Technische Universität Dresden
Fachrichtung Mathematik
Institut für Mathematische Stochastik

Large Deviations for Lévy(-Type) Processes

Diplomarbeit
zur Erlangung des ersten akademischen Grades

Diplommathematikerin

vorgelegt von

Name:	Kühn	Vorname:	Franziska
geboren am:	01.11.1989	in:	Dresden

Tag der Einreichung: 18. Februar 2014

Betreuer: Prof. Dr. rer. nat. René L. Schilling

Contents

Index of Notation **i**

1 Introduction **1**

2 Basic Definitions & Properties **3**

3 Gärtner-Ellis Approach **11**

4 Extensions of the Contraction Principle **18**

4.1 Exponential Approximations 18

4.2 Quasi-Continuity & Almost Compactness 22

5 Large Deviation Principles for Scaled Lévy Processes **25**

6 From Lévy to Lévy-Type Processes **43**

7 Large Deviations for Lévy-Driven SDEs **46**

7.1 Solutions as Markov Processes 47

7.2 Large Deviations by Exponential Approximations 55

8 Conclusion **65**

A Appendix **66**

Bibliography **73**

Index of Notation

Analysis

$\inf \varnothing$	$\inf \varnothing = +\infty$
$\lfloor x \rfloor$	integer part of x
$x^\top$	transpose of the vector x
$a \vee b$, $a \wedge b$	maximum, minimum
$\mathbb{1}_A$	indicator function of the set A
$f(t-)$	left limit $\lim_{s \uparrow t} f(s)$
càdlàg	finite left limits and right-continuous
càglàd	finite right limits and left-continuous
dom	domain
M^*	(topological) dual space of M
$\sigma(M, M^*)$	weak topology
$\Vert \cdot \Vert_\infty$	uniform norm
Γ	gamma function
$\vert\Pi\vert$	mesh size of the partition Π
f^*	Legendre transform, 11

Sets

A^c	complement of the set A
$A \mathbin{\dot\cup} B$	disjoint union of the sets A and B
$\operatorname{int} A$	interior of the set A
$\operatorname{cl} A$	closure of the set A
A^n	cartesian power of the set A
$B(x,r)$	open ball, centre x, radius r
$B[x,r]$	closed ball, centre x, radius r
$\Phi(r)$	sublevel set, 3

Probability/measure theory

$\sim$	distributed as
$\mathcal{B}(M)$	Borel σ-algebra on M
$\mathcal{A} \otimes \mathcal{B}$	product σ-algebra
$\mu \otimes \nu$	product measure
δ_x	Dirac measure at x
$\lambda\vert_B$	Lebesgue measure restricted to $B \in \mathcal{B}(\mathbb{R})$
$\mathbb{P}$ $\mathbb{E}_\mathbb{P}$	probability, expectation with respect to $\mathbb{P}$ ($\mathbb{E} = \mathbb{E}_\mathbb{P}$ for short)
$(\bar\Omega, \bar{\mathcal{A}}, \mathbb{P}^x)$	45
$\mathbb{V}$	variance
$\mathbb{E}(\cdot \mid \mathcal{F})$	conditional expectation with respect to a σ-algebra $\mathcal{F}$
a.s.	almost surely (with respect to $\mathbb{P}$)
$(B_t)_{t \geq 0}$	Brownian motion

Spaces of functions

$\mathcal{B}_b(M)$	space of borel-measurable, bounded functions $f : M \to \mathbb{R}$
$C(M)$	space of continuous functions $f : M \to \mathbb{R}$
$C_b(M)$	space of bounded, continuous functions $f : M \to \mathbb{R}$
$C_\infty(M)$	space of continuous functions $f : M \to \mathbb{R}$ satisfying $\lim_{\vert x\vert \to \infty} f(x) = 0$
$C_c^\infty(M)$	space of smooth functions $f : M \to \mathbb{R}$ with compact support
$D[0,1]$	space of càdlàg functions $f : [0,1] \to \mathbb{R}$
$\mathrm{BV}[0,1]$	space of functions $f : [0,1] \to \mathbb{R}$ of bounded variation
$AC[0,1]$	space of absolutely continuous functions $f : [0,1] \to \mathbb{R}$

References

(C1)-(C3)	13
(C4)	15
(C5)-(C7)	16
(L0)-(L2)	3
(L1'),(L2')	4
(S1),(S2)	47
(S3)	48

1 Introduction

Large deviation theory deals with the decay of probabilities of rare events on an exponential scale. Roughly speaking, large deviation estimates are of the form

$$\mathbb{P}(X_n \in B) \approx \exp(-n\, I(B)) \quad \text{for } n \text{ large},$$

where the *rate function* $I(B)$ is a measure for the asymptotic probability that the sequence $(X_n)_{n\in\mathbb{N}}$ of random variables attains values in the set B. In the last years large deviation theory has attracted more and more attention because of its variety of applications in the fields of financial mathematics, statistics, engineering, statistical physics, and chemistry. A thorough discussion of various applications can be found in the monograph [23] by Hollander. Our standard references for large deviation theory are Dembo-Zeitouni [10] and Feng-Kurtz [17].

In this work we focus on large deviation results for a certain type of stochastic processes; namely, we consider scaled *Lévy processes* and solutions of *stochastic differential equations (SDEs) driven by a Lévy process.* Both Lévy processes and solutions of SDEs driven by Lévy processes are an important subclass of *Lévy-type processes* which we briefly discuss in Chapter 6. In particular, we motivate that *symbols* are the natural generalization of *characteristic exponents* in the theory of Lévy-type processes.

Since (the proofs of) the main results are quite technical, the first part of the thesis aims at making the reader familar with some basic concepts in large deviation theory. In Chapter 2 we introduce the notion of *large deviation principle* and present fundamental results such as the *contraction principle* and *Varadhan's lemma.* Subsequently, we show that the *Legendre transform* plays – at least in a convex framework – an important role and formulate sufficient conditions for large deviations in terms of the *limiting logarithmic moment generating function.* Historically, this approach is due to Gärtner [20] and Ellis [15].
The results obtained will be used in Chapter 5 in order to establish large deviations for scaled Lévy processes. As an application, we derive two statements on the longtime behavior of Lévy processes: the *law of iterated logarithm* and the counterpart of *Strassen's law.*
Chapter 4 is concerned with some recent investigations on extensions of the contraction principle. They prove to be a crucial tool when considering large deviations for solutions

of SDEs; we will discuss this approach in Section 7.2. In Section 7.1 we follow the lines of Freidlin and Wentzell [18] and give a purely probabilistic proof of a large deviation result for SDEs driven by a Lévy process.

We emphasize that we do not attempt to state the results in their most general form since, otherwise, the proofs would become even more technical. However, we try to point the reader to notable generalizations.

To keep notation as simple as possible, we restrict ourselves to one-dimensional processes; most of the results can be proved in a similar fashion for multi-dimensional processes. Unless otherwise indicated, we assume that the processes are one-dimensional.

Finally, I would like to thank my advisor René Schilling not only for his support and valuable contributions but also for the interesting lectures (which succeeded in drawing my attention to the topic of stochastic processes) and the never-ending kindness to answer (most of) my never-ending questions. A special thanks goes to Johannes Huhn for his amiable companionship during the last year. Both to Johannes Huhn and Dirk Spitzner my thanks for reading substantial parts of my thesis and pointing out quite a few mistakes.

2

Basic Definitions & Properties

Throughout this chapter (M,d) denotes a metric space and $(\Omega,\mathcal{A},\mathbb{P})$ a probability space.

2.1 Definition Let $(\mu_\varepsilon)_{\varepsilon>0}$ be a family of probability measures on $(M,\mathcal{B}(M))$ and $\lambda:(0,\infty)\to\mathbb{R}$ be such that $\lambda(\varepsilon)\to\infty$ as $\varepsilon\to 0$. $(\mu_\varepsilon)_{\varepsilon>0}$ satisfies a *large deviation principle as $\varepsilon\to 0$ with rate function $I:M\to[0,\infty]$ and normalizing coefficient* λ if the following conditions hold.

(L0) I is lower semicontinuous, i. e. its sublevel sets $\Phi(r):=\{x\in M; I(x)\le r\}$, $r\ge 0$, are closed.

(L1) For any open set $A\subseteq M$,

$$\liminf_{\varepsilon\to 0}\frac{1}{\lambda(\varepsilon)}\log\mu_\varepsilon(A)\ge -\inf_{x\in A} I(x).$$

(L2) For any closed set $B\subseteq M$,

$$\limsup_{\varepsilon\to 0}\frac{1}{\lambda(\varepsilon)}\log\mu_\varepsilon(B)\le -\inf_{x\in B} I(x).$$

I is called a *good rate function* if its sublevel sets $\Phi(r)$, $r\ge 0$, are compact. A family of random variables $X_\varepsilon:(\Omega,\mathcal{A},\mathbb{P})\to(M,\mathcal{B}(M))$, $\varepsilon>0$, obeys a large deviation principle as $\varepsilon\to 0$ if the family of distributions $\mu_\varepsilon(\cdot):=\mathbb{P}(X_\varepsilon\in\cdot)$ satisfies a large deviation principle. Modifying (L1) and (L2) in an obvious fashion, we will also speak of a sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ satisfying a large deviation principle as $n\to\infty$ with rate function I and normalizing sequence $(\lambda_n)_{n\in\mathbb{N}}$.

Unless otherwise stated, we consider $\lambda(\varepsilon)=\varepsilon^{-1}$ and $\lambda_n=n$, respectively. Most of the results presented in Chapter 2 and 3 can be formulated for any normalizing coefficient and normalizing sequence, respectively.

Remarks (i). The rate function I is uniquely determined by the values $\mu_\varepsilon(B(x,\delta))$, $x\in M$, $\delta,\varepsilon>0$. In fact,

$$-I(x)=\lim_{\delta\to 0}\liminf_{\varepsilon\to 0}\varepsilon\log\mu_\varepsilon(B(x,\delta))=\lim_{\delta\to 0}\limsup_{\varepsilon\to 0}\varepsilon\log\mu_\varepsilon(B[x,\delta]).\tag{2.1}$$

This follows readily from the lower semicontinuity of I, cf. [18, Theorem 3.5]. On the other hand, if I satisfies (2.1), then $(\mu_\varepsilon)_{\varepsilon>0}$ obeys a *weak large deviation principle* with rate function I, i. e. (L0),(L1) hold and (L2) holds for all compact sets $B \subseteq M$, see e. g. [10, Theorem 4.1.11].

(ii). It is widely known, see e. g. [25, Theorem 1.2.11], that a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ converges in distribution to a random variable X if, and only if, for each open set A and closed set B,

$$\liminf_{n\to\infty} \mathbb{P}(X_n \in A) \geq \mathbb{P}(X \in A), \qquad \limsup_{n\to\infty} \mathbb{P}(X_n \in B) \leq \mathbb{P}(X \in B).$$

We note the analogy to (L1) and (L2). In this chapter, we will see that the so-called *exponential tightness* is the counterpart of tightness and state the analogue of Prokhorov's theorem.

2.2 Example Let $(B_t)_{t\in[0,1]}$ be a one-dimensional Brownian motion. The family of scaled processes $(\sqrt{\varepsilon}B)_{\varepsilon>0}$ satisfies a large deviation principle in $(C[0,1], \|\cdot\|_\infty)$ as $\varepsilon \to 0$ with good rate function

$$I(f) := \begin{cases} \frac{1}{2}\int_0^1 |f'(s)|^2 \, ds, & f \in AC[0,1], f(0) = 0, \\ \infty, & \text{otherwise}, \end{cases}$$

where $AC[0,1]$ denotes the set of absolutely continuous functions $f : [0,1] \to \mathbb{R}$. This large deviation principle is a special case of large deviation results we will encounter in Chapter 5. For a direct proof, based on the Cameron-Martin formula, we refer the reader to [38, Chapter 12].

The following lemma provides an alternative characterization of the large deviation lower bound (L1) and upper bound (L2). It is due to Freidlin and Wentzell [18].

2.3 Lemma *Under (L0), the conditions (L1), (L2) are equivalent to*

(L1') For any $\delta > 0$ and $x \in M$,

$$\liminf_{\varepsilon\to0} \varepsilon \log \mu_\varepsilon(B(x,\delta)) \geq -I(x).$$

(L2') For any $\delta > 0$ and $r > 0$,

$$\limsup_{\varepsilon\to0} \varepsilon \log \mu_\varepsilon(\{x \in M; d(x, \Phi(r)) \geq \delta\}) \leq -r.$$

Proof. Let $A \subseteq M$ open. Clearly, (L1) is equivalent to

$$\liminf_{\varepsilon\to0} \varepsilon \log \mu_\varepsilon(A) \geq -I(x) \qquad \text{for all} \quad x \in A.$$

Since A is open, there exists $\delta = \delta(x) > 0$ such that $B(x,\delta) \subseteq A$. If (L1') holds, then

$$\liminf_{\varepsilon\to0} \varepsilon \log \mu_\varepsilon(A) \geq \liminf_{\varepsilon\to0} \varepsilon \log \mu_\varepsilon(B(x,\delta)) \geq -I(x),$$

i.e. (L1) is satisfied. Since (L1) implies obviously (L1'), this proves (L1) $\Leftrightarrow$ (L1'). Now let $B \subseteq M$ be closed, and suppose that (L2') is satisfied. Set

$$r := \min\left\{\inf_{x\in B} I(x) - \gamma, \frac{1}{\gamma}\right\}$$

for $\gamma > 0$ sufficiently small. Since $\Phi(r)$ is compact and B closed, we have $\delta := d(B, \Phi(r)) > 0$. Consequently,

$$\mu_\varepsilon(B) \leq \mu_\varepsilon(\{x \in M; d(x, \Phi(r)) \geq \delta\}).$$

Therefore, by (L2'),

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(B) \leq -r \xrightarrow{\gamma\to 0} -\inf_{x\in B} I(x).$$

On the other hand, applying (L2) to the closed set $\{x \in M; d(x, \Phi(r)) \geq \delta\}$ yields (L2'). $\square$

Recall that a family $(\mu_\varepsilon)_{\varepsilon>0}$ of probability measures on M is called *tight* if for each $\delta > 0$ there exists a compact set $K \subseteq M$ such that $\sup_{\varepsilon>0} \mu_\varepsilon(K^c) \leq \delta$. The corresponding concept in large deviation theory is exponential tightness.

2.4 Definition A family $(\mu_\varepsilon)_{\varepsilon>0}$ of probability measures on M is called *exponentially tight* if for any $r > 0$ there exists a compact set $K_r \subseteq M$ such that

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(K_r^c) \leq -r. \tag{2.2}$$

A family of random variables $(X_\varepsilon)_{\varepsilon>0}$ is called exponentially tight if the family of distributions $(\mathbb{P}_{X_\varepsilon})_{\varepsilon>0}$ is exponentially tight.

Note that in general exponential tightness is not equivalent to goodness of the rate function. Exponential tightness implies the goodness of the rate function, but the converse is not true. For a counterexample we refer the reader to [13].

2.5 Lemma *Let $(\mu_\varepsilon)_{\varepsilon>0}$ be a sequence of probability measures on $(M, \mathcal{B}(M))$.*

(i). If μ_ε, $\varepsilon > 0$, is tight and $(\mu_\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle with a good rate function, then $(\mu_\varepsilon)_{\varepsilon>0}$ is exponentially tight.

(ii). If $(\mu_\varepsilon)_{\varepsilon>0}$ is exponentially tight and satisfies the large deviation lower bound (L1) with rate function I, then I is a good rate function.

(iii). If $(\mu_\varepsilon)_{\varepsilon>0}$ satisfies a weak large deviation principle with rate function I and $(\mu_\varepsilon)_{\varepsilon>0}$ is exponentially tight, then $(\mu_\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle with (good) rate function I.

Note that tightness holds automatically if (M, d) is a Polish space.

Proof. (i). cf. Lemma 5.2

(ii). For $r \geq 0$ let $K_{2r} \subseteq M$ be as in the definition of exponential tightness (2.2). By (L1) and (2.2),

$$-\inf_{x \in K_{2r}^c} I(x) \leq \liminf_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(K_{2r}^c) \leq \limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(K_{2r}^c) \leq -2r.$$

In particular, the sublevel set $\Phi(r) \subseteq K_{2r}$ is relatively compact. Since I is lower semicontinuous, hence $\Phi(r)$ closed, we conclude that $\Phi(r)$ is compact.

(iii). By definition it suffices to show (L2) for any closed set $B \subseteq M$. To this end, let K_r be as in the definition of exponential tightness. Obviously, $\mu_\varepsilon(B) \leq \mu_\varepsilon(B \cap K_r) + \mu_\varepsilon(K_r^c)$. Applying (L2) to the compact set $B \cap K_r$ yields

$$\begin{aligned}\limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(B) &\leq \max\left\{\limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(B \cap K_r^c), \limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(K_r^c)\right\} \\ &\leq \max\left\{-\inf_{x \in B \cap K_r} I(x), -r\right\} \leq \max\left\{-\inf_{x \in B} I(x), -r\right\}.\end{aligned}$$

The claim follows letting $r \to \infty$. □

The importance of exponential tightness is illustrated by the following analogue of Prokhorov's theorem. It was first formulated by Puhalskii [33]. Proofs can be found for instance in [17, Theorem 3.7] and [10, Lemma 4.1.23].

2.6 Theorem *Let $(\mu_n)_{n \in \mathbb{N}}$ be an exponentially tight sequence of probability measures on M. Then there exists a subsequence $(\mu_{n_k})_{k \in \mathbb{N}}$ which satisfies a large deviation principle with a good rate function.*

Theorem 2.6 provides an approach to prove large deviation results ("subsequence principle"):

(i). Show that the sequence $(\mu_n)_{n \in \mathbb{N}}$ is exponentially tight.

(ii). Choose an arbitrary subsequence $(\mu_n')_{n \in \mathbb{N}}$ and prove that the rate function I of the subsequence of $(\mu_n')_{n \in \mathbb{N}}$ satisfying a large deviation principle does not depend on the subsequence $(\mu_n')_{n \in \mathbb{N}}$.

In the theory of large deviations for stochastic processes, cf. Chapter 5 and 7, exponential tightness in the Skorohod space $D[0,1]$ is of particular interest. For (sufficient and necessary) conditions regarding exponential tightness in $D[0,1]$ we refer the reader to Feng [17, Theorem 4.1] and Puhalskii [34, Theorem 3.2.1].

2.7 Lemma *Let (M,d) be a metric space and $(\mu_\varepsilon)_{\varepsilon>0}$, $(\nu_\varepsilon)_{\varepsilon>0}$ exponentially tight families of probability measures on $(M, \mathcal{B}(M))$. Assume that $(\mu_\varepsilon)_{\varepsilon>0}$ and $(\nu_\varepsilon)_{\varepsilon>0}$ are exponentially tight and satisfy a large deviation principle in (M,d) as $\varepsilon \to 0$ with (good) rate function I and J, respectively. Then the family of product measures $(\mu_\varepsilon \otimes \nu_\varepsilon)_{\varepsilon>0}$ obeys a large deviation principle in $M \times M$ endowed with the metric*

$$d((x_1,y_1),(x_2,y_2)) := d(x_1,x_2) + d(y_1,y_2), \qquad x_1,x_2,y_1,y_2 \in M$$

as $\varepsilon \to 0$ *with (good) rate function* $K(x,y) := I(x) + J(y)$.

