1. Let $x, y \in \mathbb{R}^d$ be given. Define

$$\phi(t) = x + t\frac{y-x}{|y-x|}, \quad t \in [0, |x-y|],$$

and

$$u(t) = (\dot{\phi}(t) - b(\phi(t))), \quad t \in [0, |x-y|].$$

Then on $[0, |x-y|]$ we have $\dot{\phi} = b(\phi) + u$, and by assumptions on b we have

$$||u|| \leq 1 + B.$$

Hence

$$V(x,y,|x-y|) \leq \frac{1}{2}\int_0^{|x-y|} |u|^2(t)dt \leq \frac{1}{2}(1+B)^2|x-y|.$$

2. Let $0 < \rho$, $x_0, y_0 \in \mathbf{R}^d$, $z \in \partial G \cup \{0\}$, such that $|x_0 - z| + |y_0 - z| \leq \rho$. Define

$$\phi(t) = x_0 + t\frac{y_0 - x_0}{\rho}, \quad t \in [0,\rho],$$

and

$$u(t) = \dot{\phi}(t) - b(\phi(t)), \quad t \in [0,\rho].$$

Then for all $t \in [0,\rho]$ we have $|\phi(t) - z| \leq (1 - \frac{t}{\rho})|x_0 - z| + \frac{t}{\rho}|y_0 - z| \leq \rho$. And we have according to the calculation in part 1.

$$||u|| \leq 1 + B.$$

□

From Lemma 6.1.1 we may deduce the following continuity property for the cost function.

Lemma 6.1.2. *For any* $\delta > 0$ *there exists* $\rho > 0$ *such that*

(i) $\sup_{|x|,|y|\leq\rho} \inf_{t\in[0,1]} V(x,y,t) < \delta,$

(ii) $\sup_{\{x,y:\inf_{z\in\partial G}(|x-z|+|y-z|\leq\rho\}} \inf_{t\in[0,1]} V(x,y,t) < \delta.$

Proof. By the second part of Lemma 6.1.1 we have for two points x, y as specified in (i) or (ii) and $0 < \rho$

$$V(x,y,\rho) \leq \frac{1}{2}(1+B)^2\rho.$$

Now the result follows. □

We state the main result of this Chapter.

Theorem 6.1.3. *Assume (A1)-(A3) are satisfied. Then for any* $\delta > 0, x \in G$ *we have*

$$\lim_{\epsilon \to 0} \mathbb{P}(e^{\frac{\overline{V}+\delta}{\epsilon}} > \tau^{\epsilon,x} > e^{\frac{\overline{V}-\delta}{\epsilon}}) = 1.$$

Moreover, for all $x \in G$ *we have*

$$\lim_{\epsilon \to 0} \epsilon \log \mathbb{E}(\tau^{\epsilon,x}) = \overline{V}.$$

6.2 Proof of main result: the upper bound

In this part, we shall establish an upper bound for the exit rate of Theorem 6.1.3. This will be done in 3 steps. To begin, it will be necessary to give an exponential lower bound for the probability that $\tau^{\epsilon,x}$ is small, uniform in x in a small neighborhood of 0.

1. **We show:**

Lemma 6.2.1. *Let* $\eta > 0$. *Then there exists* $\rho_0 > 0$ *such that for* $0 < \rho < \rho_0$ *there exists* $T_0 > 0$ *such that*

$$\liminf_{\epsilon \to 0} \epsilon \log \inf_{|x| \le \rho} \mathbb{P}(\tau^{\epsilon,x} \le T_0) > -(\overline{V} + \eta).$$