Proof. K is clearly lower semicontinuous, i. e. a rate function. Moreover, the relation $\Phi_K(r) \subseteq \Phi_I(r) \times \Phi_J(r)$ implies that K is a good rate function if I and J are good rate functions. Next we prove the large deviation lower bound (L1). Let $A \subseteq M \times M$ be open, and pick $(x,y) \in A$. There exists $\delta > 0$ such that $A \supseteq B(x,\delta) \times B(y,\delta)$. By the definition of the product measure and (L1) (for $(\mu_\varepsilon)_{\varepsilon>0}$ and $(\nu_\varepsilon)_{\varepsilon>0}$), we find

$$\begin{aligned}\liminf_{\varepsilon\to0} \varepsilon \log \mu_\varepsilon \otimes \nu_\varepsilon(A) &\geq \liminf_{\varepsilon\to0} \varepsilon \log \big(\mu_\varepsilon(B(x,\delta))\nu_\varepsilon(B(y,\delta))\big) \\ &\geq -(I(x) + J(y)) = -K(x,y).\end{aligned}$$

This proves the large deviation lower bound (L1). Now let $B \subseteq M \times M$ be compact and $\gamma > 0$. For any $(x,y) \in B$ we can choose $\delta = \delta((x,y)) > 0$ such that

$$I(x') \geq \min\left\{I(x) - \gamma, \gamma^{-1}\right\} =: I^\gamma(x) \quad \text{and} \quad J(y') \geq \min\left\{J(y) - \gamma, \gamma^{-1}\right\} =: J^\gamma(y)$$

for all $x' \in B[x,\delta]$, $y' \in B[y,\delta]$. Since B is compact, there exists a finite subcover $\bigcup_{j=1}^n B(x_j,\delta_j) \times B(y_j,\delta_j)$. Consequently,

$$\begin{aligned}\limsup_{\varepsilon\to0} \varepsilon \log \mu_\varepsilon \otimes \nu_\varepsilon(B) &\leq \limsup_{\varepsilon\to0} \varepsilon \log \left(n \max_{1\leq j\leq n} \left\{\mu_\varepsilon(B[x_j,\delta_j])\, \nu_\varepsilon(B[y_j,\delta_j])\right\}\right) \\ &\leq \max_{1\leq j\leq n}\left(-\inf_{x\in B[x_j,\delta_j]} I(x) - \inf_{y\in B[y_j,\delta_j]} J(y)\right) \\ &\leq -\min_{1\leq j\leq n} (I^\gamma(x_j) + J^\gamma(y_j)) \\ &\leq -\inf_{(x,y)\in B} (I^\gamma(x) + J^\gamma(y)).\end{aligned}$$

Letting $\gamma \to 0$ proves the large deviation upper bound (L2) for any compact set B. Moreover, the exponential tightness of $(\mu_\varepsilon)_{\varepsilon>0}$ and $(\nu_\varepsilon)_{\varepsilon>0}$ obviously entails the exponential tightness of $(\mu_\varepsilon \otimes \nu_\varepsilon)_{\varepsilon>0}$. It follows from Lemma 2.5(iii) that (L2) holds for each closed set $B \subseteq M \times M$. □

It is natural to ask how a large deviation principle is transformed under a continuous mapping. The result is known as contraction principle [10, Theorem 4.2.1]. It is the counterpart of the continuous mapping theorem.

2.8 Theorem (Contraction principle) *Let* (M_1,d_1), (M_2,d_2) *be metric spaces and* $f : M_1 \to M_2$ *be a continuous function. Suppose that a family* $(\mu_\varepsilon)_{\varepsilon>0}$ *of probability measures on* M_1 *satisfies a large deviation principle with good rate function* I. *Then the sequence of image measures* $(\nu_\varepsilon)_{\varepsilon>0}$, $\nu_\varepsilon := \mu_\varepsilon \circ f^{-1}$, *on* M_2 *obeys a large deviation principle with good rate function*

$$J(y) := \inf\{I(x); x \in M_1, y = f(x)\}.$$

Proof. Since I is lower semicontinuous, it attains its minimum on compact sets. This implies that for any $y \in M_2$, $J(y) < \infty$, there exists $x \in M_1$ such that $f(x) = y$ and $J(y) = I(x)$. Consequently,

$$\Phi_J(r) = \{y \in M_2; J(y) \leq r\} = f(\Phi_I(r)) \qquad \text{for all} \quad r \geq 0.$$

In particular, $\Phi_J(r)$ is compact, i. e. J is a good rate function. Now let $A \subseteq M_1$ open. Since f is continuous, hence $f^{-1}(A)$ open, we can apply the large deviation lower bound (L1) to $f^{-1}(A)$ and obtain

$$\liminf_{\varepsilon \to 0} \varepsilon \log \nu_\varepsilon(A) = \liminf_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(f^{-1}(A)) \geq -\inf_{x \in f^{-1}(A)} I(x) = -\inf_{y \in A} J(y).$$

The upper bound (L2) follows in the same way. □

The contraction principle can be extended beyond the continuous case, cf. Chapter 4.

The following example is an application of the contraction principle and gives a glimpse how one might prove a large deviation principle for solutions of stochastic differential equations using (an extended version of) the contraction principle. We will discuss this approach in Section 7.2.

2.9 Example Let $(B_t)_{t \geq 0}$ be a one-dimensional Brownian motion, $b : \mathbb{R} \to \mathbb{R}$ a bounded, Lipschitz continuous function (with Lipschitz constant $L > 0$) and $x \in \mathbb{R}$. Denote by $(X_t^\varepsilon)_{t \in [0,1]}$ the (unique) solution of the stochastic differential equation

$$dX_t^\varepsilon = b(X_t^\varepsilon)\, dt + \sqrt{\varepsilon}\, dB_t, \qquad X_0^\varepsilon = x.$$

The family $(X^\varepsilon)_{\varepsilon > 0}$ of stochastic processes satisfies a large deviation principle in $(C[0,1], \|\cdot\|_\infty)$ as $\varepsilon \to 0$ with good rate function

$$J(\psi) := \begin{cases} \frac{1}{2} \int_0^1 |\psi'(s) - b(\psi(s))|^2\, ds, & \psi \in AC[0,1], \psi(0) = x, \\ \infty, & \text{otherwise.} \end{cases}$$

Proof. The Lipschitz continuity of b entails the existence of a (unique) solution of the given SDE, see e. g. [38]. For $f \in C[0,1]$ we denote by Ff the solution of the integral equation

$$\psi(t) = x + \int_0^t b(\psi(s))\, ds + f(t), \qquad t \in [0,1].$$

Then,

$$|(Ff_1)(t) - (Ff_2)(t)| \leq \|f_1 - f_2\|_\infty + L \int_0^t |(Ff_1)(s) - (Ff_2)(s)|\, ds.$$

Applying Gronwall's lemma, see e. g. [38, Theorem A.43], we obtain

$$\|Ff_1 - Ff_2\|_\infty \leq C\, \|f_1 - f_2\|_\infty.$$

Consequently, F defines a continuous bijection on $(C[0,1], \|\cdot\|_\infty)$. From the contraction principle, Theorem 2.8, and Example 2.2 we conclude that $F(\sqrt{\varepsilon}B) = X^\varepsilon$ satisfies a large deviation principle as $\varepsilon \to 0$ with good rate function J,

$$J(\psi) = \frac{1}{2}\int_0^1 \left|\frac{d}{ds}(F^{-1}\psi)(s)\right|^2 ds = \frac{1}{2}\int_0^1 |\psi'(s) - b(\psi(s))|^2 \, ds$$

for $\psi \in AC[0,1], \psi(0) = x$. □

Another application of the contraction principle is the following corollary which allows us to transfer large deviation results from a topology to a coarser one.

2.10 Corollary *Let (M, d_1), (M, d_2) be metric spaces such that $d_1(x,y) \leq d_2(x,y)$ for all $x, y \in M$. If a family $(\mu_\varepsilon)_{\varepsilon>0}$ of probability measures on (M, d_2) satisfies a large deviation principle in (M, d_2) with good rate function I, then it obeys a large deviation principle in (M, d_1) with good rate function I.*

In general, the converse does not hold. The *inverse contraction principle* [10, Theorem 4.2.4] gives a sufficient condition under which this is true.

Finally, we characterize large deviations in terms of asymptotics of exponential integrals. As we will see in Chapter 3, Theorem 2.11 plays a key role in large deviation theory; it is a very useful tool in many applications to large deviation theory as well as in proving large deviation principles and identifying rate functions. We follow the presentation given in [17, Theorem 3.8], see also [10, Theorem 4.3.1,Theorem 4.4.2].

2.11 Theorem *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of M-valued random variables.*

(i). Varadhan's Lemma: If $(X_n)_{n\in\mathbb{N}}$ satisfies a large deviation principle with good rate function I, then:

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{nf(X_n)}) = \sup_{x\in M}(f(x) - I(x)) \qquad \text{for all } f \in C_b(M).$$

(ii). Bryc formula: Suppose that $(X_n)_{n\in\mathbb{N}}$ is exponentially tight and that the limit

$$\Lambda(f) := \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{nf(X_n)}) \tag{2.3}$$

exists for all $f \in C_b(M)$. Then $(X_n)_{n\in\mathbb{N}}$ satisfies a large deviation principle with good rate function

$$I(x) := \sup_{f\in C_b(M)} (f(x) - \Lambda(f)).$$

Proof. We prove only (i). For a proof of Bryc's formula we refer the reader to [17, Proposition 3.8] or [10, Theorem 4.2]. Let $f \in C_b(M)$, $x \in M$ and $\varepsilon > 0$. In abuse of notation, we write[1]

$$f(\varepsilon) := \min\{f(y); y \in B[x, \varepsilon]\}.$$

[1] Revised version: Reformulated.

Then $f(\varepsilon) \to f(x)$ as $\varepsilon \to 0$, and by Markov's inequality[2] we get

$$\mathbb{P}(X_n \in B[x,\varepsilon]) \le \mathbb{E}\left(e^{nf(X_n)-nf(\varepsilon)}\right).$$

Thus,

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_n \in B[x,\varepsilon]) \le -f(\varepsilon) + \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{nf(X_n)}) \\
\Rightarrow &\lim_{\varepsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_n \in B[x,\varepsilon]) \le -f(x) + \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{nf(X_n)}).
\end{aligned} \tag{2.4}$$

Combining (2.1) and (2.4) yields

$$\sup_{x\in M}(f(x) - I(x)) \le \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{nf(X_n)}).$$

It remains to show the upper bound. For each $x \in M$ we can choose $\delta_x > 0$ such that

$$\max_{y\in B[x,\delta_x]} f(y) \le f(x) + \varepsilon, \qquad \inf_{y\in B[x,\delta_x]} I(y) \ge \min\{I(x) - \varepsilon, \varepsilon^{-1}\} =: I^\varepsilon(x). \tag{2.5}$$

By assumption, I is a good rate function, and therefore there exists a finite subcover of $\Phi(r)$[3] for $r := 2\|f\|_\infty$, i.e. $\Phi(r)$[4] $\subseteq \bigcup_{i=1}^m B(x_i, \delta_{x_i})$. It follows from the large deviation upper bound (L2) that

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(X_n)} \mathbb{1}_{(\bigcup_{i=1}^m B(x_i,\delta_{x_i}))^c}(X_n)\big) \\
&\qquad \le \limsup_{n\to\infty} \frac{1}{n} \log \left(e^{n\|f\|_\infty} \mathbb{P}\left[X_n \in \left(\bigcup_{i=1}^m B(x_i,\delta_{x_i}) \right)^c \right] \right) \\
&\qquad \le \|f\|_\infty + \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(X_n \in \left(\bigcup_{i=1}^m B(x_i,\delta_{x_i}) \right)^c \right) \le -\|f\|_\infty.
\end{aligned}$$

This implies

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{nf(X_n)}) \le \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(X_n)} \mathbb{1}_{\bigcup_{i=1}^m B(x_i,\delta_{x_i})}(X_n)\big).$$

Since

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{nf(X_n)} \mathbb{1}_{\bigcup_{i=1}^m B(x_i,\delta_{x_i})}(X_n)) \\
&\qquad \overset{(2.5)}{\le} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left(\sum_{i=1}^m e^{n(f(x_i)+\varepsilon)} \mathbb{1}_{B(x_i,\delta_{x_i})}(X_n) \right) \\
&\qquad \le \max_{1\le i\le m} \left(f(x_i) + \varepsilon + \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_n \in B[x_i,\delta_{x_i}]) \right) \\
&\qquad \overset{(L2)}{\le} \max_{1\le i\le m} \left(f(x_i) + \varepsilon - \inf_{y\in B[x_i,\delta_{x_i}]} I(y) \right) \\
&\qquad \overset{(2.5)}{\le} \sup_{x\in M} (f(x) + \varepsilon - I^\varepsilon(x))
\end{aligned}$$

the claim follows letting $\varepsilon \to 0$. □

[2]We call $\mathbb{P}(|X| \ge \delta) \le f(\delta)^{-1}\mathbb{E}f(|X|)$ *Markov inequality* for any increasing function $f : [0,\infty) \to [0,\infty)$. Common choices are $f(x) = e^{\lambda x}$, $\lambda > 0$, and $f(x) = x^n$, $n \in \mathbb{N}$.

[3]Revised version: corrected misprint.

[4]Revised version: corrected misprint.

3

Gärtner-Ellis Approach

In this chapter we establish a large deviation lower and upper bound under assumptions on the limiting behavior of the logarithmic moment generating function. This approach is due to Gärtner [20] and Ellis [15] who proved the corresponding result for real-valued random variables, see e. g. [10, Section 2.3]. Starting from the definition of the Legendre transform we motivate that the Legendre transform of the limiting logarithmic moment generating function is a natural candidate for the rate function.

3.1 Definition Let (M,d) be a metric space and denote by M^* its dual space. The *Legendre transform* (or *Fenchel-Legendre transform*, *conjugate function*) of a function $f : M^* \to [-\infty, \infty]$ is defined by

$$f^*(x) := \sup_{\lambda \in M^*} \big(\langle \lambda, x\rangle - f(\lambda)\big), \qquad x \in M.$$

3.2 Definition Let (M,d) be a metric space and $(\mu_n)_{n\in\mathbb{N}}$ a family of probability measures on $(M, \mathcal{B}(M))$. We call

$$\Lambda_{\mu_n}(\lambda) := \log\Big(\int e^{\langle \lambda, x\rangle}\, d\mu_n(x)\Big), \qquad \lambda \in M^*$$

the *logarithmic moment generating function of* μ_n and define

$$\bar{\Lambda}(\lambda) := \limsup_{n\to\infty} \frac{1}{n}\Lambda_{\mu_n}(n\lambda), \qquad \lambda \in M^*.$$

Whenever the limit exists, we write $\Lambda(\lambda)$. If X is a random variable, we denote by Λ_X the logarithmic moment generating function of the distribution $\mathbb{P}(X \in \cdot)$.

3.3 Lemma *(i).* $\bar{\Lambda} : M^* \to (-\infty, \infty]$ *is convex.*

(ii). $\bar{\Lambda}^*$ *is a convex lower semicontinuous non-negative function on* M.

Proof. (i). By the subadditivity of limes superior, it suffices to show that the logarithmic moment generating function Λ_{μ_n}, $n \in \mathbb{N}$, is convex. Pick $\lambda_1, \lambda_2 \in M^*$ and $\alpha \in (0,1)$. Applying Hölder's inequality for the conjugate exponents $1/\alpha$ and $1/(1-\alpha)$, we find

$$\begin{aligned}\Lambda_{\mu_n}(\alpha\lambda_1 + (1-\alpha)\lambda_2) &\le \log\bigg[\Big(\int e^{\langle \lambda_1, x\rangle}\, d\mu_n(x)\Big)^{\alpha}\Big(\int e^{\langle \lambda_2, x\rangle}\, d\mu_n(x)\Big)^{1-\alpha}\bigg]\\ &= \alpha\Lambda_{\mu_n}(\lambda_1) + (1-\alpha)\Lambda_{\mu_n}(\lambda_2).\end{aligned}$$

(ii). By definition, $\Lambda_{\mu_n}(0) = 0$, thus $\bar{\Lambda}^* \geq 0$. The convexity follows obviously from the definition of the Legendre transform. Observe that $M \ni x \mapsto \langle \lambda, x \rangle - \bar{\Lambda}(\lambda) \in \mathbb{R}$ is, for fixed $\lambda \in M^*$, a continuous mapping. Being the supremum of a family of continuous functions, $\bar{\Lambda}^*$ is lower semicontinuous. □

The following theorem [10, Theorem 4.5.10] shows that $\bar{\Lambda}^*$ is a good candidate for the rate function, at least if the rate function is convex.

3.4 Theorem *Let (M,d) be a metric space. Suppose that a family $(\mu_n)_{n\in\mathbb{N}}$ of probability measures on $(M, \mathcal{B}(M))$ satisfies a large deviation principle with good rate function I and*

$$\bar{\Lambda}(\lambda) = \limsup_{n\to\infty} \frac{1}{n} \Lambda_{\mu_n}(n\lambda) < \infty \qquad \textit{for all } \lambda \in M^*. \tag{3.1}$$

(i). The limit

$$\Lambda(\lambda) = \lim_{n\to\infty} \frac{1}{n} \Lambda_{\mu_n}(n\lambda)$$

exists, is finite and

$$\Lambda(\lambda) = \sup_{x\in M} \big(\langle \lambda, x \rangle - I(x) \big) =: I^*(\lambda). \tag{3.2}$$

(ii). Λ^* is the affine regularization *of I, i. e.*

$$\Lambda^*(x) = \sup\{f(x); f \textit{ convex}, f \leq I\}, \qquad x \in M.$$

In particular, if I is convex, then

$$I(x) = \Lambda^*(x) = \sup_{\lambda\in M^*} \big(\langle \lambda, x \rangle - \Lambda(\lambda) \big).$$

Proof. (i). The claim follows from Varadhan's Lemma, cf. Theorem 2.11. Note that Theorem 2.11 is applicable since we can drop the assumption on the boundedness and replace it by the tail condition (3.1), see [10, Theorem 4.3.1].

(ii). By the definition of the Legendre transform and (3.2), we have

$$\Lambda^*(x) = \sup_{\lambda\in M^*} \big(\langle \lambda, x \rangle - \Lambda(\lambda) \big) \leq \sup_{\lambda\in M^*} \big(\langle \lambda, x \rangle - (\langle \lambda, x \rangle - I(x)) \big) = I(x)$$

By Lemma 3.3, Λ^* is convex and therefore,

$$\Lambda^*(x) \leq \sup\{f(x); f \text{ convex}, f \leq I\}.$$

Clearly, $g \leq h$ implies $g^* \geq h^*$ for any two functions g, h. Applying this twice, we get $f = (f^*)^* \leq (I^*)^* \overset{(3.2)}{=} \Lambda^*$ for any $f \leq I$ convex. If I is convex, then by duality lemma, cf. Theorem A.3, $(I^*)^* = I$. □

The remaining part of this chapter is devoted to the formulation of sufficient conditions for large deviation lower and upper bounds in terms of the limiting behavior of the logarithmic moment generating function Λ. They are tailored to the application in Chapter 5. We start with the large deviation upper bound. The result is based on [8, Theorem 2.1].

3.5 Theorem *Let (M,d) be a metric space and $\mathcal{B}$ a σ-algebra on M such that*

(C1) The vector space operations are $\mathcal{B}$-measurable.

(C2) $\mathcal{B}$ contains all compact sets.

For $n \in \mathbb{N}$ let $X_n : (\Omega, \mathcal{A}, \mathbb{P}) \to (M, \mathcal{B})$ be random variables and $F \subseteq M^*$ such that

(C3) $\langle \lambda, \cdot \rangle$ is $\mathcal{B}$-measurable for all $\lambda \in F$.

Define

$$\tilde{\Lambda}(\lambda) := \begin{cases} \bar{\Lambda}(\lambda), & \lambda \in F, \\ \infty, & \lambda \in M^* \backslash F \end{cases}$$

where

$$\bar{\Lambda}(\lambda) = \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E} e^{n\langle \lambda, X_n \rangle}.$$

Then

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_n \in B) \le - \inf_{x\in B} \tilde{\Lambda}^*(x) \tag{3.3}$$

for any compact set $B \subseteq M$. If $(X_n)_{n\in\mathbb{N}}$ is exponentially tight, (3.3) *holds for all closed sets $B \in \mathcal{B}$.*

Proof. We adapt the proof given in [10, Theorem 4.5.3]. Without loss of generality we may assume $F \neq \varnothing$. Let $B \subseteq M$ be compact and fix $\delta > 0$. By definition, we have

$$\tilde{\Lambda}^*(x) = \sup_{\lambda\in F} \big(\langle \lambda, x \rangle - \bar{\Lambda}(\lambda) \big) \qquad \text{for all} \quad x \in B.$$

Therefore, there exists $\lambda_x \in F$ such that

$$\langle \lambda_x, x \rangle - \bar{\Lambda}(\lambda_x) \ge \min\left\{ \tilde{\Lambda}^*(x) - \delta, \frac{1}{\delta} \right\} =: I^\delta(x).$$

Since λ_x is continuous at x, we can choose $r_x > 0$ such that

$$\forall y \in B[x, r_x]: \quad |\langle \lambda_x, y \rangle - \langle \lambda_x, x \rangle| \le \delta. \tag{3.4}$$

Note that $B[x,r_x] \cap B$ is compact, hence $B[x,r_x] \cap B \in \mathcal{B}$. Markov's inequality yields

$$\begin{aligned} \mathbb{P}(X_n \in B[x,r_x] \cap B) &\le \mathbb{E} \exp\left(\langle \theta, X_n \rangle - \inf_{y\in B[x,r_x]} \langle \theta, y \rangle \right) \\ &= \exp\left(- \inf_{y\in B[x,r_x]} (\langle \theta, y \rangle - \langle \theta, x \rangle) \right) \mathbb{E} \exp\left(\langle \theta, X_n \rangle - \langle \theta, x \rangle \right) \end{aligned}$$

for any $\theta \in F$. In particular, for $\theta = n\lambda_x$,

$$\frac{1}{n} \log \mathbb{P}(X_n \in B[x,r_x] \cap B) \overset{(3.4)}{\le} \delta - \left(\langle \lambda_x, x \rangle - \frac{1}{n} \Lambda_{X_n}(n\lambda_x) \right).$$

By the compactness of B, there exists a finite subcover $B \subseteq \bigcup_{i=1}^m B(x_i, r_{x_i})$. From

$$\frac{1}{n} \log \mathbb{P}(X_n \in B) \le \frac{1}{n} \log \left(m \max_{1\le i\le m} \mathbb{P}(X_n \in B[x_i, r_{x_i}] \cap B) \right)$$

we conclude

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_n \in B) &\le \delta - \min_{1\le i\le m} \big(\langle \lambda_{x_i}, x_i\rangle - \bar{\Lambda}(\lambda_{x_i})\big)\\ &\le \delta - \min_{1\le i\le m} I^\delta(x_i)\\ &\le \delta - \inf_{x\in B} I^\delta(x) \xrightarrow{\delta\to 0} -\inf_{x\in B} \tilde{\Lambda}^*(x).\end{aligned}$$

This proves (3.3) for $B \subseteq M$ compact. If $(X_n)_{n\in\mathbb{N}}$ is exponentially tight, then (L2) holds for each closed set B by Lemma 2.5(iii). □

Remark For any function J satisfying $\tilde{\Lambda} \le J$ we may substitute the Legendre transform J^* for $\tilde{\Lambda}^*$ in (3.3). This turns out to be quite useful if $\tilde{\Lambda}$ is difficult to compute but easy to bound.

The proof of the large deviation lower bound (L1) requires more effort; we will have to strengthen the assumptions of Theorem 3.5. As a first step we quickly establish some basic properties under an absolute change of measure.

3.6 Lemma *Let (M,d) be a metric space, $(M,\mathcal{B})$ a measurable space and $F \subseteq M^*$ such that (C1)-(C3) are satisfied. For $n \in \mathbb{N}$ let $X_n : (\Omega,\mathcal{A},\mathbb{P}) \to (M,\mathcal{B})$ be random variables. Assume that*

$$\Lambda(\lambda) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E} e^{n\langle\lambda, X_n\rangle}$$

exists for each $\lambda \in F$. Define probability measures $\mathbb{Q}_n^\lambda$ on $(\Omega,\mathcal{A})$ by

$$d\mathbb{Q}_n^\lambda := \frac{1}{\mathbb{E}e^{n\langle\lambda,X_n\rangle}} e^{n\langle\lambda,X_n\rangle}\, d\mathbb{P}, \qquad \lambda \in F.$$

(i). The limiting logarithmic moment generating function

$$\Lambda_\lambda(\eta) := \lim_{n\to\infty} \frac{1}{n} \log\left(\int e^{n\langle\eta,X_n\rangle}\, d\mathbb{Q}_n^\lambda\right)$$

satisfies

$$\Lambda_\lambda(\eta) = \Lambda(\eta+\lambda) - \Lambda(\lambda).$$

(ii). $\tilde{\Lambda}_\lambda^*(x) = \tilde{\Lambda}^*(x) - \langle\lambda, x\rangle + \Lambda(\lambda)$ for $x \in M$.

(iii). If $(X_n)_{n\in\mathbb{N}}$ is exponentially tight, then the sequence of distributions $\mathbb{Q}_n^\lambda(X_n \in \cdot)$, $n \in \mathbb{N}$, is exponentially tight.

Proof. By the definition of $\mathbb{Q}_n^\lambda$,

$$\Lambda_\lambda(\eta) = \lim_{n\to\infty} \frac{1}{n} \log\left(\frac{1}{\mathbb{E}e^{n\langle\lambda,X_n\rangle}} \mathbb{E}e^{n\langle\eta+\lambda,X_n\rangle}\right) = -\Lambda(\lambda) + \Lambda(\lambda+\eta).$$

This proves (i). Since $\tilde{\Lambda}(\eta) = \tilde{\Lambda}_\lambda(\eta) = \infty$ for $\eta \in M^*\backslash F$, we have

$$\begin{aligned}\tilde{\Lambda}_\lambda^*(x) = \sup_{\eta\in F}\big(\langle\eta,x\rangle - \Lambda_\lambda(\eta)\big) &\overset{\text{(i)}}{=} \sup_{\eta\in F}\big(\langle\eta+\lambda,x\rangle - \langle\lambda,x\rangle - \Lambda(\lambda+\eta) + \Lambda(\lambda)\big)\\ &= \sup_{\eta\in F}\big(\langle\eta,x\rangle - \Lambda(\eta)\big) - \langle\lambda,x\rangle + \Lambda(\lambda)\\ &= \tilde{\Lambda}^*(x) - \langle\lambda,x\rangle + \Lambda(\lambda).\end{aligned}$$

It remains to show (iii). If $(X_n)_{n\in\mathbb{N}}$ is exponentially tight, then there exists for each $r \geq 0$ a compact set $K_r \subseteq M$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_n \in K_r^c) \leq -r.$$

By Hölder's inequality, we have

$$\mathbb{Q}_n^\lambda(X_n \in K_r^c) \leq \frac{1}{\mathbb{E}e^{n\langle\lambda,X_n\rangle}} \sqrt{\mathbb{P}(X_n \in K_r^c)} \sqrt{\mathbb{E}e^{2n\langle\lambda,X_n\rangle}}.$$

Consequently,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{Q}_n^\lambda(X_n \in K_r^c) \leq -\Lambda(\lambda) - \frac{r}{2} + \frac{1}{2}\Lambda(2\lambda) \xrightarrow{r\to\infty} -\infty. \qquad \square$$

Let us recall that a set $F \subseteq M^*$ *separates points in* M if

$$\{\forall \lambda \in F : \langle\lambda, x\rangle = \langle\lambda, y\rangle\} \Longrightarrow x = y$$

for any $x, y \in M$. The next lemma shows that the sequence of distributions $(\mathbb{Q}_n^\lambda(X_n \in \cdot))_{n\in\mathbb{N}}$ concentrates mass at some point $x(\lambda) \in M$ as $n \to \infty$, namely, the Gâteaux derivative of Λ at λ. The following two results are based on [8, Theorem 2.2].

3.7 Lemma *Let (M, d) be a metric space, $(M, \mathcal{B})$ be a measurable space and $F \subseteq M^*$ such that (C1)-(C3) hold and*

(C4) F separates points in M.

Let $(X_n)_{n\in\mathbb{N}}$ be an exponentially tight sequence of $\mathcal{B}$-measurable random variables. Suppose that the mapping

$$F \ni \lambda \mapsto \Lambda(\lambda) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}e^{n\langle\lambda,X_n\rangle} \in \mathbb{R}$$

*is M-*Gâteaux differentiable*, i. e. for any $\lambda \in F$ there exists $D_\lambda \in M$ such that*

$$\langle\eta, D_\lambda\rangle = \lim_{t\to 0} \frac{\Lambda(\lambda + t\eta) - \Lambda(\lambda)}{t} \qquad \textit{for all } \eta \in F.$$

Then

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{Q}_n^\lambda(X_n \in A^c) < 0$$

holds for any open set $A \in \mathcal{B}$ such that $D_\lambda \in A$.

Proof. Pick $\lambda \in F$. Let $A \in \mathcal{B}$ be open such that $D_\lambda \in A$. By Lemma 3.6(iii) it suffices to show

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{Q}_n^\lambda(X_n \in A^c \cap K) < 0$$

for any compact set $K \subseteq M$. By Theorem 3.5, we have

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{Q}_n^\lambda(X_n \in A^c \cap K) \leq -\inf_{x\in A^c\cap K} \tilde{\Lambda}_\lambda^*(x),$$

i. e. the claim follows if

$$c := \inf_{x\in A^c\cap K} \tilde{\Lambda}_\lambda^*(x) > 0.$$

Suppose that $c = 0$. Since $\tilde{\Lambda}^*_\lambda$ is lower semicontinuous, cf. Lemma 3.3, there exists $x_0 \in A^c \cap K$ such that $\tilde{\Lambda}^*_\lambda(x_0) = 0$. Consequently, by Lemma 3.6(ii),

$$\langle \lambda, x_0 \rangle - \Lambda(\lambda) = \tilde{\Lambda}^*(x_0) \geq \langle \lambda + t\eta, x_0 \rangle - \Lambda(\lambda + t\eta)$$

for any $\eta \in F$, $t > 0$. Thus,

$$\langle \eta, D_\lambda \rangle = \lim_{t \to 0} \frac{\Lambda(\lambda + t\eta) - \Lambda(\lambda)}{t} \geq \langle \eta, x_0 \rangle.$$

Since the same calculation holds for $-\eta$ and F separates points in M, we conclude $x_0 = D_\lambda$. Obviously, $D_\lambda = x_0 \in A^c \cap K$ contradicts $D_\lambda \in A$. □

Before we finally state the large deviation lower bound, we introduce the notion of subdifferentiability.

3.8 Definition Let (M, d) be a metric space and $f : M \to [-\infty, \infty]$ be a function.

(i). f is called *proper* $:\Leftrightarrow \forall y \in M : f(y) > -\infty, \exists x \in M : f(x) < \infty$.

(ii). If f is convex and proper, the *subdifferential of f at x* is defined by

$$\partial f(x) := \{\lambda \in M^*; \forall y \in M : f(y) - f(x) \geq \langle \lambda, y - x \rangle\}.$$

3.9 Theorem *Let $(M, \|\cdot\|)$ be a Banach space, $\mathcal{B}$ a σ-algebra on M, $F \subseteq M^*$ a linear subspace and $M_0 \subseteq M$ a closed subspace. Suppose that (C1)-(C4) are satisfied as well as*

(C5) $M_0^* = F|_{M_0}$

(C6) M_0 *separates points in* F.

(C7) $(M, \mathcal{B}) \ni x \mapsto \|x\| \in (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *is measurable.*

Let $(X_n)_{n \in \mathbb{N}}$ be an exponentially tight sequence of $(M, \mathcal{B})$-measurable random variables. If

$$\Lambda(\lambda) = \lim_{n \to \infty} \frac{1}{n} \log \mathbb{E} e^{n\langle \lambda, X_n \rangle}, \qquad \lambda \in F$$

is M_0-Gâteaux differentiable and $\operatorname{dom} \tilde{\Lambda}^* \subseteq M_0$*, then the large deviation lower bound holds for every open set $A \in \mathcal{B}$:*

$$\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}(X_n \in A) \geq -\inf_{x \in A} \tilde{\Lambda}^*(x).$$

Proof. Let $A \in \mathcal{B}$ be open. Without loss of generality we may assume that there exists $x \in A$ such that $\tilde{\Lambda}^*(x) < \infty$. By assumption, $\operatorname{dom} \tilde{\Lambda}^* \subseteq M_0$, hence $x \in M_0$. Choose $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq A$. From (the proof of) Lemma 3.3 we know that $\tilde{\Lambda}^*|_{M_0}$ is a lower semicontinuous, convex, proper function. By Theorem A.5, there exists $y \in M_0$ such that $\partial \tilde{\Lambda}^*(y) \neq \varnothing$ and

$$|x - y|^1 < \varepsilon, \qquad \|\tilde{\Lambda}^*(x) - \tilde{\Lambda}^*(y)\| < \varepsilon. \tag{3.5}$$

Pick $\lambda \in \partial\tilde{\Lambda}^*(y)$. Note that the function $\tilde{\Lambda}|_F = \Lambda|_F$ is convex and lower semicontinuous. Since $M_0^* = F|_{M_0}$, an application of the duality lemma A.3 yields $(\tilde{\Lambda}^*|_{M_0})^* = \tilde{\Lambda}$ and therefore, by Lemma A.4(i), $y \in \partial\tilde{\Lambda}(\lambda)$. On the other hand, the M_0-Gâteaux differentiability implies $\partial\tilde{\Lambda}(\lambda) = \{D_\lambda\}$. Thus, $D_\lambda = y$. Set

$$U := B(x,\varepsilon) \cap \{z \in M; |\langle \lambda, D_\lambda - z\rangle| < \delta\}$$

for $\delta > 0$. By (C3) and (C7), we have $U \in \mathcal{B}$. Obviously, $D_\lambda \in U$ and

$$\begin{aligned}\mathbb{P}(X_n \in A) \geq \mathbb{P}(X_n \in U) &= \mathbb{E}e^{n\langle \lambda, X_n\rangle} \int \mathbb{1}_U(X_n) e^{-n\langle X_n, \lambda\rangle}\, d\mathbb{Q}_n^\lambda \\ &\geq \mathbb{E}e^{n\langle \lambda, X_n\rangle} e^{-n(\langle \lambda, D_\lambda\rangle + \delta)} \mathbb{Q}_n^\lambda(X_n \in U).\end{aligned}$$

where $\mathbb{Q}_n^\lambda$ is defined as in Lemma 3.6. Applying Lemma 3.7, we get

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_n \in A) &\geq -(\langle \lambda, D_\lambda\rangle - \Lambda(\lambda)) + \delta \\ &\geq -\sup_{\mu\in F} \big(\langle \mu, D_\lambda\rangle - \Lambda(\mu)\big) + \delta \\ &= -\tilde{\Lambda}^*(D_\lambda) + \delta = -\tilde{\Lambda}^*(y) + \delta.\end{aligned}$$

Finally, by (3.5),

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_n \in A) \geq -\tilde{\Lambda}^*(x) + \delta + \varepsilon.$$

Since $\varepsilon, \delta > 0$ are arbitrary, we are done. □

Remark If $(X_n)_{n\in\mathbb{N}}$ meets the assumtions of Theorem 3.5 and Theorem 3.9, then $(X_n)_{n\in\mathbb{N}}$ satisfies the large deviation lower bound (L1) and upper bound (L2) for any sets $A, B \in \mathcal{B}$ with convex rate function $\tilde{\Lambda}^*$. We say that $(X_n)_{n\in\mathbb{N}}$ satisfies a *large deviation principle in* (M,d) *with respect to* $\mathcal{B}$. Note that $\tilde{\Lambda}^*$ is even a good rate function, cf. Lemma 2.5.

3.10 Corollary (Abstract Gärtner-Ellis theorem) *Let* $(M, \|\cdot\|)$ *be a Banach space and* $(X_n)_{n\in\mathbb{N}}$ *an exponentially tight sequence of* $\mathcal{B}(M)$*-measurable random variables. If*

$$\Lambda(\lambda) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}e^{n\langle \lambda, X_n\rangle}$$

is M*-Gâteaux differentiable for each* $\lambda \in M^*$*, then* $(X_n)_{n\in\mathbb{N}}$ *satisfies a large deviation principle with convex good rate function* Λ^**.*

Proof. By the Hahn-Banach theorem, $F := M^*$ separates points in M. Consequently, the claim follows by applying Theorem 3.9 and Theorem 3.5. □

Remark In fact, Corollary 3.10 holds for any locally convex Hausdorff topological space, see e. g. [10, Theorem 4.6.14].

4
Extensions of the Contraction Principle

In Example 2.9 we have shown that the contraction principle can be used to derive large deviation results for SDEs of the form

$$dX_t^\varepsilon = b(X_t^\varepsilon)\,dt + \sqrt{\varepsilon}\,dB_t.$$

If we consider more interesting SDEs, for example

$$dX_t^\varepsilon = b(X_t^\varepsilon)\,dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dB_t,$$

this approach fails since the corresponding function F, defined in the proof of Example 2.9, is, in general, not continuous. Consequently, it is of interest to know in which way the contraction principle can be extended beyond the continuous case.
We present two generalizations of the contraction principle. The first applies to (measurable) functions which can be approximated by continuous functions and is due to Dembo-Zeitouni [10]. Garcia [21] introduces the notion of quasi-continuous almost compact functions and proves a contraction principle for this class of functions. It is worth mentioning that his paper contains a thorough discussion of some more extensions, see [21, Section 2].
In Chapter 7 we will use these extensions in order to deduce large deviation results for solutions of SDEs.

4.1 Exponential Approximations

Throughout this section $(\Omega, \mathcal{A}, \mathbb{P})$ denotes a probability space and (M, d) a metric space. Unless otherwise mentioned, the results are taken from [10, Section 4.2.2].

4.1 Definition Let $(X_{\varepsilon,m})_{\varepsilon>0,m\in\mathbb{N}}$ and $(Y_\varepsilon)_{\varepsilon>0}$ be families of M-valued random variables on $(\Omega, \mathcal{A}, \mathbb{P})$. $(X_{\varepsilon,m})_{\varepsilon>0,m\in\mathbb{N}}$ is called *exponentially good approximation of* $(Y_\varepsilon)_{\varepsilon>0}$ if

$$\lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}(d(X_{\varepsilon,m}, Y_\varepsilon) > \delta) = -\infty \qquad \text{for all } \delta > 0. \tag{4.1}$$

If $X_{\varepsilon,m}$ does not depend on $m \in \mathbb{N}$, we say that $(X_\varepsilon)_{\varepsilon>0}$ and $(Y_\varepsilon)_{\varepsilon>0}$ are *exponentially equivalent*.

The following theorem is the main result of this section and provides a relation between large deviation principles of exponentially good approximations.

4.2 Theorem *Let $(X_{\varepsilon,m})_{\varepsilon>0,m\in\mathbb{N}}$ be an exponentially good approximation of $(Y_\varepsilon)_{\varepsilon>0}$ such that $(X_{\varepsilon,m})_{\varepsilon>0}$ satisfies a large deviation principle with rate function I_m as $\varepsilon\to0$.*

(i). $(Y_\varepsilon)_{\varepsilon>0}$ satisfies a weak large deviation principle with rate function

$$I(y) := \sup_{\delta>0}\liminf_{m\to\infty}\inf_{z\in B(y,\delta)} I_m(z). \tag{4.2}$$

(ii). If I is a good rate function and

$$\inf_{y\in B} I(y) \leq \limsup_{m\to\infty}\inf_{y\in B} I_m(y) \tag{4.3}$$

holds for each closed set $B\subseteq M$, then $(Y_\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle with good rate function I.

Proof. (i). It suffices to show

$$I(y) = -\inf_{\delta>0}\limsup_{\varepsilon\to0}\varepsilon\log\mathbb{P}(Y_\varepsilon\in B[y,\delta]) = -\inf_{\delta>0}\liminf_{\varepsilon\to0}\varepsilon\log\mathbb{P}(Y_\varepsilon\in B(y,\delta)) \tag{4.4}$$

for any $y\in M$, cf. (2.1). Fix $\delta>0$ and $y\in M$. From

$$\mathbb{P}(X_{\varepsilon,m}\in B(y,\delta)) \leq \mathbb{P}(Y_\varepsilon\in B(y,2\delta)) + \mathbb{P}(d(X_{\varepsilon,m},Y_\varepsilon)>\delta)$$

we conclude by the large deviation lower bound (L1) for $(X_{\varepsilon,m})_{\varepsilon>0}$

$$\begin{aligned}-\inf_{z\in B(y,\delta)} I_m(z) &\leq \liminf_{\varepsilon\to0}\varepsilon\log\mathbb{P}(X_{\varepsilon,m}\in B(y,\delta))\\ &\leq \max\left\{\liminf_{\varepsilon\to0}\varepsilon\log\mathbb{P}(Y_\varepsilon\in B(y,2\delta)), \limsup_{\varepsilon\to0}\varepsilon\log\mathbb{P}(d(X_{\varepsilon,m},Y_\varepsilon)>\delta)\right\}.\end{aligned}$$

Since $(X_{\varepsilon,m})_{\varepsilon>0,m\in\mathbb{N}}$ is an exponentially good approximation, we get

$$\inf_{\delta>0}\liminf_{\varepsilon\to0}\varepsilon\log\mathbb{P}(Y_\varepsilon\in B(y,2\delta)) \geq \inf_{\delta>0}\limsup_{m\to\infty}\left(-\inf_{z\in B(y,\delta)} I_m(z)\right) = -I(y). \tag{4.5}$$

Interchanging the roles of $X_{\varepsilon,m}$ and Y_ε, we find

$$\begin{aligned}&\limsup_{\varepsilon\to0}\varepsilon\log\mathbb{P}(Y_\varepsilon\in B[y,\delta])\\ &\qquad\leq \max\left\{\limsup_{\varepsilon\to0}\varepsilon\log\mathbb{P}(X_{\varepsilon,m}\in B[y,2\delta]), \limsup_{\varepsilon\to0}\varepsilon\log\mathbb{P}((X_{\varepsilon,m},Y_\varepsilon)>\delta)\right\}.\end{aligned}$$

Therefore, by the large deviation upper bound for $(X_{\varepsilon,m})_{\varepsilon>0}$ and (4.1),

$$\inf_{\delta>0}\limsup_{\varepsilon\to0}\varepsilon\log\mathbb{P}(Y_\varepsilon\in B[y,\delta]) \leq \inf_{\delta>0}\limsup_{m\to\infty}\left(-\inf_{z\in B[y,2\delta]} I_m(z)\right) = -I(y). \tag{4.6}$$

Combining (4.5) and (4.6) yields (4.4).