Proof. (of Lemma 6.2.1) Let $\delta = \frac{\eta}{6}$, and let ρ_0 be so small that the inequalities of Lemma 6.1.2 hold. Use Lemma 6.1.2, (i) to choose for $x \in G$ such that $|x| \le \rho$ a path $\psi^x \in \mathcal{C}([0, t_x])$, with $0 \le t_x \le 1$, satisfying $\psi^x(0) = x, \psi^x(t_x) = 0$, and

$$I_{x,t_x}(\psi^x) \le \frac{\eta}{3}.$$

Use (A3) and Lemma 6.1.2, (ii) to choose $z \in \mathbf{R}^d \setminus \overline{G}$, $T_1 > 0, \phi \in \mathcal{C}([0, T_1])$ such that $\phi(0) = 0, \phi(T_1) = z$, and such that

$$I_{0,T_1}(\phi) \le \overline{V} + \frac{\eta}{3}.$$

Let $\hat{\phi}$ be the solution of the differential equation $\dot{\hat{\phi}} = b(\hat{\phi})$ with $\hat{\phi}(0) = z$. Next, let $T_0 = T_1 + 1$, and set for $x \in G$ such that $|x| \le \rho$

$$\phi_t^x = \begin{cases} \psi^x(t), & 0 \le t \le t_x, \\ \phi(t - t_x), & t_x \le t \le T_1 + t_x, \\ \hat{\phi}(t - T_1 - t_x), & T_1 + t_x \le t \le T_0. \end{cases}$$

Then we have

$$I_{x,T_0}(\phi) \le I_{x,t_x}(\psi^x) + I_{0,T_1}(\phi) \le \overline{V} + \frac{2\eta}{3}.$$

Now let $\Delta = d(z, \overline{G})$, and consider the open set

$$\Psi = \cup_{|x| \le \rho} \{\psi \in \mathcal{C}([0, T_0]) : ||\psi - \phi^x|| < \frac{\Delta}{2}\}.$$

Since ϕ^x visits z, by definition of Ψ every path $\psi \in \Psi$ leaves G before time T_0. Hence by virtue of Corollary 5.2.7 we have

$$\begin{aligned}
\liminf_{\epsilon\to 0} \epsilon \log \inf_{|x|\le\rho} \mathbb{P}(\tau^{\epsilon,x} \le T_0) &\ge \liminf_{\epsilon\to 0} \epsilon \log \inf_{|x|\le\rho} \mathbb{P}(X^{\epsilon,x} \in \Psi) \\
&\ge -\sup_{|x|\le\rho} \inf_{\psi\in\Psi} I_{x,T_0}(\psi) \\
&\ge -\sup_{|x|\le\rho} I_{x,T_0}(\phi^x) \\
&\ge -(\overline{V} + \frac{2\eta}{3}) > -(\overline{V} + \eta).
\end{aligned}$$

□

2. We next need to show that the probability that the diffusion stays inside G without hitting a small neighborhood of 0 is exponentially vanishing. For this purpose, for $\rho > 0, x \in \mathbf{R}^d$ such that $B_\rho(0) \subset G$ let

$$\sigma_\rho^x = \inf\{t : |X_t^{\epsilon,x}| \le \rho \text{ or } X_t^{\epsilon,x} \in \partial G\}.$$

We show:

Lemma 6.2.2. *We have*

$$\lim_{t\to\infty} \limsup_{\epsilon\to 0} \epsilon \log \sup_{x\in G} \mathbb{P}(\sigma_\rho^x > t) = -\infty.$$

Proof. (of Lemma 6.2.2) (i) For $t \ge 0$ consider the closed set in $\mathcal{C}([0,t])$

$$\Psi_t = \{\phi \in \mathcal{C}([0,t]) : \phi(s) \in \overline{G \setminus B_\rho(0)}, s \in [0,t]\}.$$

We show:

$$\limsup_{\epsilon\to 0} \epsilon \log \sup_{x\in G} \mathbb{P}(\sigma_\rho^x > t) \le -\inf_{\psi\in\Psi_t} I_{\psi(0),t}(\psi).$$

Indeed, by definition and Corollary 5.2.7

$$\begin{aligned}
\limsup_{\epsilon\to 0} \epsilon \log \sup_{x\in G} \mathbb{P}(\sigma_\rho^x > t) &\le \limsup_{\epsilon\to 0} \epsilon \log \sup_{x\in \overline{G\setminus B_\rho(0)}} \mathbb{P}(\sigma_\rho^x > t) \\
&\le -\inf_{x\in\overline{G\setminus B_\rho(0)}} \inf_{\psi\in\Psi_t} I_{x,t}(\psi) \\
&= -\inf_{\psi\in\Psi_t} I_{\psi(0),t}(\psi).
\end{aligned}$$