(ii). It follows from the first part of this theorem that $(Y_\varepsilon)_{\varepsilon>0}$ satisfies a weak large deviation principle; it remains to show (L2) for any closed set $B \subseteq M$. Fix $\delta > 0$. The large deviation upper bound (L2) for $(X_{\varepsilon,m})_{\varepsilon>0}$ implies

$$\begin{aligned}
&\limsup_{\varepsilon\to0} \varepsilon \log \mathbb{P}(Y_\varepsilon \in B) \\
&\qquad \leq \max\Big\{\limsup_{\varepsilon\to0} \varepsilon \log \mathbb{P}(X_{\varepsilon,m} \in B + B[0,\delta]), \limsup_{\varepsilon\to0} \varepsilon \log \mathbb{P}(d(X_{\varepsilon,m}, Y_\varepsilon) > \delta)\Big\} \\
&\qquad \leq \max\Big\{-\inf_{y\in B+B[0,\delta]} I_m(y), \limsup_{\varepsilon\to0} \varepsilon \log \mathbb{P}(d(X_{\varepsilon,m}, Y_\varepsilon) > \delta)\Big\}.
\end{aligned}$$

Consequently, by (4.1) and (4.3),

$$\begin{aligned}
\limsup_{\varepsilon\to0} \varepsilon \log \mathbb{P}(Y_\varepsilon \in B) &\leq -\lim_{\delta\to0} \limsup_{m\to\infty} \inf_{y\in B+B[0,\delta]} I_m(y) \\
&\overset{(4.3)}{\leq} -\lim_{\delta\to0} \inf_{y\in B+B[0,\delta]} I(y) = -\inf_{y\in B} I(y).
\end{aligned}$$

In the last step we used that I is a good rate function, cf. [10, Lemma 4.1.6]. This finishes the proof. □

In particular, Theorem 4.2 entails the following corollary which we will apply in Chapter 5 in order to prove large deviation results for scaled Lévy processes. It is compiled from [8].

4.3 Corollary *Let $(M, \|\cdot\|)$ be a normed space and $\mathcal{B}$ a σ-algebra on M satisfying (C1), (C2) and (C7). Let $(X_\varepsilon)_{\varepsilon>0}$ and $(Y_\varepsilon)_{\varepsilon>0}$ be exponentially equivalent families of $\mathcal{B}$-measurable random variables. If $(X_\varepsilon)_{\varepsilon>0}$ is exponentially tight and obeys a large deviation principle with respect to $\mathcal{B}$ with good rate function I as $\varepsilon \to 0$, then $(Y_\varepsilon)_{\varepsilon>0}$ obeys a large deviation principle with respect to $\mathcal{B}$ with good rate function I as $\varepsilon \to 0$.*

Proof. We start with the proof of the large deviation lower bound (L1). Obviously, it suffices to show

$$\lim_{\varepsilon\to0} \varepsilon \log \mathbb{P}(Y_\varepsilon \in B(x,\delta)) \geq -I(x)$$

for any $\delta > 0$ and $x \in M$ such that $I(x) < \infty$. By (C1) and (C7), $B(x,\delta) \in \mathcal{B}$ and

$$(x,y) \mapsto d(x,y) := \|x - y\|$$

is $\mathcal{B} \otimes \mathcal{B}/\mathcal{B}(\mathbb{R})$-measurable. Fix $\varrho \in (0,1)$. We have

$$\begin{aligned}
\mathbb{P}(Y_\varepsilon \in B(x,\delta)) &\geq \mathbb{P}(X_\varepsilon \in B(x,\delta/2), d(X_\varepsilon, Y_\varepsilon) \leq \delta/2) \\
&\geq \mathbb{P}(X_\varepsilon \in B(x,\delta/2)) - \mathbb{P}(d(X_\varepsilon, Y_\varepsilon) > \delta/2) =: A_\varepsilon - B_\varepsilon.
\end{aligned}$$

By (L1) and (4.1),

$$A_\varepsilon \geq e^{-(I(x)+\varrho)\varepsilon^{-1}} > 0 \qquad\qquad B_\varepsilon \leq e^{-2(I(x)+1)\varepsilon^{-1}}$$

for $\varepsilon > 0$ sufficiently small. In particular, $B_\varepsilon/A_\varepsilon \to 0$ as $\varepsilon \to 0$. Consequently,

$$\begin{aligned}
\liminf_{\varepsilon\to0} \varepsilon \log \mathbb{P}(Y_\varepsilon \in B(x,\delta)) &\geq \liminf_{\varepsilon\to0} \varepsilon \log \left(A_\varepsilon \left(1 - \frac{B_\varepsilon}{A_\varepsilon}\right)\right) \\
&\geq \liminf_{t\to\infty} \varepsilon \log A_\varepsilon \geq -(I(x) + \gamma).
\end{aligned}$$

Since $\varrho > 0$ is arbitrary, this proves the large deviation lower bound. Now let $B \subseteq M$ closed, $\delta > 0$. For any $r \geq 0$, let $K_r \subseteq M$ compact as in the definition of exponential tightness. From

$$\mathbb{P}(Y_\varepsilon \in B) \leq \mathbb{P}(X_\varepsilon \in (B + B[0,\delta]) \cap K_r) + \mathbb{P}(X_\varepsilon \in K_r^c) + \mathbb{P}(d(X_\varepsilon, Y_\varepsilon) > \delta)$$

we conclude by the exponential tightness and (L2)

$$\begin{aligned}\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}(Y_\varepsilon \in B) &\leq \max\left\{\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}(X_\varepsilon \in (B + B[0,\delta]) \cap K_r), -r\right\} \\ &\leq \max\left\{-\inf_{x\in(B+B[0,\delta])\cap K_r} I(x), -r\right\} \\ &\leq \max\left\{-\inf_{x\in B+B[0,\delta]} I(x), -r\right\}.\end{aligned}$$

Note that $(B + B[0,\delta]) \cap K_r \in \mathcal{B}$ is closed. Letting $r \to \infty$ and $\delta \to 0$ yields

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}(Y_\varepsilon \in B) \leq -\lim_{\delta\to 0} \inf_{x\in B+B[0,\delta]} I(x) = -\inf_{x\in B} I(x).$$

Here we used the lower semicontinuity of I, cf. [10, Lemma 4.1.6]. This completes the proof. $\square$

Now we are ready to prove our first extension of the contraction principle.

4.4 Theorem (Extended contraction principle I) *Let (M_1, d_1), (M_2, d_2) be metric spaces and $(X_\varepsilon)_{\varepsilon>0}$ a family of random variables obeying a large deviation principle in (M_1, d_1) with good rate function I. For $m \in \mathbb{N}$ let $f_m : M_1 \to M_2$ be continuous functions and $f : M_1 \to M_2$ measurable such that*

$$\limsup_{m\to\infty} \sup_{x:I(x)\leq r} d_2(f_m(x), f(x)) = 0 \qquad \textit{for all } \; r \geq 0. \tag{4.7}$$

Then for any family of random variables $(Y_\varepsilon)_{\varepsilon>0}$ for which $(f_m(X_\varepsilon))_{\varepsilon>0,m\in\mathbb{N}}$ is an exponentially good approximation holds a large deviation principle with good rate function

$$J(y) = \inf\{I(x); y = f(x)\}.$$

Proof. Since the functions f_m, $m \in \mathbb{N}$, are continuous, the contraction principle entails that $(f_m(X_\varepsilon))_{\varepsilon>0}$ satisfies a large deviation principle with good rate function

$$J_m(y) := \inf\{I(x); y = f_m(x)\}.$$

Moreover, by (4.7), f is continuous on any sublevel set $\Phi_I(r) := \{x \in M_1; I(x) \leq r\}$, $r \geq 0$. Hence, J is a good rate function with sublevel sets $f(\Phi_I(r))$. In view of Theorem 4.2 it suffices to check (4.3) and to identify the rate function. To this end, fix $B \subseteq M_2$ closed and $\delta > 0$, and set

$$c := \liminf_{m\to\infty} \inf_{y\in B} J_m(y).$$

Suppose that $c < \infty$; then we can choose a sequence $(x_m)_{m\in\mathbb{N}} \subseteq M_1$ and $r > 0$ such that $f_m(x_m) \in B$ and $I(x_m) = \inf_{y\in B} J_m(y) \leq r$. From (4.7) we conclude $f(x_m) \in B + B[0,\delta]$ for $m = m(\delta)$ sufficiently large. Thus,

$$\inf_{y\in B+B[0,\delta]} J(y) \leq J(f(x_m)) \leq I(x_m) = \inf_{y\in B} J_m(y).$$

Taking $\delta \to 0$ and $m \to \infty$, we find

$$\inf_{y\in B} J(y) \leq \liminf_{m\to\infty} \inf_{y\in B} J_m(y) = c.$$

Obviously, this inequality is trivially satisfied if $c = \infty$. In particular, (4.3) holds. In order to identify the rate function, we use the preceding inequality for $B := B[y,\delta]$ and let $\delta \to 0$. For more details, see [10, Theorem 4.2.23]. □

4.2 Quasi-Continuity & Almost Compactness

In this section we prove a contraction principle for quasi-continuous almost compact functions. The results are adapted from Garcia [21] who considers topological spaces whereas we restrict ourselves to metric spaces. Throughout this section (M_1, d_1), (M_2, d_2) denote metric spaces and $(X_\varepsilon)_{\varepsilon>0}$ a family of M_1-valued random variables. For a function $f : M_1 \to M_2$ and $x \in M_1$ we set

$$f^x := \{y \in M_2; \exists (x_n)_{n\in\mathbb{N}} \subseteq M_1 : x_n \to x, f(x_n) \to y\}.$$

We start with the definition of quasi-continuity and show its relevance for the large deviation lower bound.

4.5 Definition Let $f : M_1 \to M_2$ be a function and $x \in M_1$. We call f *quasi-continuous at* x if for every $y \in f^x$ there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq M_1$ such that $x_n \to x$, $f(x_n) \to y$, and f is continuous at x_n for all $n \in \mathbb{N}$. We call any such sequence *companion sequence* for the tuple (x, y). We say that f is *quasi-continuous* if it is quasi-continuous at any $x \in M_1$.

4.6 Theorem *Let $f : M_1 \to M_2$ be quasi-continuous. Assume that for any $x \in M_1$ and $y \in f^x$ there exists a companion sequence $(x_n)_{n\in\mathbb{N}}$ such that $I(x_n) \to I(x)$. If $(X_\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle in (M_1, d_1) with rate function I, then $(X_\varepsilon, f(X_\varepsilon))_{\varepsilon>0}$ satisfies the large deviation lower bound (L1) with rate function*

$$J(x,y) := \begin{cases} I(x), & y \in f^x, \\ \infty, & \textit{otherwise}. \end{cases}$$

Proof. First of all, the lower semicontinuity of J follows directly from the lower semicontinuity of I, cf. [21, Lemma 4.2]. Moreover, it suffices to show (L1) for $A = A_1 \times A_2$ where $A_1 \subseteq M_1$, $A_2 \subseteq M_2$ are open sets. By the large deviation lower for $(X_\varepsilon)_{\varepsilon>0}$, we have

$$\begin{aligned} \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}((X_\varepsilon, f(X_\varepsilon)) \in A_1 \times A_2) &\geq \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\big(X_\varepsilon \in \operatorname{int}(A_1 \cap f^{-1}(A_2))\big) \\ &\geq - \inf_{z\in\operatorname{int}(A_1\cap f^{-1}(A_2))} I(z). \end{aligned}$$

Consequently, we are done if

$$\inf_{z \in \mathrm{int}(A_1 \cap f^{-1}(A_2))} I(z) \leq J(x,y) \qquad \text{for all } \ (x,y) \in A_1 \times A_2. \tag{4.8}$$

Fix $(x,y) \in A_1 \times A_2$ and $\varepsilon > 0$. If $y \notin f^x$, (4.8) holds trivially. Otherwise, there exists a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \to x$, $f(x_n) \to y$, $I(x_n) \to I(x)$, and f is continuous at x_n. Since A_2 is open, $B(y,\delta) \subseteq A_2$ for $\delta > 0$ sufficiently small. For $n \in \mathbb{N}$ sufficiently large we have $x_n \in A_1$, $f(x_n) \in B(y,\delta) \subseteq A_2$, $I(x_n) \leq I(x)+\varepsilon$. As f is continuous at x_n, this implies $x_n \in \mathrm{int}(A_1 \cap f^{-1}(A_2))$. Hence,

$$\inf_{z \in \mathrm{int}(A_1 \cap f^{-1}(A_2))} I(z) \leq I(x_n) \leq I(x) + \varepsilon = J(x,y) + \varepsilon. \qquad \square$$

4.7 Definition A function $f : M_1 \to M_2$ is called *almost compact at* $x \in M_1$ if for any sequence $(x_n)_{n\in\mathbb{N}}, x_n \to x$, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $(f(x_{n_k}))_{k\in\mathbb{N}}$ converges.

4.8 Example (i). Càdlàg and càglàd functions are quasi-continuous and almost compact.

(ii). The function

$$f(x) := \begin{cases} \sin\left(\frac{1}{x}\right), & x \neq 0, \\ f(0), & x = 0, \end{cases}$$

is quasi-continuous and almost compact for any $f(0) \in [-1,1]$.

(iii). $f(x) := -\mathbb{1}_{(-\infty,0)}(x) + \mathbb{1}_{(0,\infty)}(x)$, $x \in \mathbb{R}$, is not quasi-continuous but almost compact at $x = 0$.

(iv). $f(x) := \frac{1}{x}\mathbb{1}_{(0,\infty)}(x)$, $x \in \mathbb{R}$, is quasi-continuous but not almost compact at $x = 0$.

4.9 Theorem *Suppose that $(X_\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle in (M_1,d_1) with rate function I. Let $f : M_1 \to M_2$ be almost compact on $\mathrm{dom}\, I$. Then $(X_\varepsilon, f(X_\varepsilon))_{\varepsilon>0}$ satisfies the large deviation upper bound (L2) with rate function J defined in Theorem 4.6.*

Proof. Fix $B \subseteq M_1 \times M_2$ closed. By (L2) for $(X_\varepsilon)_{\varepsilon>0}$,

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}((X_\varepsilon, f(X_\varepsilon)) \in B) \leq -\inf_{x\in\tilde{B}} I(x)$$

for $\tilde{B} := \mathrm{cl}(\{x \in M_1; (x,f(x)) \in B\})$. For any $x \in \tilde{B}$, there exists $(x_n)_{n\in\mathbb{N}} \subseteq M_1$ such that $(x_n, f(x_n)) \in B$ and $x_n \to x$. If $x \in \mathrm{dom}\, I$ then the almost compactness implies $(x_n, f(x_n)) \to (x,y)$ for some $y \in M_2$ and a suitable subsequence of $(x_n)_{n\in\mathbb{N}}$. As B is closed, $(x,y) \in B$. This shows

$$I(x) = J(x,y) \geq \inf_{(x_1,y_1)\in B} J(x_1,y_1) \qquad \text{for all } \ x \in \tilde{B} \cap \mathrm{dom}\, I.$$

Note that this inequality is trivially satisfied for $x \notin \mathrm{dom}\, I$. Thus,

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}((X_\varepsilon, f(X_\varepsilon)) \in B) \leq -\inf_{(x_1,y_1)\in B} J(x_1,y_1). \qquad \square$$

Combining Theorem 4.6 and Theorem 4.9, we find

4.10 Theorem (Extension of the contraction principle II) *Suppose that $f : M_1 \to M_2$ satisfies the following conditions.*

(i). f is almost compact on $\operatorname{dom} I$.

(ii). For each $x \in M_1$ there exists a companion sequence $(x_n)_{n\in\mathbb{N}}$ such that $I(x_n) \to I(x)$. In particular, f is quasi-continuous.

If $(X_\varepsilon)_{\varepsilon>0}$ obeys a large deviation principle with (good) rate function I, then $(X_\varepsilon, f(X_\varepsilon))_{\varepsilon>0}$ satisfies a large deviation principle with (good) rate function J defined in Theorem 4.6. In particular, $(f(X_\varepsilon))_{\varepsilon>0}$ obeys a large deviation principle with (good) rate function

$$I_0(y) := \inf\{J(x,y); x \in M_1 : y \in f^x\} = \inf\{I(x); x \in M_1 : y \in f^x\}.$$

Let us finally mention the following theorem which shows basically that almost compactness preserves exponential tightness.

4.11 Theorem ([21, Theorem 6.3], [22, Theorem 7.1])

(i). If $(X_\varepsilon)_{\varepsilon>0}$ is exponentially tight and $f : M_1 \to M_2$ almost compact, then $(f(X_\varepsilon))_{\varepsilon>0}$ is exponentially tight.

(ii). Let $(X_\varepsilon)_{\varepsilon>0}$ be exponentially tight and $(f_m)_{m\in\mathbb{N}}$ a sequence of almost compact functions. If $(f_m(X_\varepsilon))_{\varepsilon>0, m\in\mathbb{N}}$ is an exponentially good approximation of $(Y_\varepsilon)_{\varepsilon>0}$, then $(X_\varepsilon, Y_\varepsilon)_{\varepsilon>0}$ is exponentially tight.

5

Large Deviation Principles for Scaled Lévy Processes

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space and $(L_t)_{t\geq 0}$ a real-valued *Lévy process* on $(\Omega, \mathcal{A}, \mathbb{P})$, i. e. a family of random variables $L_t : \Omega \to \mathbb{R}$, $t \geq 0$, such that

(i). $L_0 = 0$

(ii). $L_t - L_s \sim L_{t-s}$ for $0 \leq s < t$ (stationary increments)

(iii). $(L_{t_j} - L_{t_{j-1}})_{j=1,\ldots,n}$ are independent for $0 = t_0 < t_1 < \ldots < t_n$, $n \in \mathbb{N}$ (independent increments)

(iv). $t \mapsto L_t(\omega)$ is càdlàg for all $\omega \in \Omega$.

For $x \in \mathbb{R}$ we call $(x + L_t)_{t\geq 0}$ *Lévy process started at x*. Throughout this chapter

$$L_t = \gamma t + \sigma B_t + \int_0^t \int_{|z|>1} z\, N(dz, ds) + \int_0^t \int_{0<|z|\leq 1} z\, \tilde{N}(dz, ds)$$

denotes the *Lévy-Itô decomposition* of $(L_t)_{t\geq 0}$, where $(B_t)_{t\geq 0}$ is a Brownian motion, N the *jump counting measure* of $(L_t)_{t\geq 0}$, $\tilde{N}$ its *compensated jump counting measure*, and (γ, σ^2, ν) the *Lévy triplet* comprising the *drift* $\gamma \in \mathbb{R}$, the *diffusion coefficient* $\sigma \geq 0$, and the *Lévy measure* ν on $(\mathbb{R}\backslash\{0\}, \mathcal{B}(\mathbb{R}\backslash\{0\}))$ which satisfies

$$\int_{\mathbb{R}\backslash\{0\}} (y^2 \wedge 1)\, \nu(dy) < \infty.$$

By *Lévy-Khinchine's formula*, the characteristic function of L_t is given by

$$\mathbb{E} e^{\imath \xi L_t} = e^{-t\psi(\xi)}, \qquad \xi \in \mathbb{R}, t \geq 0 \tag{5.1}$$

where

$$\psi(\xi) := -\imath \gamma \xi + \frac{1}{2}\sigma^2 \xi^2 + \int_{\mathbb{R}\backslash\{0\}} \left(1 - e^{\imath y \xi} + \imath \xi y \mathbb{1}_{|y|\leq 1}\right) \nu(dy), \qquad \xi \in \mathbb{R}$$

is called the *characteristic exponent* of $(L_t)_{t\geq 0}$. We say that $(L_t)_{t\geq 0}$ is a *Lévy process without Gaussian component* if $\sigma = 0$. For a thorough discussion of Lévy processes we refer the reader to [36].

In this chapter we consider large deviations for scaled Lévy processes of the form

$$\omega \mapsto \frac{L(t\cdot,\omega)}{S(t)} \in D[0,1]$$

where S is a suitable increasing function. One of the first results is due to Wentzell [40] who studied a certain type of time-homogeneous Markov processes. We discuss this approach in Chapter 7. Lynch-Sethuraman [29] proved a large deviation principle for Lévy processes of bounded variation and the scaling function $S(t) = t$ with respect to the weak*-topology on $\mathrm{BV}[0,1] \cap D[0,1]$. These results were generalized by Mogulskii [31] to Lévy processes with possibly infinite variation and the J_1-topology. Moreover, Mogulskii stated a large deviation principle with respect to the uniform norm under the growth condition

$$\lim_{t\to\infty} \frac{S(t)}{t} = 0 \quad \text{and} \quad \lim_{t\to\infty} \frac{S(t)}{\sqrt{t}} = \infty. \tag{5.2}$$

In this work we follow the approach suggested by de Acosta [8] who proved large deviations for Lévy processes taking values in a separable Banach space relative to the scaling function $S(t) = t$. More recently, Feng and Kurtz [17] developed a technique using tools of viscosity solutions and nonlinear semigroups associated with Markov processes. In particular, all earlier results on scaled Lévy processes are covered, see [17, Section 10.1].

In order to show large deviations for scaled Lévy processes, one has to pose an exponential integrability condition on L_1; namely, for $S(t) = t$,

$$\mathbb{E}e^{\lambda|L_1|} < \infty \qquad \text{for all} \quad \lambda \in (0,\lambda_0]^1 \neq \varnothing$$

for results with respect to the weak*- or J_1-topology, and

$$\mathbb{E}e^{\lambda|L_1|} < \infty \qquad \text{for all} \quad \lambda \geq 0$$

for large deviations with respect to the uniform topology.

We start with a large deviation result relative to the scaling function $S(t) = t$. Subsequently, we modify the proof appropriately in order to prove a large deviation principle under the growth condition (5.2). Finally, two results on the longtime behavior of Lévy processes are stated: the *law of iterated logarithm* and the counterpart of *Strassen's law*; both under an exponential integrability condition. They are typical applications of large deviation results.

We denote by $D[0,1]$ the space of real-valued càdlàg functions on $[0,1]$ endowed with the uniform norm

$$\|f\|_\infty := \sup_{t\in[0,1]} |f(t)|$$

and the σ-algebra $\mathcal{B} := \sigma(\pi_t; t \in [0,1])$ generated by the projections $\pi_t : D[0,1] \to \mathbb{R}$, $f \mapsto f(t)$, $t \in [0,1]$. Let us remark that $\mathcal{B}$ equals the Borel σ-algebra generated by the J_1-metric, see e. g. [14, Proposition 3.7.1] or [4, Theorem 12.5].

[1] Revised version: Corrected misprint throughout this chapter.

5.1 Theorem *Let $(L_t)_{t\geq 0}$ be a Lévy process started at x such that*

$$\mathbb{E}e^{\lambda|L_1|} < \infty \qquad \text{for all} \;\; \lambda \geq 0. \tag{5.3}$$

Then $(L(t\cdot)/t)_{t>0}$ satisfies a large deviation principle in $(D[0,1], \|\cdot\|_\infty)$ with respect to $\mathcal{B}$ as $t \to \infty$ with good rate function I,

$$I(f) := \begin{cases} \int_0^1 \Psi^*(f'(s))\,ds, & f \in AC[0,1], f(0) = x, \\ \infty, & \text{otherwise}, \end{cases} \tag{5.4}$$

where

$$\Psi(w) := \gamma w + \frac{1}{2}\sigma^2 w^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{yw} - 1 - yw\,\mathbb{1}_{|y|\leq 1}\right)\nu(dy), \qquad w \in \mathbb{R} \tag{5.5}$$

denotes the logarithmic moment generating function of L_1, i. e.

$$\liminf_{t\to\infty} \frac{1}{t}\log\mathbb{P}\left(\frac{L(t\cdot)}{t} \in A\right) \geq -\inf_{f\in A} I(f) \tag{5.6}$$

$$\limsup_{t\to\infty} \frac{1}{t}\log\mathbb{P}\left(\frac{L(t\cdot)}{t} \in B\right) \leq -\inf_{f\in B} I(f) \tag{5.7}$$

for any open set $A \in \mathcal{B}$ and closed set $B \in \mathcal{B}$.

First of all, we note that it suffices, because of the spatial homogeneity of the Lévy process $(L_t)_{t\geq 0}$, to consider $x = 0$. We split the proof into several steps:

(i). The sequence of discretizations $(Z_n)_{n\in\mathbb{N}}$ defined by

$$\frac{Z_n(s,\omega)}{n} := \frac{1}{n}L(\lfloor n\cdot s\rfloor,\omega) = \frac{1}{n}\left(\sum_{j=0}^{n-1} L(j,\omega)\mathbb{1}_{[j/n,(j+1)/n)}(s) + L(n,\omega)\mathbb{1}_{\{1\}}(s)\right)$$

is exponentially tight in $(D[0,1], \|\cdot\|_\infty)$, cf. Lemma 5.3.

(ii). $(Z_n/n)_{n\in\mathbb{N}}$ satisfies a large deviation principle in $(D[0,1], \|\cdot\|_\infty)$ with respect to $\mathcal{B}$ with good rate function I,

$$I(f) := \sup_{\alpha\in \mathrm{BV}[0,1]\cap D[0,1]} \left(\int_0^1 f\,d\alpha - \int_0^1 \Psi(\alpha(1) - \alpha(s))\,ds\right), \tag{5.8}$$

cf. Theorem 5.5.

(iii). $(Z_{\lfloor t\rfloor}/\lfloor t\rfloor)_{t>0}$ and $(L(t\cdot)/t)_{t>0}$ are exponentially equivalent, cf. Lemma 5.6.

(iv). $(L(t\cdot)/t)_{t>0}$ satisfies a large deviation principle with good rate function I and I equals the rate function defined in (5.4), cf. Theorem 5.7.

Remark The mapping $\omega \mapsto L(t\cdot,\omega) \in D[0,1]$ is $\mathcal{A}/\mathcal{B}$-measurable for each $t \geq 0$. In particular, the probabilities appearing in (5.6) and (5.7) are well-defined. If $\omega \mapsto L(t\cdot,\omega)$ is measurable with respect to the Borel-σ-algebra generated by the uniform norm $\|\cdot\|_\infty$ on $D[0,1]$, then (5.6) and (5.7) hold for any open set A and closed set B, respectively. Mind that $\mathcal{B} \subsetneq \mathcal{B}((D[0,1], \|\cdot\|_\infty))$, cf. [4, p. 157].

In order to show the exponential tightness of $(Z_n/n)_{n\in\mathbb{N}}$, we need the following lemma [17, Lemma 3.3].

5.2 Lemma *Let (M,d) be a metric space and $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of tight probability measures on $(M,\mathcal{B}(M))$. If for any $r>0$, $\varepsilon>0$, there exists $K_{r,\varepsilon}\subseteq M$ compact such that*

$$\limsup_{n\to\infty}\frac{1}{n}\log\mu_n(\{x\in M; d(x,K_{r,\varepsilon})\geq\varepsilon\})\leq -r$$

then $(\mu_n)_{n\in\mathbb{N}}$ is exponentially tight.

Proof. Since μ_n is tight for each $n\in\mathbb{N}$, we may assume that

$$\sup_{n\in\mathbb{N}}\frac{1}{n}\log\mu_n(\{x\in M; d(x,K_{r,\varepsilon})\geq\varepsilon\})\leq -r$$

i. e.

$$\mu_n(\{x\in M; d(x,K_{r,\varepsilon})\geq\varepsilon\})\leq e^{-nr}.$$

Define

$$K_r := \mathrm{cl}\left(\bigcap_{k\in\mathbb{N}}\left(K_{r+k,1/k}+B(0,1/k)\right)\right).$$

It is not diffcult to show that K_r is complete and totally bounded, hence compact. Moreover,

$$\mu_n(K_r^c)\leq\sum_{k\in\mathbb{N}}\mu_n(\{x\in M; d(x,K_{r+k,1/k})\geq 1/k\})\leq\sum_{k\in\mathbb{N}}e^{-n(r+k)}\leq\frac{e}{e-1}e^{-nr}.$$

This proves the exponential tightness of $(\mu_n)_{n\in\mathbb{N}}$. □

5.3 Lemma *(i). For each $n\in\mathbb{N}$, Z_n/n is tight in $(D[0,1],\|\cdot\|_\infty)$.*

(ii). $(Z_n/n)_{n\in\mathbb{N}}$ is exponentially tight.

Remark It is widely known that any probability measure on a Polish space is tight. Since $(D[0,1],\|\cdot\|_\infty)$ is not a Polish space – it is not separable – this result does not apply.

Proof. Since the mapping

$$(\mathbb{R}^n,\|\cdot\|)\ni x\mapsto(T_nx)(t):=\sum_{j=1}^{n-1}x_j\mathbb{1}_{[j/n,(j+1)/n)}(t)+x_n\mathbb{1}_{\{1\}}(t)\in(D[0,1],\|\cdot\|_\infty)$$

is continuous, it follows that $T_n(K)$ is compact for any compact set $K\subseteq\mathbb{R}^n$. For $K\subseteq\mathbb{R}$ compact, we have

$$\mathbb{P}\left(\frac{Z_n}{n}\notin T_n(K^n)\right)\leq\sum_{j=1}^n\mathbb{P}\left(\frac{L_j}{n}\notin K\right).$$

Since L_j/n is tight – its distribution is a probability measure on $(\mathbb{R},\mathcal{B}(\mathbb{R}))$, hence tight – for $j=1,\ldots,n$, we conclude that Z_n/n is tight in $(D[0,1],\|\cdot\|_\infty)$. This proves (i). We

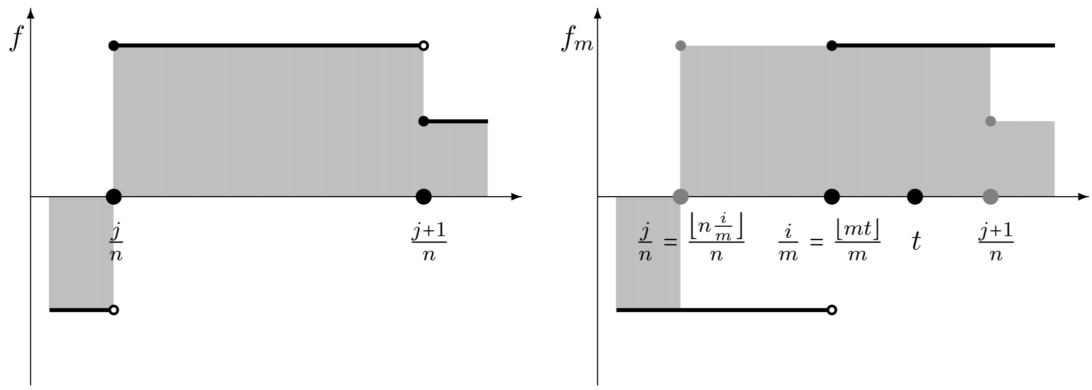

Figure 5.1: The function $f \in T_n(K^n)$ and its approximation f_m.

proceed to show that the assumptions of Lemma 5.2 are satisfied. To this end fix $r > 0$ and $\varepsilon > 0$. For $K \subseteq \mathbb{R}$ and $n \geq m$, we have

$$\begin{aligned}\mathbb{P}\left(d\left(\frac{Z_n}{n}, T_m(K^m)\right) > \varepsilon\right) &\leq \mathbb{P}\left(\frac{Z_n}{n} \notin T_n(K^n)\right) + \mathbb{P}\left(\frac{Z_n}{n} \in T_n(K^n), d\left(\frac{Z_n}{n}, T_m(K^m)\right) > \varepsilon\right) \\ &=: I_1 + I_2. \end{aligned} \tag{5.9}$$

We choose $K := [-r, r]$ and estimate the terms separately. Applying Etemadi's inequality, cf. Corollary A.2, and Markov's inequality yields

$$I_1 = \mathbb{P}\left(\sup_{1\leq j\leq n}\left|\frac{L_j}{n}\right| > r\right) \leq 3\sup_{1\leq j\leq n}\mathbb{P}\left(|L_j| > \frac{nr}{3}\right) \leq 3\sup_{1\leq j\leq n}\mathbb{E}e^{|L_j|-nr/3} \leq 3e^{-nr/3}\beta_1^n$$

where $\beta_1 := \mathbb{E}e^{|L_1|} < \infty$ because of (5.3). In order to estimate I_2 we observe that if we set $f_m := f(\lfloor m\cdot\rfloor/m)$, then

$$d(f, T_m(K^m)) \leq \|f - f_m\|_\infty \qquad \text{for all} \quad f \in T_n(K^n). \tag{5.10}$$

Moreover,

$$\begin{aligned}\|f - f_m\|_\infty &= \max_{1\leq i\leq m-1}\sup_{t\in[i/m,(i+1)/m)}|f(t) - f_m(t)| \\ &= \max_{1\leq i\leq m-1}\sup_{t\in[i/m,(i+1)/m)}\left|f\left(\frac{\lfloor nt\rfloor}{n}\right) - f\left(\frac{\lfloor mt\rfloor}{m}\right)\right| \\ &\leq \max_{1\leq i\leq m-1}\sup_{1\leq j\leq\lfloor n/m\rfloor+1}\left|f\left(\frac{\lfloor n\frac{i}{m}\rfloor}{n} + \frac{j}{n}\right) - f\left(\frac{\lfloor n\frac{i}{m}\rfloor}{n}\right)\right|.\end{aligned} \tag{5.11}$$

For the last line we used that

$$f\left(\frac{\lfloor n\frac{i}{m}\rfloor}{n}\right) = f\left(\frac{i}{m}\right) = f\left(\frac{\lfloor mt\rfloor}{m}\right) \qquad \text{for all} \quad t \in \left[\frac{i}{m}, \frac{i+1}{m}\right)$$

as $f \in T_n(K^n)$, see Figure 5.1. Combining (5.10) and (5.11), we get

$$\begin{aligned} I_2 &\leq \mathbb{P}\left(\sup_{1\leq i\leq m-1}\sup_{1\leq j\leq \lfloor n/m\rfloor+1}\left|Z_n\left(\frac{\lfloor n\frac{i}{m}\rfloor}{n}+\frac{j}{n}\right)-Z_n\left(\frac{\lfloor n\frac{i}{m}\rfloor}{n}\right)\right|>n\varepsilon\right) \\ &\leq \sum_{i=1}^{m-1}\mathbb{P}\left(\sup_{1\leq j\leq \lfloor n/m\rfloor+1}\left|L\left(\left\lfloor n\frac{i}{m}\right\rfloor+j\right)-L\left(\left\lfloor n\frac{i}{m}\right\rfloor\right)\right|>n\varepsilon\right). \end{aligned}$$

By the stationarity and independence of the increments and Markov's inequality,

$$\begin{aligned} I_2 \leq m\mathbb{P}\left(\sup_{1\leq j\leq \lfloor n/m\rfloor+1}|L_j|>n\varepsilon\right) &\leq 3m\sup_{1\leq j\leq \lfloor n/m\rfloor+1}\mathbb{P}\left(|L_j|>\frac{n\varepsilon}{3}\right) \\ &\leq 3m\sup_{1\leq j\leq \lfloor n/m\rfloor+1}\mathbb{E}e^{r|L_j|-nr\varepsilon/3} \\ &\leq 3m\beta_2^{\lfloor n/m\rfloor+1}e^{-nr\varepsilon/3} \end{aligned}$$

where $\beta_2 := \mathbb{E}e^{r|L_1|}<\infty$. Consequently,

$$\begin{aligned} \limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(d\left(\frac{Z_n}{n},T_m(K^m)\right)>\varepsilon\right) &\leq \max\left\{\log\beta_1-\frac{r}{3},\frac{1}{m}\log\beta_2-\frac{r\varepsilon}{3}\right\} \\ &\xrightarrow{r,m\to\infty} -\infty. \end{aligned}$$

Now the claim follows from Lemma 5.2. □

As a next step we prove a large deviation principle for $(Z_n/n)_{n\in\mathbb{N}}$ using the results of Chapter 3. In particular, we have to verify the conditions (C1)-(C7). Let us recall *Riesz' representation theorem*, see e. g. [27, Theorem 2.6.1]. It states that the topological dual of $(C[0,1],\|\cdot\|_\infty)$ is isomorphic to $\mathrm{BV}[0,1]\cap D[0,1]$, the space of càdlàg functions of bounded variation on $[0,1]$. The mapping $I:\mathrm{BV}[0,1]\cap D[0,1]\to D[0,1]^*$,

$$(I\alpha)(f) := \langle\alpha,f\rangle := \int_0^1 f\,d\alpha := \lim_{n\to\infty}\sum_{j=0}^{n-1} f\left(\frac{j+1}{n}\right)\left[\alpha\left(\frac{j+1}{n}\right)-\alpha\left(\frac{j}{n}\right)\right],$$

allows us to consider any set $F\subseteq \mathrm{BV}[0,1]\cap D[0,1]$ as a subset of $D[0,1]^*$. Observe that $(I\alpha)(f)$ is well-defined, see for instance [11, Proposition 2.1.6]. We call $(I\alpha)(f)$ *Lebesgue-Stieltjes integral of f with respect to α.*

5.4 Lemma *For $(M,\|\cdot\|) := (D[0,1],\|\cdot\|_\infty)$, $M_0 := C[0,1]$, $F := \mathrm{BV}[0,1]\cap D[0,1]$ and $\mathcal{B} := \sigma(\pi_t; t\in[0,1])$ the conditions (C1)-(C7) are satisfied.*

Proof. (C1) The mapping $(f,g)\mapsto f+g$ is $\mathcal{B}\otimes\mathcal{B}/\mathcal{B}$-measurable since its composition with the projection π_t, $t\geq 0$, is $\mathcal{B}\otimes\mathcal{B}/\mathcal{B}(\mathbb{R})$-measurable. Obviously, $f\mapsto\lambda f$ is $\mathcal{B}/\mathcal{B}$-measurable.

(C2) cf. Lemma A.10.

(C3) Let $\alpha\in\mathrm{BV}[0,1]\cap D[0,1]$. The mapping

$$f\mapsto\sum_{j=0}^{n-1} f\left(\frac{j+1}{n}\right)\left[\alpha\left(\frac{j+1}{n}\right)-\alpha\left(\frac{j}{n}\right)\right]$$

is $\mathcal{B}/\mathcal{B}(\mathbb{R})$-measurable. Since the right side converges to $\langle\alpha,f\rangle$ as $n\to\infty$, the claim follows.

(C4) Obvious.

(C5) This follows from Riesz' representation theorem.

(C6) Clear by Riesz' representation theorem.

(C7) Let $(t_k)_{k\in\mathbb{N}}$ be dense in $[0,1]$. Then, by the right-continuity,

$$B_{D[0,1]}[0,\varepsilon] = \bigcap_{k\in\mathbb{N}} \pi_{t_k}^{-1}(B_{\mathbb{R}}[0,\varepsilon]) \cap \pi_1^{-1}[B_{\mathbb{R}}(0,\varepsilon)] \in \mathcal{B} \qquad \text{for all } \varepsilon > 0. \qquad \square$$

The next result is based on [8, Lemma 4.2].

5.5 Theorem *The sequence $(Z_n/n)_{n\in\mathbb{N}}$ satisfies a large deviation principle in $(D[0,1], \|\cdot\|_\infty)$ with respect to $\mathcal{B}$ as $n\to\infty$ with good rate function I defined by*

$$I(f) := \sup_{\alpha\in \mathrm{BV}[0,1]\cap D[0,1]} \left(\int_0^1 f\, d\alpha - \int_0^1 \Psi(\alpha(1)-\alpha(s))\, ds \right). \tag{5.12}$$

Proof. In view of Theorem 3.5, Theorem 3.9 and Lemma 5.4, it suffices to show that

(i). The limit
$$\Lambda(\alpha) := \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E} e^{\langle \alpha, Z_n\rangle}$$
exists and equals $\int_0^1 \Psi(\alpha(1)-\alpha(s))\,ds$ for each $\alpha \in \mathrm{BV}[0,1]\cap D[0,1]$.

(ii). Λ is $C[0,1]$-Gâteaux differentiable on $\mathrm{BV}[0,1]\cap D[0,1]$.

(iii). $\operatorname{dom} I \subseteq \mathrm{BV}[0,1]\cap D[0,1]$

We defer the proof of (iii) to Theorem 5.7; there we will show that $\operatorname{dom} I \subseteq AC[0,1]$. As $AC[0,1] \subseteq \mathrm{BV}[0,1]\cap D[0,1]$, this proves (iii). Note that

$$Z_n = \sum_{j=1}^{n-1} L_j \mathbb{1}_{\left[\frac{j}{n}, \frac{j+1}{n}\right)} + L_n \mathbb{1}_{\{1\}} = \sum_{j=1}^{n} (L_j - L_{j-1}) \mathbb{1}_{\left[\frac{j}{n}, 1\right]}.$$

Consequently, by the stationarity and independence of the increments,

$$\begin{aligned} \mathbb{E} e^{\langle \alpha, Z_n\rangle} &= \mathbb{E} \exp\left(\sum_{j=1}^{n} (L_j - L_{j-1})(\alpha(1)-\alpha(j/n)) \right) \\ &= \prod_{j=1}^{n} \mathbb{E} \exp\left(L_1(\alpha(1)-\alpha(j/n))\right). \end{aligned}$$

Since $(L_t)_{t\geq 0}$ is a Lévy process with finite exponential moments, we have

$$\mathbb{E} e^{\lambda L_1} = e^{\Psi(\lambda)} \qquad \text{for all } \lambda \in \mathbb{R},$$

where Ψ is given by (5.5), see [36, Theorem 25.17]. Therefore,

$$\Lambda(\alpha) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \Psi(\alpha(1)-\alpha(j/n)) = \int_0^1 \Psi(\alpha(1)-\alpha(s))\, ds.$$

Pick $\beta \in \mathrm{BV}[0,1] \cap D[0,1]$ and set

$$u(t,s) := \Psi\Big((\alpha(1) - \alpha(s)) + t(\beta(1) - \beta(s))\Big), \quad t \in [-1,1], s \in [0,1].$$

Obviously, $\alpha, \beta \in \mathrm{BV}[0,1]$ implies $\|\alpha\|_\infty + \|\beta\|_\infty \le C < \infty$. By (5.3), we have

$$-\infty < \log \mathbb{E}e^{-2C|L_1|} \le |u(t,s)| \le \log \mathbb{E}e^{2C|L_1|} < \infty.$$

By the differentiability lemma for parameter-dependent integrals, cf. [37, Theorem 11.5], $u(\cdot, s)$ is differentiable and

$$|\partial_t u(t,s)| \le 2C \frac{1}{\mathbb{E}e^{-2C|L_1|}} \sqrt{\mathbb{E}(L_1^2)} \sqrt{\mathbb{E}e^{2C|L_1|}} < \infty \qquad \text{for all} \quad t \in [-1,1].$$

Another application of the differentiabilty lemma yields

$$\begin{aligned}\frac{\Lambda(\alpha + t\beta) - \Lambda(\alpha)}{t} &\xrightarrow{t \to 0} \int_0^1 \partial_t u(0,s)\, ds \\ &= \int_0^1 (\beta(1) - \beta(s)) \frac{1}{\mathbb{E}e^{L_1(\alpha(1) - \alpha(s))}} \mathbb{E}(L_1 e^{L_1(\alpha(1)-\alpha(s))})\, ds.\end{aligned}$$

This shows that Λ is $C[0,1]$-Gâteaux differentiable at α, and its derivative equals, by the integration-by-parts formula Lemma A.9,

$$D_\alpha(t) := \int_0^t \frac{1}{\mathbb{E}e^{L_1(\alpha(1)-\alpha(s))}} \mathbb{E}(L_1 e^{L_1(\alpha(1)-\alpha(s))})\, ds, \quad t \in [0,1]. \qquad \square$$

Remark Observe that $(Z_n/n)_{n \in \mathbb{N}}$ satisfies a large deviation principle with good rate function I as $n \to \infty$ if, and only if, $(Z_{\lfloor t \rfloor}/\lfloor t \rfloor)_{t>0}$ satisfies a large deviation principle with good rate function I as $t \to \infty$.