(ii) **It remains to show:**

$$\lim_{t\to\infty} \inf_{\psi\in\Psi_t} I_{\psi(0),t}(\psi) = \infty. \tag{6.4}$$

(ii.1) Let $(\Phi_t)_{t\ge 0}$ be the flow of the differential equation $\dot{\phi} = b(\phi)$. According to (A2) and (A1) for any $x \in \overline{G \setminus B_\rho(0)}$ there is $t_x \ge 0$ such that $\Phi_{t_x}(x) \in B_{\rho/2}(0)$. For $x \in \overline{G \setminus B_\rho(0)}$ define now

$$W_x = \Phi_{t_x}^{-1}(B_{\rho/2}(0)).$$

By definition, W_x is an open neighborhood of x. Now choose $x_1, \cdots, x_n \in \overline{G \setminus B_\rho(0)}$ such that $\overline{G \setminus B_\rho(0)} \supset \cup_{1 \leq i \leq n} W_{x_i}$ and set $T = \max_{1 \leq i \leq n} t_{x_i}$. Then any solution trajectory of $\dot{\phi} = b(\phi)$ starting in $\overline{G \setminus B_\rho(0)}$ hits $B_{\rho/2}(0)$ before time T.

(ii.2) **Assume:** (6.4) does not hold. Then there exists $M > 0$ such that for any $n \in \mathbb{N}$ there exists $\psi^n \in \Psi_{nT}$ such that $I_{\psi^n(0),nT}(\psi^n) \leq M$. Now for $0 \leq k \leq n-1$ let

$$\psi^{n,k}(t) = \psi^n(t - kT), \quad 0 \leq t \leq T.$$

Then $\psi^{n,k} \in \Psi_T$ and

$$M \geq I_{\psi^n(0),nT}(\psi^n) = \sum_{k=0}^{n-1} I_{\psi^n(kT),T}(\psi^{n,k}) \geq n \min_{0 \leq k \leq n-1} I_{\psi^{n,k}(0),T}(\psi^{n,k}).$$

Hence there exists a sequence $(\phi^n)_{n \in \mathbb{N}}$ in Ψ_T such that

$$\lim_{n \to \infty} I_{\phi^n(0),T}(\phi^n) = 0.$$

By compactness of $\{\phi \in \mathcal{C}([0,T]) : \phi(0) \in \overline{G \setminus B_\rho(0)}, I_{\phi(0),T}(\phi) \leq 1\}$, which follows directly from the goodness of the rate function, the sequence possesses a cluster point $\phi^* \in \Psi_T$. By lower semicontinuity of $\psi \mapsto I_{\psi(0),T}$ we obtain $I_{\phi^*(0),T}(\phi^*) = 0$. But this means that ϕ^* is a solution of the differential equation $\dot{\phi} = b(\phi)$, with $\phi^*(0) \in \overline{G \setminus B_\rho(0)}$. Hence by what has been proved in (ii.1), ϕ^* reaches $B_{\rho/2}(0)$ before time T. This contradicts $\phi^* \in \Psi_T$. □

3. We are now in a position to establish the upper bound. **We show:** for $x \in G, \delta > 0$ we have

$$\lim_{\epsilon \to 0} \mathbb{P}(\tau^{\epsilon,x} < e^{\frac{\overline{V}+\delta}{\epsilon}}) = 1.$$

(i) To do this, **we first show:** for any $\delta > 0$ there exists $T > 0$ and $\epsilon_0 > 0$ such that

$$\inf_{x \in G} \mathbb{P}(\tau^{\epsilon,x} \leq T) \geq e^{-\frac{\overline{V}+\frac{\delta}{2}}{\epsilon}}.$$