Finally, we claim that the large deviation principle for $(Z_{\lfloor t \rfloor}/\lfloor t \rfloor)_{t>0}$ entails the large deviation principle for $(L(t\cdot)/t)_{t>0}$. In view of Corollary 4.3, it suffices to show that $(Z_{\lfloor t \rfloor}/\lfloor t \rfloor)_{t>0}$ and $(L(t\cdot)/t)_{t>0}$ are exponentially equivalent. The following result is essentially [8, Lemma 4.3].

5.6 Lemma *$(Z_{\lfloor t \rfloor}/\lfloor t \rfloor)_{t>0}$ and $(L(t\cdot)/t)_{t>0}$ are exponentially equivalent.*

Proof. Let $\varepsilon > 0$ and $r \ge 0$. Obviously,

$$\left\| \frac{Z_{\lfloor t \rfloor}}{\lfloor t \rfloor} - \frac{L(t\cdot)}{t} \right\|_\infty \le \left(\frac{1}{\lfloor t \rfloor} - \frac{1}{t} \right) \|Z_{\lfloor t \rfloor}\|_\infty + \left\| \frac{Z_{\lfloor t \rfloor}}{t} - \frac{L(t\cdot)}{t} \right\|_\infty =: A_t + B_t \tag{5.13}$$

We estimate $\mathbb{P}(A_t > \varepsilon)$ and $\mathbb{P}(B_t > \varepsilon)$ separately. As in the proof of Lemma 5.3 we find

$$\mathbb{P}(A_t > \varepsilon) \le \mathbb{P}\left(\sup_{0 \le k \le \lfloor t \rfloor} |L_k| > t(t-1)\varepsilon \right) \le 3 \exp\left(-t(t-1)\frac{\varepsilon}{3} \right) \beta_1^{\lfloor t \rfloor} \tag{5.14}$$

where $\beta_1 := \mathbb{E}e^{|L_1|}$. In order to estimate B_t we note that

$$\sup_{s \in [0,1]} |Z_{\lfloor t \rfloor}(s) - L(ts)| \le \sup_{0 \le k \le \lfloor t \rfloor} \sup_{u \in [0,2]} |L_{k+u} - L_k|$$

as $ts - \lfloor \lfloor t \rfloor s \rfloor \leq 2$. By the stationarity of the increments and Etemadi's inequality, cf. Corollary A.2, this implies

$$\mathbb{P}\left(B_t > \varepsilon\right) \leq \mathbb{P}\left(\sup_{0 \leq k \leq \lfloor t \rfloor} \sup_{u \in [0,2]} |L_{k+u} - L_k| > t\varepsilon \right) \leq 3(\lfloor t \rfloor + 1) \sup_{u \in [0,2]} \mathbb{P}\left(|L_u| > \frac{t\varepsilon}{3} \right).$$

If we set $\mu := \mathbb{E}L_1$, then we have for sufficiently large t

$$\mathbb{P}\left(|L_u| > \frac{t\varepsilon}{3} \right) \leq \mathbb{P}\left(|L_u - u\mu| > \frac{t\varepsilon}{6} \right) \qquad \text{for all} \quad u \in [0,2].$$

Since $(L_u - u\mu)_{u \geq 0}$ is a martingale, we know that $(e^{r|L_u - u\mu|})_{u \geq 0}$ is a submartingale. By Markov's inequality,

$$\sup_{u \in [0,2]} \mathbb{P}\left(|L_u - u\mu| > \frac{t\varepsilon}{6} \right) \leq e^{-tr\varepsilon/6} \mathbb{E} e^{r|L_2 - 2\mu|} =: \beta_2 e^{-tr\varepsilon/6}. \tag{5.15}$$

Combining (5.13), (5.14) and (5.15) yields

$$\begin{aligned} &\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}\left(\left\| \frac{Z_{\lfloor t \rfloor}}{\lfloor t \rfloor} - \frac{L(t\cdot)}{t} \right\|_\infty > 2\varepsilon \right) \\ &\qquad \leq \max\left\{ \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}(A_t > \varepsilon), \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}(B_t > \varepsilon) \right\} \leq -\frac{r\varepsilon}{6} \xrightarrow{r\to\infty} -\infty. \end{aligned}$$

This finishes the proof. □

Proof of Theorem 5.1. By Theorem 5.5, $(Z_{\lfloor t \rfloor]}/\lfloor t \rfloor)_{t>0}$ satisfies a large deviation principle in $(D[0,1], \|\cdot\|_\infty)$ with respect to $\mathcal{B}$ with good rate function I defined in (5.8). Since $(Z_{\lfloor t \rfloor]}/\lfloor t \rfloor)_{t>0}$ and $(L(t\cdot)/t)_{t>0}$ are exponentially equivalent, cf. Lemma 5.6, the claim follows by applying Corollary 4.3 and Theorem 5.7 below. □

5.7 Theorem *The good rate function I defined in* (5.8) *equals*

$$J(f) := \begin{cases} \int_0^1 \Psi^*(f'(s))\, ds, & f \in AC[0,1], f(0) = 0, \\ \infty, & \textit{otherwise}. \end{cases}$$

Proof. First, we show that $I(f) < \infty$ implies that f is absolutely continuous and $f(0) = 0$. Then, there exists $g \in L^1([0,1], \lambda|_{[0,1]})$ such that

$$f(t) = \int_0^t g(s)\, ds, \qquad t \in [0,1]. \tag{5.16}$$

Fix $\varepsilon > 0$, $0 < s_1 < t_1 \leq \ldots < s_n < t_n \leq 1$, and $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$. We define

$$\alpha(t) := \sum_{j=1}^n c_j \mathbb{1}_{[s_j, t_j)}(t), \qquad t \in [0,1]. \tag{5.17}$$

Obviously, $\alpha \in \mathrm{BV}[0,1] \cap D[0,1]$ and

$$\int_0^1 f\, d\alpha = \sum_{j=1}^n c_j (f(s_j) - f(t_j)). \tag{5.18}$$

Moreover, by the definition of α,

$$\begin{aligned}\int_0^1 \log \mathbb{E} e^{L_1(\alpha(1)-\alpha(s))}\,ds &= \sum_{j=1}^n \int_0^1 \log \mathbb{E} e^{-c_j L_1} \mathbb{1}_{[s_j,t_j)}(s)\,ds \\ &\leq \log \mathbb{E} e^{\|c\|_\infty |L_1|} \sum_{j=1}^n (t_j - s_j). \end{aligned} \tag{5.19}$$

From the very definition of the rate function I, cf. (5.8), we see

$$\int_0^1 f\,d\alpha \leq I(f) + \int_0^1 \log \mathbb{E} e^{L_1(\alpha(1)-\alpha(s))}\,ds.$$

Using (5.18) and (5.19), we find

$$\sum_{j=1}^n c_j (f(s_j) - f(t_j)) \leq I(f) + \log \mathbb{E} e^{|L_1|\|c\|_\infty} \sum_{j=1}^n (t_j - s_j).$$

In particular, for $c_j := r \operatorname{sgn}(f(s_j) - f(t_j))$, $r > 0$,

$$\sum_{j=1}^n |f(t_j) - f(s_j)| \leq \frac{I(f)}{r} + \frac{\log \mathbb{E} e^{|L_1| r}}{r} \sum_{j=1}^n (t_j - s_j).$$

Choosing $r > 0$ sufficiently large and $\delta > 0$ sufficiently small, we see that

$$\sum_{j=1}^n (t_j - s_j) < \delta \Rightarrow \sum_{j=1}^n |f(t_j) - f(s_j)| < \varepsilon,$$

i. e. f is absolutely continuous. A similar calculation shows

$$|f(t)| \leq \frac{I(f)}{r} + t \frac{\log \mathbb{E} e^{|L_1| r}}{r}.$$

Letting $t \to 0$ and $r \to \infty$ yields $f(0) = 0$. This proves (5.16). Now let $f \in D[0,1]$ be given by (5.16). By Lemma A.9,

$$\begin{aligned}\int_0^1 f\,d\alpha - \int_0^1 \Psi(\alpha(1) - \alpha(s))\,ds &= \int_0^1 \left[g(s)(\alpha(1) - \alpha(s)) - \Psi(\alpha(1) - \alpha(s))\right] ds \\ &\leq \int_0^1 \Psi^*(g(s))\,ds = \int_0^1 \Psi^*(f'(s))\,ds\end{aligned}$$

for any $\alpha \in \mathrm{BV}[0,1] \cap D[0,1]$. Thus, $I(f) \leq J(f)$. It remains to prove $I(f) \geq J(f)$ for $f \in AC[0,1]$, $f(0) = 0$. By the monotone convergence theorem, it suffices to show

$$\int_0^1 \Lambda_k(f'(s))\,ds \leq I(f)$$

where

$$\Lambda_k(x) := \sup_{|\alpha| \leq k} (\alpha x - \Psi(\alpha)), \qquad x \in \mathbb{R}, k \in \mathbb{N}.$$

Note that Λ_k is convex and locally bounded, hence continuous, see e. g. [35, Corollary 10.1.1]. Since

$$\sum_{j=0}^{n-1} \frac{f\left(\frac{j+1}{n}\right) - f\left(\frac{j}{n}\right)}{\frac{1}{n}} \mathbb{1}_{\left[\frac{j}{n}, \frac{j+1}{n}\right)}(s) \to f'(s) \qquad \text{a.s.}$$

we get from the dominated convergence theorem

$$\int_0^1 \Lambda_k(f'(s))\,ds = \lim_{n\to\infty}\sum_{j=0}^{n-1}\frac{1}{n}\Lambda_k\left(n\left[f\left(\frac{j+1}{n}\right)-f\left(\frac{j}{n}\right)\right]\right).$$

As $\alpha \mapsto \alpha x - \Psi(\alpha)$ is continuous, we can choose $|\alpha(x)| \leq k$ such that

$$\Lambda_k(x) = \alpha(x)x - \Psi(\alpha(x)).$$

Consequently, for suitable $\alpha_0^n,\ldots,\alpha_{n-1}^n$,

$$\begin{aligned}\int_0^1 \Lambda_k(f'(s))\,ds &= \lim_{n\to\infty}\sum_{j=0}^{n-1}\left[\alpha_j^n\left(f\left(\frac{j+1}{n}\right)-f\left(\frac{j}{n}\right)\right)-\frac{1}{n}\Psi(\alpha_j^n)\right]\\ &= \lim_{n\to\infty}\left(\int_0^1 f\,d\alpha^n - \int_0^1 \Psi(\alpha^n(1)-\alpha^n(s))\,ds\right) \leq I(f),\end{aligned}$$

where $\alpha^n \in \mathrm{BV}[0,1]\cap D[0,1]$, $n\in\mathbb{N}$, is a step function of the form (5.17). This finishes the proof. □

It is natural to ask whether there are other scalings than $S(t) = t$ which yield a large deviation principle. The following large deviation principle was stated by Mogulskii [31] without providing a detailed proof. We show that de Acosta's approach remains valid in this setting.

5.8 Theorem *Let $(L_t)_{t\geq 0}$ be a Lévy process such that $\mathbb{E}L_1 = 0$, $\mathbb{E}L_1^2 > 0$ and $\mathbb{E}e^{\lambda|L_1|} < \infty$ for all $\lambda \in (0,\lambda_0] \neq \emptyset$. If $S:[0,\infty)\to[0,\infty)$ is an increasing function such that*

$$\frac{S(t)}{t}\xrightarrow{t\to\infty}0 \quad \text{and} \quad \frac{S(t)}{\sqrt{t}}\xrightarrow{t\to\infty}\infty, \tag{5.20}$$

then $(L(t\cdot)/S(t))_{t\geq 0}$ satisfies a large deviation principle in $(D[0,1],\|\cdot\|_\infty)$ with respect to $\mathcal{B}$ with normalizing coefficient $S^2(t)/t$ and good rate function I given by

$$I(f) := \begin{cases}\frac{1}{2\mathbb{E}L_1^2}\int_0^1 f'(t)^2\,dt, & f\in AC[0,1], f(0)=0,\\ \infty, & \text{otherwise},\end{cases} \tag{5.21}$$

i. e.

$$\begin{aligned}\liminf_{t\to\infty}\frac{t}{S^2(t)}\log\mathbb{P}\left(\frac{L(t\cdot)}{S(t)}\in A\right) &\geq -\inf_{f\in A} I(f)\\ \limsup_{t\to\infty}\frac{t}{S^2(t)}\log\mathbb{P}\left(\frac{L(t\cdot)}{S(t)}\in B\right) &\leq -\inf_{f\in B} I(f)\end{aligned}$$

for any open set $A\in\mathcal{B}$ and closed set $B\in\mathcal{B}$.

Note that the rate function I coincides with the rate function of the scaled Brownian motion, cf. Example 5.10. In order to prove Theorem 5.8 we need the following lemma.

5.9 Lemma *Let $(L_t)_{t\geq 0}$ be as in Theorem 5.8. Then there exists $C \geq 0$ such that*

$$\mathbb{E}e^{\lambda L_t} \leq \exp\left[\left(\frac{1}{2}\mathbb{E}L_1^2\lambda^2 + C\lambda^3\right)t\right] \qquad \text{for all } \lambda \in (0,\lambda_0]. \tag{5.22}$$

In particular, for some constant C',

$$\mathbb{E}e^{\lambda L_t} \leq e^{C'\lambda^2 t} \qquad \text{for all } \lambda \in (0,\lambda_0]. \tag{5.23}$$

Proof. Let $0 \leq \lambda \leq \lambda_0$. First of all, $e^{\lambda|L_1|} \in L^1$ implies $e^{\lambda|L_t|} \in L^1$ for any $t \geq 0$, cf. [36, Theorem 25.17]. Moreover, $\mathbb{E}e^{\lambda L_t} = e^{t\Psi(\lambda)}$ where

$$\Psi(\lambda) := \frac{1}{2}\sigma^2\lambda^2 + \int_{\mathbb{R}\setminus\{0\}} (e^{\lambda y} - 1 - \lambda y)\,\nu(dy),$$

cf. [36, Theorem 25.17]. (Recall that $\mathbb{E}L_1 = 0$.) The integral appearing on the right-hand side is finite as $e^{\lambda|L_1|} \in L^1$, cf. [36, Theorem 25.17]. Using Taylor's formula, we find

$$\begin{aligned}\Psi(\lambda) &= \underbrace{\frac{1}{2}\left(\sigma^2 + \int_{\mathbb{R}\setminus\{0\}} y^2\,\nu(dy)\right)}_{\mathbb{E}L_1^2}\lambda^2 + \frac{1}{6}\lambda^3 \int_{\mathbb{R}\setminus\{0\}} e^{\lambda\xi(y)}y^3\,\nu(dy) \\ &\leq \frac{1}{2}\mathbb{E}L_1^2\lambda^2 + \frac{1}{6}\lambda^3 \int_{\mathbb{R}\setminus\{0\}} e^{\lambda_0|y|}|y|^3\,\nu(dy)\end{aligned}$$

for some intermediate value $\xi(y)$ between 0 and y. As $\mathbb{E}e^{\lambda_0|L_1|} < \infty$ the latter integral is finite, cf. [36, Theorem 25.3]. Consequently, the claim follows. □

Proof of Theorem 5.8. Essentially, we have to show that Lemma 5.3, Theorem 5.5 and Lemma 5.6 are satisfied for the discrete approximations $Z_n/S(n)$ of $L(t\cdot)/S(t)$. Without mentioning it explicitely, we pick up the notation from the the corresponding results. The remaining part of the proof goes through as in the proof of Theorem 5.1.

(i). Lemma 5.3: Obviously, $Z_n/S(n)$ is tight. In order prove the exponential tightness of $(Z_n/S(n))_{n\in\mathbb{N}}$, we modify the estimates of I_1 and I_2. By Etemadi's inequality and Markov's inequality,

$$\begin{aligned}I_1 = \mathbb{P}\left(\sup_{1\leq j\leq n}\left|\frac{L_j}{S(n)}\right| > r\right) &\leq 3\sup_{1\leq j\leq n}\left[\mathbb{P}\left(L_j > \frac{S(n)r}{3}\right) + \mathbb{P}\left(-L_j > \frac{S(n)r}{3}\right)\right] \\ &\leq 3\exp\left(-\frac{S(n)r}{3}\lambda(n)\right)\sup_{1\leq j\leq n}\left(\mathbb{E}e^{\lambda(n)L_j} + \mathbb{E}e^{-\lambda(n)L_j}\right)\end{aligned}$$

for any $\lambda(n) > 0$. For $\lambda(n) := S(n)/n$ we obtain from (5.23)

$$I_1 \leq 6\exp\left(-\frac{r - C}{3}\frac{S(n)^2}{n}\right)$$

for a constant $C > 0$ which does not depend on $n \in \mathbb{N}$ and $r \geq 0$. I_2 is treated in a similar way. Consequently, we get

$$\limsup_{n\to\infty} \frac{n}{S^2(n)} \log \mathbb{P}\left(d\left(\frac{Z_n}{S(n)}, T_m(K^m)\right) > \varepsilon\right) \xrightarrow{m\to\infty} -\infty.$$

This shows the exponential tightness of $(Z_n/S(n))_{n\in\mathbb{N}}$, cf. Lemma 5.2.

(ii). Theorem 5.5: For $\alpha \in \mathrm{BV}[0,1] \cap D[0,1]$ set

$$\Lambda(\alpha) := \lim_{n\to\infty} \frac{n}{S^2(n)} \log \mathbb{E} \exp\left(\frac{S(n)^2}{n} \langle \alpha, \frac{Z_n}{S(n)} \rangle \right).$$

First of all, we have to show that the limit exists. As in the proof of Theorem 5.5 we see

$$\log \mathbb{E} \exp\left(\frac{S(n)}{n} \langle \alpha, Z_n \rangle \right) = \sum_{j=1}^n \Psi\left(\frac{S(n)}{n} (\alpha(1) - \alpha(j/n)) \right).$$

where

$$\Psi(w) = \frac{1}{2}\sigma^2 w^2 + \int_{\mathbb{R}\setminus\{0\}} (e^{wy} - 1 - wy)\, \nu(dy), \qquad |w| \le \lambda_0,$$

is the logarithmic moment generating function of L_1. Observe that the expectation values are finite for sufficiently large $n \in \mathbb{N}$ since $S(n)/n \to 0$ as $n \to \infty$. Consequently,

$$\begin{aligned}
&\frac{n}{S^2(n)} \log \mathbb{E} \exp\left(\frac{S(n)}{n} \langle \alpha, Z_n \rangle \right) \\
&\quad = \frac{1}{2}\sigma^2 \frac{1}{n} \sum_{j=1}^n (\alpha(1) - \alpha(j/n))^2 \\
&\quad + \frac{n}{S(n)^2} \sum_{j=1}^n \int_{\mathbb{R}\setminus\{0\}} \left(e^{\frac{S(n)}{n}(\alpha(1)-\alpha(j/n))y} - 1 - \frac{S(n)}{n}(\alpha(1) - \alpha(j/n))y \right) \nu(dy) \\
&\quad =: J_1 + J_2
\end{aligned}$$

Since J_1 can be seen as a Riemann sum, we find

$$J_1 \xrightarrow{n\to\infty} \frac{1}{2}\sigma^2 \int_0^1 (\alpha(1) - \alpha(s))^2 \, ds.$$

Using Taylor's formula and the growth condition (5.20), it is not difficult to show that

$$J_2 \xrightarrow{n\to\infty} \frac{1}{2} \left(\int_{\mathbb{R}\setminus\{0\}} y^2 \, \nu(dy) \right) \int_0^1 (\alpha(1) - \alpha(s))^2 \, ds.$$

Indeed: By Taylor's formula,

$$\begin{aligned}
J_2 &= \frac{1}{2}\frac{1}{n} \left(\int_{\mathbb{R}\setminus\{0\}} y^2 \, \nu(dy) \right) \sum_{j=1}^n (\alpha(1) - \alpha(j/n))^2 \\
&\quad + \frac{S(n)}{n^2} \left(\int_{\mathbb{R}\setminus\{0\}} e^{\frac{S(n)}{n}(\alpha(1)-\alpha(j/n))x} x^3 \, \nu(dx) \right) \sum_{j=1}^n (\alpha(1) - \alpha(j/n))^3 \\
&=: J_2^1 + J_2^2
\end{aligned}$$

Obviously, it suffices to show $J_2^2 \to 0$. Since

$$|J_2^2| \le \frac{S(n)}{n} 8 \|\alpha\|_\infty^3 \int_{\mathbb{R}\setminus\{0\}} (\mathbb{1}_{(-\infty,0)} + e^{\lambda_0 x}) |x|^3 \, \nu(dx)$$

the claim follows with (5.20). This proves

$$\Lambda(\alpha) = \frac{1}{2}\mathbb{E}L_1^2 \int_0^1 (\alpha(1) - \alpha(s))^2 \, ds \qquad \text{for all} \quad \alpha \in \mathrm{BV}[0,1] \cap D[0,1].$$

In view of Theorem 5.5 and Theorem 5.7, it follows easily that Λ is $C[0,1]$-Gâteaux differentiable and its Legendre transform is given by

$$\Lambda^*(f) = \begin{cases} \frac{1}{2\mathbb{E}L_1^2}\int_0^1 f'(t)^2\,dt, & f \in AC[0,1], f(0) = 0, \\ \infty, & \text{otherwise.} \end{cases}$$

(iii). Lemma 5.6: As in step (i) the estimates of A_t and B_t are modified appropriately using Lemma 5.9. □

Before we proceed with two applications, let us give some examples.

5.10 Example (i). Both Brownian motion and Poisson process have exponential moments of all orders. Therefore, the corresponding scaled processes satisfy a large deviation principle in $(D[0,1], \|\cdot\|_\infty)$ as $t \to \infty$ relative to the scaling function $S(t) = t$ as well as any scaling function meeting the growth condition (5.20). For the scaled Brownian motion the domain of the rate function equals the *Cameron-Martin space* $\mathcal{H}^1 := \{f \in AC[0,1]; f(0) = 0, \int_0^1 f'(t)^2\,dt < \infty\}$ and

$$I(f) = \frac{1}{2}\int_0^1 f'(t)^2\,dt \qquad \text{for all } \; f \in \mathcal{H}^1.$$

By the scaling property,

$$\frac{B(ts)}{t} \sim \frac{B(s)}{\sqrt{t}}, \qquad s,t > 0,$$

and it is therefore not difficult to see that $(B(\cdot)/\sqrt{t})_{t>0}$ satisfies the same large deviation principle as $(B(t\cdot)/t)_{t>0}$ for any Brownian motion $(B_s)_{s\geq 0}$.

(ii). Let $(L_t)_{t\geq 0}$ be a *Gamma process*, i.e. a Lévy process such that

$$L_t \sim \frac{\alpha^t}{\Gamma(t)} x^{t-1} e^{-\alpha x} \mathbb{1}_{(0,\infty)}(x)\,dx, \qquad t > 0,$$

for some $\alpha > 0$. Then $\mathbb{E}e^{\lambda|L_1|} < \infty$ for $0 \leq \lambda < \alpha$. Consequently, by Theorem 5.8, $(L(t\cdot)/S(t))_{t>0}$ obeys a large deviation principle relative to any scaling function S satisfying (5.20). Lynch and Sethuraman proved that a large deviation principle holds for $(L(t\cdot)/t)_{t>0}$ with respect to the weak*-topology, cf. [29, Example 6.2].

(iii). Neither Theorem 5.1 nor Theorem 5.8 does apply to Lévy processes with infinite moments of order k for some $k \in \mathbb{N}$. In particular, α-stable processes are not covered.

Theorem 5.8 shows that, under the growth condition (5.20) on the scaling function S, scaled Lévy processes share the rate function with the scaled Brownian motion. Therefore, it is a natural guess that some asymptotics of the Brownian motion carry over to Lévy processes. In fact, using Theorem 5.8, it is not difficult to prove the analogue of the (functional) law of iterated logarithm. The assumption on the exponential moments is due to our approach; there are more general statements, see e. g. [36, Proposition 48.9].

5.11 Theorem (Law of iterated logarithm) *Let* $(L_t)_{t\geq 0}$ *be a Lévy process such that* $\mathbb{E}e^{\lambda|L_1|} < \infty$ *for all* $\lambda \in (0, \lambda_0] \neq \varnothing$. *Then*

$$\limsup_{t\to\infty} \frac{L_t - t\mathbb{E}L_1}{\sqrt{2t\log\log t}} = \sqrt{\mathbb{V}L_1} \qquad a.s. \tag{5.24}$$

$$\liminf_{t\to\infty} \frac{L_t - t\mathbb{E}L_1}{\sqrt{2t\log\log t}} = -\sqrt{\mathbb{V}L_1} \qquad a.s. \tag{5.25}$$

Proof. Clearly, it suffices to consider the case $\mathbb{E}L_1 = 0$, $\mathbb{V}L_1 = 1$. Moreover, applying (5.24) to $-L$ yields (5.25); it remains to prove (5.24). The idea of the proof is taken from [38, Corollary 11.2] where the result is shown for Brownian motion. Set $S(s) := \sqrt{2s\log\log s}$, and pick $q > 1$. Note that S satisfies the growth condition (5.20).

(i). Let $\varepsilon > 0$. Consider $B := \{f \in D[0,1]; f(0) = 0, \sup_{s\in[0,1]} f(s) \geq 1 + \varepsilon\}$. Obviously, $B \in \mathcal{B}$ is closed (with respect to the uniform topology)[2] and therefore, by Theorem 5.8,

$$\mathbb{P}\left(\sup_{s\leq q^n} \frac{L(s)}{S(q^n)} \geq (1+\varepsilon)\right) = \mathbb{P}\left(\frac{L(q^n\cdot)}{S(q^n)} \in B\right) \leq \exp\left(-\frac{S(q^n)^2}{q^n}\left(\inf_{f\in B} I(f) - \varepsilon\right)\right) \tag{5.26}$$

for $n \geq n_0 = n_0(\varepsilon)$ sufficiently large where I denotes the rate function (5.21). For $f \in B$ such that $I(f) < \infty$ we have by Jensen's inequality

$$(1+\varepsilon)^2 \leq \sup_{s\in[0,1]} |f(s) - f(0)|^2 \leq \int_0^1 |f'(s)|^2\,ds = 2I(f).$$

On the other hand, if we set $t \mapsto f(t) := (1+\varepsilon)t$, then $f \in B$ and $I(f) = 1/2(1+\varepsilon)^2$. Hence, $\inf_{f\in B} I(f) = 1/2(1+\varepsilon)^2$. By (5.26) and the definition of S, we conclude

$$\sum_{n=1}^\infty \mathbb{P}\left(\sup_{s\leq q^n} \frac{L(s)}{S(q^n)} \geq (1+\varepsilon)\right) \leq \sum_{n=1}^\infty \frac{1}{(n\log q)^{2\inf_{f\in B} I(f) - 2\varepsilon}} < \infty.$$

Now Borel-Cantelli's lemma shows

$$\limsup_{n\to\infty} \frac{\sup_{s\leq q^n} L_s}{S(q^n)} \leq (1+\varepsilon) \qquad \text{a.s.}$$

Clearly, any $t > 1$ is contained in an interval of the form $[q^{n-1}, q^n]$; thus

$$\frac{L_t}{S(t)} \leq \frac{\sup_{s\leq q^n} L_s}{S(q^n)} \frac{S(q^n)}{S(q^{n-1})}$$

as S is increasing. Letting $q \to 1$ and $\varepsilon \to 0$ along countable sequences yields

$$\limsup_{t\to\infty} \frac{L(t)}{S(t)} \leq (1+\varepsilon)\sqrt{q} \xrightarrow{q\to 1, \varepsilon\to 0} 1 \qquad \text{a.s.}$$

(ii). By the stationarity of increments, we have[3]

$$\mathbb{P}\left(\frac{L(q^n) - L(q^{n-1})}{S(q^n - q^{n-1})} > (1-\varepsilon)\right) \geq \mathbb{P}\left(\frac{L((q^n - q^{n-1})\cdot)}{S(q^n - q^{n-1})} \in A\right)$$

[2] Revised version: Reformulated.

[3] Revised version: Corrected misprint.

where $A := \{f \in D[0,1]; f(1) > 1-\varepsilon\}$[4]. Obviously, $A \in \mathcal{B}$ is open. Applying Theorem 5.8, we obtain

$$\mathbb{P}\left(\frac{L(q^n)-L(q^{n-1})}{S(q^n-q^{n-1})} > (1-\varepsilon)\right) \geq \exp\left(-\frac{S(q^n-q^{n-1})^2}{q^n-q^{n-1}}(I(f)+\varepsilon/4)\right)$$

for $n \geq n_0 = n_0(\varepsilon)$ sufficiently large and any $f \in A$. If we choose $f(t) := (1-\varepsilon/2)t$, then $f \in A$ and $I(f) = 1/2(1-\varepsilon/2)^2$. Thus,

$$\mathbb{P}\left(\frac{L(q^n)-L(q^{n-1})}{S(q^n-q^{n-1})} > (1-\varepsilon)\right) \geq \frac{1}{(\log(q^n(1-q^{-1})))^{1-\varepsilon/2+\varepsilon^2/4}}.$$

Since the increments are independent, Borel Cantelli's lemma yields for $\varepsilon > 0$ sufficiently small

$$L(q^n) \geq (1-\varepsilon)S(q^n-q^{n-1}) + L(q^{n-1})$$

for infinitely many $n \in \mathbb{N}$. Applying the first part of this proof to the Lévy process $-L$, we find

$$-L(q^{n-1}) \leq 2S(q^{n-1}) \leq \frac{2}{\sqrt{q}}S(q^n) \qquad \text{a.s.}$$

for $n \geq n_1$ sufficiently large. Hence,

$$\frac{L(q^n)}{S(q^n)} \geq (1-\varepsilon)\frac{S(q^n-q^{n-1})}{S(q^n)} - \frac{2}{\sqrt{q}}$$

Letting $q \to \infty$ and $\varepsilon \to 0$ along countable sequences, we conclude

$$\limsup_{t\to\infty} \frac{L(t)}{S(t)} \geq \limsup_{n\to\infty} \frac{L(q^n)}{S(q^n)} \geq 1 \qquad \text{a.s.} \qquad \square$$

5.12 Theorem (Functional law of iterated logarithm) *Let $(L_t)_{t\geq 0}$ be a Lévy process such that $\mathbb{E}L_t = 0$ and $\mathbb{E}e^{\lambda|L_1|} < \infty$ for $\lambda \in (0,\lambda_0] \neq \varnothing$. The set*[5]

$$\left\{\frac{L(t\cdot,\omega)}{\sqrt{2t\log\log t}}; t > e\right\}$$

is relatively compact in $(D[0,1], \|\cdot\|_\infty)$ a.s., and the set of limit points (as $t \to \infty$) is for almost all $\omega \in \Omega$ given by the sublevel set $\Phi(\mathbb{E}(L_1^2)/2)$ of the good rate function I defined in (5.21).

Remark Obviously, it suffices to consider the case $\mathbb{E}L_1^2 = 1$ (otherwise we apply the result to $L_t/\sqrt{\mathbb{E}L_1^2}$). If $\mathbb{E}L_1^2 = 0$ the statement is obvious.

5.13 Lemma *The set of all limit points $\mathcal{L}(\omega)$ satisfies $\mathcal{L}(\omega) \subseteq \Phi(1/2)$.*

Proof. Since $\Phi(1/2) = \bigcap_{r>0}\Phi(1/2+r)$, it suffices to show $\mathcal{L}(\omega) \subseteq \Phi(1/2+r)$. Set $S(t) := \sqrt{2t\log\log t}$ and $Z_t := \frac{L(t\cdot)}{S(t)}$. Fix $q > 1$, $\delta > 0$. By Theorem 5.8 and (L2'), we find for $\gamma < r$

$$\mathbb{P}\left(d(Z_{q^n}, \Phi(1/2+r)) > \delta\right) \leq \exp\left(-2\log\log q^n\left(\frac{1}{2}+r-\gamma\right)\right).$$

[4] Revised version: Corrected misprint.

[5] Revised version: Corrected misprint.

for $n \geq n_0(\gamma,\omega)$ sufficiently large. Thus, by Borel-Cantelli's theorem,

$$d(Z_{q^n}(\cdot,\omega),\Phi(1/2+r)) \leq \delta$$

for $n \geq n_0$. It remains to fill the gaps in the sequence $(q^n)_{n\in\mathbb{N}}$. Note that

$$\begin{aligned}\sup_{q^{n-1}\leq t\leq q^n} \|Z_t - Z_{q^n}\|_\infty &= \sup_{q^{n-1}\leq t\leq q^n} \sup_{0\leq r\leq 1} \left| \frac{L(rt)}{S(t)} - \frac{L(rq^n)}{S(q^n)} \right| \\ &\leq \underbrace{\sup_{q^{n-1}\leq t\leq q^n} \sup_{0\leq r\leq 1} \frac{|L(rt)-L(rq^n)|}{S(q^n)}}_{=:A_n} + \underbrace{\sup_{q^{n-1}\leq t\leq q^n} \sup_{0\leq r\leq 1} \frac{|L(rt)|}{S(q^n)} \left| \frac{S(q^n)}{S(t)} - 1 \right|}_{=:B_n}.\end{aligned}$$

From

$$B_n \leq \sup_{s\leq q^n} \frac{|L(s)|}{S(q^n)} \left| \frac{S(q^n)}{S(q^{n-1})} - 1 \right|$$

it follows easily from the law of iterated logarithm, cf. Theorem 5.11, that $B_n \leq \frac{\delta}{2}$ for $n \geq n_1(\omega,\delta,q)$ sufficiently large. In order to estimate A_n we note that

$$\mathbb{P}\left(A_n > \frac{\delta}{2}\right) = \mathbb{P}\left(\sup_{q^{-1}\leq t\leq 1} \sup_{0\leq r\leq 1} \frac{|L(q^n rt) - L(q^n r)|}{S(q^n)} > \frac{\delta}{2} \right) \leq \mathbb{P}\left(\frac{L(q^n \cdot)}{S(q^n)} \in B \right)$$

where

$$B := \left\{ f \in D[0,1];\ \sup_{q^{-1}\leq t\leq 1} \sup_{0\leq r\leq 1} |f(rt) - f(r)| \geq \frac{\delta}{2} \right\}.$$

It is not difficult to see that $\inf_{f\in B} I(f) = \frac{\delta^2}{8}\frac{q}{q-1}$, cf. [38, Lemma 12.16]. Therefore, by Theorem 5.8,

$$\mathbb{P}\left(A_n > \frac{\delta}{2}\right) \leq \exp\left(-2\log\log q^n \cdot \left(\frac{\delta^2}{8}\frac{q}{q-1} - \gamma \right) \right)$$

for any $\gamma > 0$ and $n \geq n_2(\gamma)$ sufficiently large. If $q > 1$ is close to 1, this implies $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n > \delta/2) < \infty$. Applying Borel-Cantelli's theorem yields

$$\sup_{q^{n-1}\leq t\leq q^n} \|Z_t - Z_{q^n}\|_\infty \leq \delta$$

for $n \geq n_3(\omega)$ sufficiently large. Finally,

$$d\big(Z_s(\cdot,\omega),\Phi(\tfrac{1}{2}+r)\big) \leq \|Z_s(\cdot,\omega) - Z_{q^n}(\cdot,\omega)\|_\infty + d\big(Z_{q^n}(\cdot,\omega),\Phi(\tfrac{1}{2}+r)\big) \leq 2\delta$$

for $s \geq s_0(\omega) := q^{N(\omega)+1}$, $N(\omega) := \max_{j=0,1,2,3} n_j(\omega)$. Since $\Phi(1/2+r)$ is closed, this proves the claim. □

5.14 Lemma $\mathcal{L}(\omega) \supseteq \Phi(1/2)$.

Proof. Since the sublevel sets are compact, we have $\operatorname{cl}\left(\bigcup_{r<1/2}\Phi(r)\right) \subseteq \Phi(1/2)$. On the other hand, any $f \in \Phi(1/2)$ can be approximated by $(1-\varepsilon)f \in \Phi((1-\varepsilon)^2)$, $\varepsilon > 0$. Therefore, $\operatorname{cl}\left(\bigcup_{r<1/2}\Phi(r)\right) = \Phi(1/2)$. Consequently, it suffices to show that for any $r < 1/2$, $\varepsilon > 0$, $f \in \Phi(r)$, there is a. s. a sequence $s_n = s_n(\omega) \to \infty$ such that

$$\limsup_{n\to\infty} \|Z_{s_n} - f\|_\infty \leq \varepsilon.$$

We set $s_n := q^n$ for some $q > 1$. Obviously,

$$\|Z_{s_n} - f\|_\infty \le \sup_{q^{-1}\le t\le 1}\left|\frac{L(ts_n) - L(s_{n-1})}{S(s_n)} - f(t)\right| + \frac{|L(s_{n-1})|}{S(s_n)} + \sup_{t\le q^{-1}}|f(t)| + \sup_{t\le q^{-1}}\frac{|L(ts_n)|}{S(s_n)}.$$

We estimate the terms separately. By the stationarity of the increments, we have

$$\begin{aligned}\mathbb{P}(A_n) := \mathbb{P}\left(\sup_{q^{-1}\le t\le 1}\left|\frac{L(ts_n) - L(s_{n-1})}{S(s_n)} - f(t)\right| < \frac{\varepsilon}{4}\right) &= \mathbb{P}\left(\sup_{q^{-1}\le t\le 1}\left|\frac{L(tq^n - q^{n-1})}{S(q^n)} - f(t)\right| < \frac{\varepsilon}{4}\right)\\ &= \mathbb{P}\left(\sup_{0\le t\le 1-q^{-1}}\left|\frac{L(q^n t)}{S(q^n)} - f(t+q^{-1})\right| < \frac{\varepsilon}{4}\right)\\ &\ge \mathbb{P}\left(\frac{L(q^n\cdot)}{S(q^n)} \in B\left(g, \frac{\varepsilon}{8}\right)\right)\end{aligned}$$

where

$$g(t) := \begin{cases} f(t) - f(q^{-1}), & 0 \le t \le 1 - q^{-1},\\ f(1) - f(q^{-1}), & 1 - q^{-1} \le t \le 1,\end{cases}$$

and $q > 1$ is sufficiently large such that $|f(q^{-1})| \le \varepsilon/8$. Obviously, $I(g) < \frac{1}{2}$ and therefore we conclude by Theorem 5.8 that $\sum_{n\in\mathbb{N}}\mathbb{P}(A_n) = \infty$. Taking a subsequence, if necessary, we obtain by applying Borel-Cantelli's theorem,

$$\limsup_{n\to\infty}\sup_{q^{-1}\le t\le 1}\left|\frac{L(ts_n) - L(s_{n-1})}{S(s_n)} - f(t)\right| \le \frac{\varepsilon}{4}.$$

By Lemma 5.9 (for $\lambda(n) = S(q^{n-1})/q^{n-1}$) and Markov's inequality, we find

$$\mathbb{P}\left(\left|\frac{L(q^{n-1})}{S(q^n)}\right| > \frac{\varepsilon}{4}\right) \le 2\exp\left(\log\log q^{n-1}\left(-\sqrt{q}\frac{\varepsilon}{4} + C\right)\right).$$

Similarly, by Etemadi's inequality,

$$\mathbb{P}\left(\sup_{0\le t\le q^{-1}}\left|\frac{L(tq^n)}{S(q^n)}\right| > \frac{\varepsilon}{4}\right) \le 6\exp\left(\log\log q^{n-1}\left(-\sqrt{q}\frac{\varepsilon}{12} + C\right)\right).$$

Moreover, by Hölder's inequality,

$$\sup_{t\le q^{-1}}|f(t)| \le \int_0^{1/q}|f'(s)|\,ds \le \frac{\sqrt{2r}}{\sqrt{q}} < \frac{1}{\sqrt{q}}.$$

For $q > 1$ sufficiently large, we find by Borel-Cantelli's lemma

$$\limsup_{n\to\infty}\left(\frac{|L(s_{n-1})|}{S(s_n)} + \sup_{t\le q^{-1}}|f(t)| + \sup_{t\le q^{-1}}\frac{|L(ts_n)|}{S(s_n)}\right) < \frac{3}{4}\varepsilon.$$

This finishes the proof. $\square$

6

From Lévy to Lévy-Type Processes

One natural way to generalize the concept of Lévy processes are Lévy-type processes. Roughly speaking, a Lévy-type process is a time-homogeneous Markov process which locally resembles a Lévy process. In order to give a precise definition we have to introduce some notions which are closely connected with (time-homogeneous) Markov processes. For a survey on Lévy-type processes we refer the reader to [5]. Throughout this chapter $(\Omega, \mathcal{A}, \mathbb{P})$ denotes a probability space.