First use Lemma 6.2.1 to choose $T_0 > 0$ and $\rho > 0$ such that

$$\liminf_{\epsilon \to 0} \epsilon \log \inf_{|x| \leq \rho} \mathbb{P}(\tau^{\epsilon,x} \leq T_0) > -(\overline{V} + \frac{\delta}{4}),$$

and hence $\epsilon_0 > 0$ such that for $\epsilon < \epsilon_0$

$$\inf_{|x| \leq \rho} \mathbb{P}(\tau^{\epsilon,x} \leq T_0) \geq \exp(-\frac{\overline{V} + \frac{\delta}{4}}{\epsilon}).$$

Next, use Lemma 6.2.2 to choose, for $\rho > 0$ given, $T_1 > 0$ such that

$$\limsup_{\epsilon \to 0} \epsilon \log \sup_{x \in G} \mathbb{P}(\sigma_\rho^x > T_1) < 0,$$

and hence $r > 0$ and $\epsilon_0 > 0$ (eventually smaller than the one above) such that for $\epsilon < \epsilon_0$

$$\epsilon \log \sup_{x \in G} \mathbb{P}(\sigma_\rho^x > T_1) < -r,$$

and such that $1 - e^{-\frac{r}{\epsilon}} > e^{-\frac{\delta}{4\epsilon}}$. Now note that for $x \in G$ on the set $\{\sigma_\rho^x < \tau^{\epsilon,x}\}$ we have

$$\tau^{\epsilon,x} = \sigma_\rho^x + \tau^{\epsilon, X^{\epsilon,x}_{\sigma_\rho^x}} \circ \vartheta_{\sigma_\rho^x},$$

where ϑ_s denotes the usual shift on path space by time s. Hence by the strong Markov property for $\epsilon < \epsilon_0$

$$\begin{aligned}
\mathbb{P}(\tau^{\epsilon,x} \leq T_0 + T_1) &\geq \mathbb{P}(\sigma_\rho^x \leq T_1)\, \mathbb{P}(\tau^{\epsilon, X^{\epsilon,x}_{\sigma_\rho^x}} \leq T_0) \\
&\geq \inf_{x \in G} \mathbb{P}(\sigma_\rho^x \leq T_1) \inf_{|x| \leq \rho} \mathbb{P}(\tau^{\epsilon,x} \leq T_0) \\
&\geq e^{-\frac{\overline{V} + \frac{\delta}{4}}{\epsilon}} (1 - e^{-\frac{r}{\epsilon}}) \\
&\geq e^{-\frac{\overline{V} + \frac{\delta}{4}}{\epsilon}} e^{-\frac{\delta}{4\epsilon}} \\
&= e^{-\frac{\overline{V} + \frac{\delta}{2}}{\epsilon}}.
\end{aligned}$$

It remains to set $T = T_0 + T_1$.

(ii) Abbreviate $q = \inf_{x \in G} \mathbb{P}(\tau^{\epsilon,x} \leq T)$. For $k \in \mathbb{N}, x \in G, \epsilon > 0$ consider the events $\{\tau^{\epsilon,x} > kT\}$. Then, since $\tau^{\epsilon,x} = \tau^{\epsilon, X^{\epsilon,x}_{kT}} \circ \vartheta_{kT} + kT$ in the set $\{\tau^{\epsilon,x} > kT\}$, by conditioning on $\mathcal{F}_{kT}$ and the strong Markov property

$$\begin{aligned}
\mathbb{P}(\tau^{\epsilon,x} > (k+1)T) &= (1 - \mathbb{P}(\tau^{\epsilon,x} \leq (k+1)T | \tau^{\epsilon,x} > kT)\ \mathbb{P}(\tau^{\epsilon,x} > kT) \\
&\leq (1-q) \mathbb{P}(\tau^{\epsilon,x} > kT).
\end{aligned}$$

Hence by recursion for $k \in \mathbb{N}, \epsilon < \epsilon_0$

$$\sup_{x \in G} \mathbb{P}(\tau^{\epsilon,x} > kT) \leq (1-q)^k.$$

Therefore we obtain the following bound for the expected exit time

$$\sup_{x \in G} \mathbb{E}(\tau^{\epsilon,x}) \leq T \sum_{k=0}^\infty \sup_{x \in G} \mathbb{P}(\tau^{\epsilon,x} > kT) \leq T \sum_{k=0}^\infty (1-q)^k = \frac{T}{q}.$$

In particular, since $q \geq e^{-\frac{\overline{V} + \frac{\delta}{2}}{\epsilon}}$, we have

$$\sup_{x \in G} \mathbb{E}(\tau^{\epsilon,x}) \leq T e^{\frac{\overline{V} + \frac{\delta}{2}}{\epsilon}}.$$

Finally, using Chebyshev's inequality, we arrive at

$$\mathbb{P}(\tau^{\epsilon,x} \geq e^{\frac{\overline{V} + \delta}{\epsilon}}) \leq e^{-\frac{\overline{V} + \delta}{\epsilon}}\ \mathbb{E}(\tau^{\epsilon,x}) \leq T e^{-\frac{\delta}{2\epsilon}},$$

valid for any $x \in G, \epsilon < \epsilon_0$. It remains to let $\epsilon \to 0$ to obtain the upper bound.

6.3 Proof of main result: the lower bound

We now establish the lower bound for the exit rate of Theorem 6.1.3. Again we proceed in three main steps. We first need an estimate which shows that starting in G, the diffusion has high probability of being attracted to a small neighborhood of 0 before getting to ∂G.

1. Recall for $x \in G, \rho > 0$ such that $\overline{K}_\rho(0) \subset G$

$$\sigma_\rho^x = \inf\{t : X_t^{\epsilon,x} \in \partial G \cup \overline{K}_\rho(0)\}.$$

We show:

Lemma 6.3.1. *For any $x \in G, \rho > 0$ such that $\overline{K}_\rho(0) \subset G$ we have*

$$\lim_{\epsilon \to 0} \mathbb{P}(X_{\sigma_\rho^x}^{\epsilon,x} \in \overline{K}_\rho(0)) = 1.$$

Proof. (of Lemma 6.3.1) Given $\rho > 0$, we may and do consider $x \in G \setminus \overline{K}_\rho(0)$, for otherwise the claim is trivial. Let ϕ be the solution of the differential equation $\dot{\phi} = b(\phi), \phi(0) = x$,

$$T = \inf\{t \geq 0 : \phi(t) \in \overline{K}_{\frac{\rho}{2}}(0)\}.$$

According to (A2), $T < \infty$. Further, according to (A1), we have

$$\Delta = \rho \wedge d(\phi([0,T]), \partial G) > 0,$$

where $d(K, L)$ denotes the well defined distance of two compact sets $K, L \subset \mathbf{R}^d$. Then for $0 \leq t \leq T$ by the global Lipschitz continuity of b

$$|X_t^{\epsilon,x} - \phi(t)| \leq B \int_0^t |X_s^{\epsilon,x} - \phi(s)| ds + \sqrt{\epsilon} |W_t|.$$

By Gronwall's Lemma

$$\sup_{0 \leq t \leq T} |X_t^{\epsilon,x} - \phi(t)| \leq e^{BT} \sqrt{\epsilon} \sup_{0 \leq t \leq T} |W_t|.$$

Therefore

$$\begin{aligned}
\mathbb{P}(X_{\sigma_\rho^x}^{\epsilon,x} \notin \overline{K}_\rho(0)) &\leq \mathbb{P}(\sup_{0 \leq t \leq T} |X_t^{\epsilon,x} - \phi(t)| > \frac{\Delta}{2}) \\
&\leq \mathbb{P}(\sup_{0 \leq t \leq T} |W_t| > \frac{\Delta}{2\sqrt{\epsilon}} e^{-BT}) \to 0
\end{aligned}$$

as $\epsilon \to 0$, due to Chebyshev's and Doob's inequalities. This implies the desired result. □

2. We next have to establish an auxiliary result which says that during bounded time intervals, the diffusion cannot get away too far from its starting point.

We show:

Lemma 6.3.2. *Let $\rho > 0, c > 0$. Then there exists a constant $T(c,\rho)$ such that*

$$\limsup_{\epsilon\to 0} \epsilon \log \sup_{x\in G} \mathbb{P}(\sup_{0\le t\le T(c,\rho)} |X_t^{\epsilon,x} - x| \ge \rho) < -c.$$

Proof. (of Lemma 6.3.2) Fix $\rho > 0, \epsilon > 0$. Then for $0 \le t \le T \le \frac{\rho}{2B}, x \in G$

$$|X_t^{\epsilon,x} - x| \quad \le \quad |\int_0^t b(X_s^{\epsilon,x})ds| + \sqrt{\epsilon}|W_t| \le \frac{\rho}{2} + \sqrt{\epsilon}|W_t|.$$

Hence for $x \in G$ by the reflection principle and Theorem 4.1.1

$$\begin{aligned}
\mathbb{P}(\sup_{0\le t\le T} |X_t^{\epsilon,x} - x| \ge \rho) &\le \mathbb{P}(\sup_{0\le t\le T} |W_t| \ge \frac{\rho}{2\sqrt{\epsilon}}) \\
&= \mathbb{P}(\sup_{0\le s\le 1} |W_s| \ge \frac{\rho}{2\sqrt{\epsilon T}}) \\
&\le d\,\mathbb{P}(\sup_{0\le s\le 1} |W_s^1| \ge \frac{\rho}{2d\sqrt{\epsilon T}}) \\
&\le 2\,d\,\mathbb{P}(\sup_{0\le s\le 1} W_s^1 \ge \frac{\rho}{2d\sqrt{\epsilon T}}) \\
&= 4\,d\,\mathbb{P}(W_1^1 \ge \frac{\rho}{2d\sqrt{\epsilon T}}) \\
&\le 4\,d\,e^{-\frac{\rho^2}{8d^2\epsilon T}}.
\end{aligned}$$

This finally implies for $0 \le T \le \frac{\rho}{2B}$

$$\begin{aligned}
\limsup_{\epsilon\to 0} \epsilon \log \sup_{x\in G} \mathbb{P}(\sup_{0\le t\le T} |X^{\epsilon,x} - x| \ge \rho) &\le \limsup_{\epsilon\to 0} \epsilon \log 4de^{-\frac{\rho^2}{8d^2T\epsilon}} \\
&= -\frac{\rho^2}{8d^2T}.
\end{aligned}$$

For $T = T(c,\rho)$ small enough, the last line of the preceding inequality is bounded above by $-c$. □

3. We need a final auxiliary result relating the quasi-potential with the probability that, starting from the boundary of a small sphere centered at 0, to hit the boundary of G before hitting an even smaller sphere centered at 0.

Lemma 6.3.3. *Let $N \subset \partial G$ be closed. Then*

$$\lim_{\rho\to 0} \limsup_{\epsilon\to 0} \epsilon \log \sup_{|x|=2\rho} \mathbb{P}(X^{\epsilon,x}_{\sigma_\rho^x} \in N) \le -\inf_{z\in N} V(0,z).$$

Proof. (of Lemma 6.3.3) For $\delta > 0$ fixed let $V_N = \min\{(\inf_{z\in N} V(0,z) - \delta), \frac{1}{\delta}\}$. Note first that by definition for $x, y, z \in \mathbf{R}^d$

$$V(x,z) \le V(x,y) + V(y,z).$$

Therefore, using Lemma 6.1.2 (i) to estimate the last term in the second expression for ρ_0 small enough and $0 < \rho < \rho_0$

$$\inf_{z\in N, |y|=2\rho} V(y,z) \geq \inf_{z\in N} V(0,z) - \sup_{|y|=2\rho} V(0,y) \geq V_N.$$

Use Lemma 6.2.2 to choose $T > 0$ such that for any $0 < \rho < \rho_0$

$$\limsup_{\epsilon\to 0} \epsilon \log \sup_{|y|=2\rho} \mathbb{P}(\sigma_\rho^y > T) < -V_N.$$

Consider the closed set of $\mathcal{C}([0,T])$

$$\Phi = \{\phi \in \mathcal{C}([0,T]) : \phi(t) \in N \text{ for some } t \in [0,T]\}.$$

Then by Corollary 5.2.7 for $0 < \rho < \rho_0$

$$\begin{aligned} \limsup_{\epsilon\to 0} \epsilon \log \sup_{|y|=2\rho} \mathbb{P}(X^{\epsilon,y} \in \Phi) &\leq -\inf_{|y|=2\rho} \inf_{\phi\in\Phi} I_{y,T}(\phi) \\ &\leq -\inf_{|y|=2\rho, z\in N} V(y,z)) \leq -V_N. \end{aligned}$$

We may summarize by stating that for $0 < \rho < \rho_0$

$$\begin{aligned} \limsup_{\epsilon\to 0} \epsilon \log \sup_{|x|=2\rho} \mathbb{P}(X^{\epsilon,x}_{\sigma^x_\rho} \in N) &\leq \limsup_{\epsilon\to 0} \epsilon \log[\sup_{|y|=2\rho} \mathbb{P}(\sigma^y_\rho > T) + \sup_{|y|=2\rho} \mathbb{P}(X^{\epsilon,y} \in \Phi)] \\ &\leq -V_N. \end{aligned}$$

Hence the claimed inequality follows. □

4. Let now $\overline{V} > 0, \delta > 0, x \in G$. **We show:**

$$\lim_{\epsilon\to 0} \mathbb{P}(\tau^{\epsilon,x} \leq e^{\frac{\overline{V}-\delta}{\epsilon}}) = 0.$$

To do this, choose $\rho > 0$ such that $\overline{K}_{2\rho}(0) \subset G$. Define recursively for $x \in G, m \geq 0$

$$\begin{aligned} \vartheta_0^x &= 0, \\ \tau_m^x &= \inf\{t \geq \vartheta_m^x : X_t^{\epsilon,x} \in \overline{K}_\rho(0) \cup \partial G\}, \\ \vartheta_{m+1}^x &= \begin{cases} \infty, & X^{\epsilon,x}_{\tau^x_m} \in \partial G, \\ \inf\{t \geq \tau^x_m : |X^{\epsilon,x}_t| = 2\rho\}, & |X^{\epsilon,x}_{\tau^x_m}| = \rho. \end{cases} \end{aligned}$$

Then $(X^{\epsilon,x}_{\tau^x_m})_{m\geq 0}$ is a Markov chain, where we use the convention $X^{\epsilon,x}_{\tau^x_m} = X^{\epsilon,x}_{\tau^{\epsilon,x}}$, if $\tau^x_m = \infty$. Fix $\delta > 0$. Using Lemma 6.3.3, applied to $N = \partial G$, choose $\rho_0 > 0$ such that for $0 < \rho < \rho_0$

$$\limsup_{\epsilon\to 0} \epsilon \log \sup_{|x|=2\rho} \mathbb{P}(X^{\epsilon,x}_{\sigma^x_\rho} \in \partial G) \leq -\overline{V} + \frac{\delta}{2}.$$

Now also fix $\rho < \rho_0$. Moreover, choose $T_0 = T(\overline{V}, \rho)$ according to Lemma 6.3.2. Then there exists $\epsilon_0 > 0$ such that for any $0 \leq \epsilon < \epsilon_0, m \geq 1$ by the strong Markov property

$$\sup_{x\in G} \mathbb{P}(\tau^{\epsilon,x} = \tau^x_m) \leq \sup_{|y|=2\rho} \mathbb{P}(X^{\epsilon,y}_{\sigma^y_\rho} \in \partial G) \leq e^{-\frac{\overline{V}-\frac{\delta}{2}}{\epsilon}},$$

and also

$$\sup_{x\in G} \mathbb{P}(\vartheta_m^x - \tau_{m-1}^x \leq T_0) \leq \sup_{x\in G} \mathbb{P}(\sup_{0\leq t\leq T_0} |X^{\epsilon,x} - x| \geq \rho) \leq e^{-\frac{\overline{V}-\frac{\delta}{2}}{\epsilon}}.$$

Now let $k \in \mathbb{N}$. Then for $x \in G$

$$\{\tau^{\epsilon,x} \leq kT_0\} \subset \{\tau^{\epsilon,x} = \tau_0^x\} \cup \cup_{m=1}^k [\{\tau^{\epsilon,x} = \tau_m^x\} \cup \{\vartheta_m^x - \tau_{m-1}^x \leq T_0\}].$$

Hence for $k \in \mathbb{N}, x \in G$

$$\begin{aligned}\mathbb{P}(\tau^{\epsilon,x} \leq kT_0) &\leq \mathbb{P}(\tau^{\epsilon,x} = \tau_0^x) + \sum_{m=1}^k [\mathbb{P}(\tau^{\epsilon,x} = \tau_m^x) + \mathbb{P}(\vartheta_m^x - \tau_{m-1}^x \leq T_0)] \\ &\leq \mathbb{P}(\tau^{\epsilon,x} = \tau_0^x) + 2ke^{-\frac{\overline{V}-\frac{\delta}{2}}{\epsilon}}.\end{aligned}$$

Now take

$$k = [\frac{1}{T_0} e^{\frac{\overline{V}-\delta}{\epsilon}}] + 1.$$

Then our estimate further yields for $x \in G$ and with the help of Lemma 6.3.1

$$\begin{aligned}\mathbb{P}(\tau^{\epsilon,x} \leq e^{\frac{\overline{V}-\delta}{\epsilon}}) &\leq \mathbb{P}(\tau^{\epsilon,x} \leq kT_0) \\ &\leq \mathbb{P}(X_{\sigma_\rho^x}^{\epsilon,x} \notin \overline{K}_\rho(0)) + \frac{4}{T_0} e^{-\frac{\delta}{2\epsilon}} \to 0\end{aligned}$$

as $\epsilon \to 0$. The lower bound for $\mathbb{E}(\tau^{\epsilon,x})$ now follows from Chebyshev's inequality.

5. We finally have to treat the case $\overline{V} = 0$. Let $\delta > 0, x \in G$. Choose $\rho > 0$ such that $\overline{K}_{2\rho}(0) \subset G$. Further let $c > 0$. Then Lemma 6.3.1 and Lemma 6.3.2 combined with the Markov property allow us to choose $\epsilon_0 > 0$ such that for $0 < \epsilon < \epsilon_0$

$$\mathbb{P}(\tau^{\epsilon,x} > e^{-\frac{\delta}{\epsilon}}) \geq \mathbb{P}(X_{\sigma_\rho^x}^{\epsilon,x} \in \overline{K}_\rho(0)) \cdot \inf_{|y|\leq\rho} \mathbb{P}(\sup_{0\leq t\leq T(c,\rho)} |X_t^{\epsilon,y} - y| \leq \rho) \to 1$$

as $\epsilon \to 0$. This completes the proof of our main result.

Bibliography

[1] R. Azencott. Grandes déviations et applications. In *Ecole d'ete de probabilites de Saint-Flour VIII-1978, Lect. Notes Math. 774, 1-176.* 1980.

[2] P. Baldi and M. Chaleyat-Maurel. An extension of Ventsel-Freidlin estimates. In *Stochastic analysis and related topics, Proc. Workshop, Silivri/Turk., Lect. Notes Math. 1316, 305-327.* 1988.

[3] P. Baldi and B. Roynette. Some exact equivalents for the Brownian motion in Hölder norm. *Probab. Theory Related Fields*, 93(4):457–484, 1992.

[4] Z. Ciesielski. On Haar functions and on the Schauder basis of the space $C_{\langle 0,1\rangle}$. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 7:227–232, 1959.

[5] A. Dembo and O. Zeitouni. *Large deviations techniques and applications.* Springer-Verlag, New York, second edition, 1998.

[6] S. Herrmann and B. Roynette. Boundedness and convergence of some self-attracting diffusions. *Mathematische Annalen*, 325(1):81–96, 2003.

[7] Pierre Priouret. Remarques sur les petites perturbations de systèmes dynamiques. In *Seminaire de probabilites XVI, Univ. Strasbourg 1980/81, Lect. Notes Math. 920, 184-200.* 1982.

[8] O. Raimond. Self-attracting diffusions: case of the constant interaction. *Probab. Theory Rel. Fields*, 107(2):177–196, 1997.

[9] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, third edition, 1999.

[10] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 1979.