6.1 Definition Let $(X_t, \mathcal{F}_t)_{t\geq 0}$ be an adapted stochastic process and $(\mathbb{P}^x)_{x\in\mathbb{R}}$ a family of probability measures on $(\Omega, \mathcal{A})$. $(X_t, \mathcal{F}_t)_{t\geq 0}$ is a *(time-homogeneous) Markov process* if

$$\mathbb{E}^x(f(X_t) \mid \mathcal{F}_s) = \mathbb{E}^{X_s} f(X_{t-s}) \quad \mathbb{P}^x - \text{a.s.} \qquad \text{for all } f \in C_b(\mathbb{R}), x \in \mathbb{R}, s \leq t. \tag{6.1}$$

Equation (6.1) is also called *Markov property*. The *transition semigroup* $(T_t)_{\geq 0}$ of the Markov process $(X_t, \mathcal{F}_t)_{t\geq 0}$ is defined by

$$T_t f(x) := \mathbb{E}^x f(X_t), \qquad f \in \mathcal{B}_b(\mathbb{R}), t \geq 0, x \in \mathbb{R}.$$

We associate the *generator* $A : \operatorname{dom} A \to C_\infty(\mathbb{R})$ with $(X_t)_{t\geq 0}$[1],

$$\begin{aligned} \operatorname{dom} A &:= \left\{ f \in C_\infty(\mathbb{R}); \exists g \in C_\infty(\mathbb{R}) : \lim_{t\to 0} \left\| \frac{T_t f - f}{t} - g \right\|_\infty = 0 \right\} \\ Af &:= \lim_{t\to 0} \frac{T_t f - f}{t}. \end{aligned}$$

Using the Markov property (6.1), it is not difficult to see that $(T_t)_{t\geq 0}$ defines indeed a *semigroup*; that is $T_t T_s = T_{t+s}$ for all $s, t \geq 0$. For some basic properties of the transition semigroup and the generator see e. g. [38, Chapter 7].

6.2 Definition A Markov process[2] $(X_t, \mathcal{F}_t)_{t\geq 0}$ is called a *Lévy-type process with symbol q* if

$$x \mapsto q(x, \xi) := -\lim_{t\to 0} \frac{\mathbb{E}^x e^{\imath\xi(X_t - x)} - 1}{t}$$

defines a continuous function for each $\xi \in \mathbb{R}$ and $q(x, \xi)$ is of the form

$$q(x, \xi) = -\imath\gamma(x)\xi + \frac{1}{2}a(x)\xi^2 + \int_{\mathbb{R}\setminus\{0\}} (1 - e^{\imath y\xi} + \imath y\xi \mathbb{1}_{|y|\leq 1})\, N(x, dy)$$

[1] Revised version: Corrected misprint.

[2] More precisely: a *Feller process*. For the sake of simplicity we omit this detail here.

where, for fixed $x \in \mathbb{R}$, $(\gamma(x), a(x), N(x,\cdot))$ is a Lévy triplet, i.e. $\gamma(x) \in \mathbb{R}$, $a(x) \geq 0$ and $N(x,\cdot)$ is a measure on $(\mathbb{R}\backslash\{0\}, \mathcal{B}(\mathbb{R}\backslash\{0\}))$ satisfying $\int_{\mathbb{R}\backslash\{0\}} (y^2 \wedge 1) N(x,dy) < \infty$.

6.3 Example Let $(L_t)_{t\geq 0}$ be a Lévy process with Lévy triplet (γ, σ^2, ν). The Lévy-Khinchine formula (5.1) implies

$$-\lim_{t\to 0} \frac{\mathbb{E}^x e^{\imath\xi(L_t - x)} - 1}{t} = -\imath\gamma\xi + \frac{1}{2}\sigma^2\xi^2 + \int_{\mathbb{R}\backslash\{0\}} (1 - e^{\imath y\xi} + \imath y\xi \mathbb{1}_{|y|\leq 1})\,\nu(dy).$$

Moreover, $(L_t)_{t\geq 0}$ is a Markov process with respect to the canonical filtration, see [5, Theorem 2.6] for more details. Consequently, $(L_t)_{t\geq 0}$ is a Lévy-type process, and its symbol coincides with the characteristic exponent. In particular, the symbol q does not depend on the variable x. This indicates that Lévy processes are homogeneous in space.

Another typical example for Lévy-type processes are solutions of *SDEs driven by a Lévy process*, i.e. SDEs of the form

$$dX_t = f(X_{t-})\,dL_t \tag{6.2}$$

where $(L_t)_{t\geq 0}$ is a Lévy process. Our standard references for stochastic integration and SDEs are Ikeda-Watanabe [24] and Protter [32]. The next result is compiled from [32, Theorem V.7].

6.4 Theorem *Let $f : \mathbb{R} \to \mathbb{R}^n$ be bounded and $(L_t)_{t\geq 0}$ be an n-dimensional Lévy process. Suppose that f is* locally Lipschitz continuous*, i.e. for any $R > 0$ there exists $L = L(R) > 0$ such that*

$$|f(x) - f(y)| \leq L\,|x - y| \qquad \textit{for all} \;\; x, y \in B[0,R].$$

Then there exists a unique solution $(X_t)_{t\geq 0}$ of the SDE

$$dX_t = f(X_{t-})\,dL_t, \qquad X_0 = x \in \mathbb{R}. \tag{6.3}$$

Note that Theorem 6.4 entails in particular existence and uniqueness of solutions for SDEs of the form

$$dX_t = b(X_{t-})\,dt + \sigma(X_{t-})\,dB_t + \eta(X_{t-})\,dL_t \tag{6.4}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion and $(L_t)_{t\geq 0}$ an independent Lévy process. This follows simply from the fact that $(t, B_t, L_t)_{t\geq 0}$ is a Lévy process. We call b *drift coefficient* and σ *diffusion coefficient*.

In order to consider solutions of SDEs as Lévy-type processes we need to overcome a minor technical difficulty: we have to enlarge the underlying probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Define

$$\bar{\Omega} := \mathbb{R} \times \Omega \qquad \bar{\mathcal{A}} := \mathcal{B}(\mathbb{R}) \otimes \mathcal{A} \qquad \mathbb{P}^x := \delta_x \otimes \mathbb{P}.$$

Any random variable X on $(\Omega, \mathcal{A})$ can be extended to $(\bar{\Omega}, \bar{\mathcal{A}})$ by setting $X(x,\omega) := X(\omega)$. We further define a process $(X_t)_{t\geq 0}$ on $(\bar{\Omega}, \bar{\mathcal{A}})$ by $X_t(x,\omega) := X_t^x(\omega)$ where $(X_t^x)_{t\geq 0}$ is the unique solution of (6.3). Clearly,

$$X_t = x + \int_0^t f(X_{s-})\,dL_s \quad \mathbb{P}^x - \text{a.s.}$$

In abuse of notation we call $(X_t)_{t\geq 0}$ unique solution of (6.2). The following theorem is taken from [39, Section 3.1].

6.5 Theorem *Let $f : \mathbb{R} \to \mathbb{R}^n$ be bounded and (locally) Lipschitz continuous and $(L_t)_{t\geq 0}$ be an n-dimensional Lévy process with symbol ψ. Then the unique solution $(X_t)_{t\geq 0}$ of the SDE* (6.2) *is a Lévy-type process with symbol $q(x,\xi) = \psi(f(x)\xi)$, $x, \xi \in \mathbb{R}$.*

6.6 Example Let $b, \sigma, \eta : \mathbb{R} \to \mathbb{R}$ be bounded and (locally) Lipschitz continuous. If $(B_t)_{t\geq 0}$ is a Brownian motion and $(L_t)_{t\geq 0}$ an independent Lévy process with symbol ψ, then the symbol q of the unique solution of (6.4) is given by

$$q(x,\xi) = -b(x)\xi + \frac{1}{2}\sigma^2(x)\xi^2 + \psi(\eta(x)\xi).$$

7

Large Deviations for Lévy-Driven SDEs

In this chapter we consider large deviation results for solutions of SDEs of the form

$$dX_t^\varepsilon = f(X_{t-}^\varepsilon)\, dL_t^\varepsilon \tag{7.1}$$

where $L_t^\varepsilon := \varepsilon L_{t/\varepsilon}$, $\varepsilon > 0$, is a scaled Lévy process. They are a special case of so-called Freidlin-Wentzell results. Wentzell [40] studied families of time-homogeneous Markov processes $(X^\varepsilon)_{\varepsilon>0}$ with generators of the form

$$A^\varepsilon f(x) = b(x)f'(x) + \frac{\varepsilon}{2}a(x)f''(x) + \frac{1}{\varepsilon}\int_{\mathbb{R}\setminus\{0\}} \big(f(x+\varepsilon y) - f(x) - \varepsilon y f'(x)\big)\, N(x, dy) \tag{7.2}$$

for $f \in C_c^\infty(\mathbb{R})$. Speaking in terms of symbols, cf. Definition 6.2, this corresponds to

$$q^\varepsilon(x,\xi) = -\imath\, b(x)\xi + \frac{\varepsilon}{2}a(x)\xi^2 + \frac{1}{\varepsilon}\int_{\mathbb{R}\setminus\{0\}} (1 - e^{\imath\varepsilon y\xi} + \imath\varepsilon y\xi)\, N(x,dy),$$

cf. [5, Corollary 2.23]. We have seen in Theorem 6.5 and Example 6.6 that the symbols of solutions of (7.1) are indeed of this form.

The original proof of Wentzell is based on a change of measure (for the large deviation lower bound) and an approximation of the solutions by polygons (for the large deviation upper bound). Since both are typical large deviation techniques, we present this approach in Section 7.1 following the monograph [18] by Freidlin and Wentzell. Let us remark that the (quite restrictive) assumptions have been relaxed since then. In particular for SDEs driven by Brownian motion there exists rich literature, see e. g. Azencott [1] and Baldi-Chaleyat-Maurel [3] (ε-dependent coefficients), Cutland [7] (time-dependent coefficients), Millet-Sanz-Nualart [30] (anticipating SDEs), and Kulik-Soboleva [26] (discontinuous drift coefficient). The most general extension has been obtained by Feng-Kurtz [17, Section 10.3].

Liptser and Puhalskii [28] studied large deviations for SDEs driven by Brownian motion and random measures; this comprises SDEs driven by Lévy processes. Their results cover SDEs of a quite general form: the coefficients may be time-dependent functionals of the past of the process X^ε. Later on we will see that the assumptions formulated by Freidlin and Wentzell [18] are hard to check; it is therefore worth mentioning that Liptser and Puhalskii give sufficient conditions for their assumptions in terms of the coefficients of the SDE. More recently, de Acosta [9] has developed a generalization of the Gärtner-Ellis approach presented in Chapter 3 which works in a non-convex framework. As an application, a large deviation principle for SDEs driven by Brownian motion and random measures is proved. Dembo and Zeitouni [10] obtained a large deviation result for SDEs driven by Brownian motion using an extension of the contraction principle, cf. Section 4.1. A generalisation to SDEs driven by Lévy processes is presented in Section 7.2. We will close this section with a glimpse into a more general setting; namely, large deviations for stochastic integrals with respect to semimartingales, cf. Theorem 7.12. For a thorough discussion we refer the reader to Garcia [22] and Ganguly [19].

Throughout this chapter $(\Omega, \mathcal{A}, \mathbb{P})$ denotes a complete probability space, $(\bar{\Omega}, \bar{\mathcal{A}}, \mathbb{P}^x)$ the corresponding enlarged probability space, cf. Chapter 6, and $(B_t)_{t\geq 0}$ a Brownian motion on $(\Omega, \mathcal{A}, \mathbb{P})$.

7.1 Solutions as Markov Processes

Here we give a purely probabilistic proof of a large deviation principle for solutions of (7.1) following the presentation in the monograph [18] by Freidlin and Wentzell. We restrict ourselves to solutions of (7.1) – instead of considering time-homogeneous Markov processes with generators of the form (7.2) – and re-write the statement in terms of symbols.

7.1 Theorem *Let $f : \mathbb{R} \to \mathbb{R}^n$ be bounded and locally Lipschitz continuous. Let $(L_t)_{t\geq 0}$ be a Lévy process with Lévy triplet (γ, σ^2, ν) and symbol ψ such that $\mathbb{E}e^{\lambda |L_1|} < \infty$ for all $\lambda \in \mathbb{R}$. Denote by $(X_t^\varepsilon)_{t\geq 0}$ the unique solution of the SDE*

$$dX_t = f(X_{t-})\, dL_t^\varepsilon$$

where $L_t^\varepsilon := \varepsilon L_{t/\varepsilon}$ is the scaled Lévy process. The symbol of the solution of the SDE

$$dX_t = f(X_{t-})\, dL_t$$

is given by $q(x,\xi) = \psi(f(x)\xi)$. Set $Q(x,\xi) := q(x, -\imath\xi)$, and denote by $Q^*(x,\cdot)$ the Legendre transform of the convex function $Q(x,\cdot)$. Suppose that the following conditions are satisfied.

(S1) *$Q^*(x,\beta) < \infty$ for all $x, \beta \in \mathbb{R}$; for any $R > 0$ there exist constants $C_1, C_2 > 0$ such that*

$$Q^*(x,\beta) + \left|\frac{\partial}{\partial\beta} Q^*(x,\beta)\right| \leq C_1 \quad \text{and} \quad \frac{\partial^2}{\partial\beta^2} Q^*(x,\beta) > C_2 \qquad \text{for all } \; x \in \mathbb{R}, |\beta| \leq R.$$

(S2) *Continuity condition:*

$$\Delta Q^*(\delta) := \sup_{|x-y|<\delta} \sup_{\beta\in\mathbb{R}} \frac{Q^*(x,\beta) - Q^*(y,\beta)}{1 + Q^*(y,\beta)} \xrightarrow{\delta\to 0} 0.$$

Then the family $(X^\varepsilon)_{\varepsilon>0}$ of processes on $(\bar{\Omega}, \bar{\mathcal{A}}, \mathbb{P}^x)$ satisfies a large deviation principle in $(D[0,1], \|\cdot\|_\infty)$ with good rate function

$$I_x(\varphi) := I(\varphi) := \begin{cases} \int_0^1 Q^*(\varphi(t), \varphi'(t))\, dt, & \varphi \in AC[0,1], \varphi(0) = x, \\ \infty, & \text{otherwise.} \end{cases} \tag{7.3}$$

Remarks (i). Since $(L_t)_{t \geq 0}$ has exponential moments, its symbol ψ is twice differentiable. This implies in particular that $\xi \mapsto Q(x,\xi)$ is twice differentiable and so is its Legendre transform $\beta \mapsto Q^*(x,\beta)$, see e. g. [35]. Therefore, we do not have to *assume* differentiability in (S1).

(ii). Using the boundedness of f and the definition of Q, it is not difficult to see that the following condition is automatically satisfied:

(S3) There exists $\bar{Q} : \mathbb{R} \to [0,\infty)$ such that $\bar{Q}(0) = 0$ and $Q(x,\xi) \leq \bar{Q}(\xi)$ for all $x, \xi \in \mathbb{R}$.

(iii). The large deviation lower and upper bound hold uniformly in the initial point $x \in \mathbb{R}$, see [18, Theorem 5.2.1]. Moreover, we can replace the Lipschitz continuity by uniform continuity if we *assume* that there exists a solution to the SDE $dX_t = f(X_{t-})\, dL_t^\varepsilon$ for each $\varepsilon > 0$ (we do not need uniqueness).

(iv). The assumptions on the Legendre transform Q^* are quite restrictive. In fact, Theorem 7.1 does not even apply to the scaled Poisson process – i. e. $f = 1$, $(L_t)_{t \geq 0}$ Poisson process – since the Legendre transform Q^* of the logarithmic moment generating function Q of the Poisson distribution does not satisfy $Q^* < \infty$.

Theorem 7.1 claims that I, defined in (7.3), is a good rate function; we defer the proof to the appendix, see Lemma A.8. To keep notation simple, we restrict ourselves to the case $n = 1$ and assume that $(L_t)_{t \geq 0}$ admits the Lévy-Itô decomposition

$$L_t = t + B_t + \int_0^t \int z\, \tilde{N}(dz, ds) =: t + B_t + J_t, \qquad t \geq 0, \tag{7.4}$$

cf. Chapter 5. Then Q equals

$$Q(x,\xi) = f(x)\xi + \frac{1}{2} f(x)^2 \xi^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{f(x)\xi y} - 1 - f(x)\xi y \right) d\nu(y). \tag{7.5}$$

We prove the large deviation lower and upper bound separately. In order to obtain the large deviation lower bound, we need the following lemma. Let us remark that Freidlin and Wentzell use a different argumentation based on the fact that $(X_t^\varepsilon)_{t \geq 0}$ is a (non-homogeneous) Markov process with respect to the family of measures $(\mathbb{Q}^{x,\varepsilon})_{x \in \mathbb{R}}$ defined in Lemma 7.2.

7.2 Lemma *(i). For any bounded Borel-measurable function $\alpha : [0,1] \times \mathbb{R} \to \mathbb{R}$ and $\varepsilon > 0$ the process*

$$M_t^{\alpha,\varepsilon} := \exp\left(\frac{1}{\varepsilon} \int_0^t \alpha(r-, X_{r-}^\varepsilon)\, dX_r^\varepsilon - \frac{1}{\varepsilon} \int_0^t Q(X_r^\varepsilon, \alpha(r, X_r^\varepsilon))\, dr \right), \quad t \in [0,1]$$

is a $\mathbb{P}^x$-martingale with respect to the canonical filtration $\mathcal{F}_t^\varepsilon := \sigma\{B_s^\varepsilon, J_s^\varepsilon; s \leq t\}$. In particular, $\mathbb{E}^x M_t^{\alpha,\varepsilon} = 1$.

(ii). Let $\varphi \in C[0,1]$ be piecewise differentiable, and set $\alpha(t,y) := \frac{\partial}{\partial\beta}Q^*(y,\varphi'(t))$. Then $(X_t^\varepsilon - \varphi_t)_{t\in[0,1]}$ is a martingale with respect to the probability measure $d\mathbb{Q}^{x,\varepsilon} := M_1^{\alpha,\varepsilon}\,d\mathbb{P}^x$, and its variance equals

$$\mathbb{E}_{\mathbb{Q}^{x,\varepsilon}}\big((X_t^\varepsilon - \varphi_t)^2\big) = \varepsilon\,\mathbb{E}_{\mathbb{Q}^{x,\varepsilon}}\left(\int_0^1 \frac{\partial^2}{\partial\xi^2}Q(X_r^\varepsilon, \alpha(r, X_r^\varepsilon))\,dr\right). \tag{7.6}$$

Proof. Denote by $\tilde{N}_\varepsilon$ the compensated jump counting measure of the Lévy process L_t^ε. It follows from Itô's formula that $(M_t^{\alpha,\varepsilon})_{t\in[0,1]}$ satisfies

$$dM_s^{\alpha,\varepsilon} = \frac{1}{\varepsilon}M_{s-}^{\alpha,\varepsilon}f(X_{s-})\alpha(s-,X_{s-})\,dB_s^\varepsilon + M_{s-}^{\alpha,\varepsilon}(e^{\varepsilon^{-1}\alpha(s-,X_{s-})f(X_{s-})z} - 1)\,d\tilde{N}_\varepsilon(dz,ds). \tag{7.7}$$

This shows that $(M_t^{\alpha,\varepsilon})_{t\in[0,1]}$ is a positive local martingale, hence $\mathbb{E}^x M_t^{\alpha,\varepsilon} < \infty$, cf. [38, Proposition 17.3]. In order to deduce that $(M_t^{\alpha,\varepsilon})_{t\in[0,1]}$ is a martingale, we show that $\sup_{t\in[0,1]}(M_t^{\alpha,\varepsilon})^2 \in L^1(\mathbb{P}^x)$. Set

$$N_t^{\alpha,\varepsilon} := \exp\left(\varepsilon^{-1}\int_0^t \alpha(r-,X_{r-}^\varepsilon)\,d\left(X_r^\varepsilon - \int_0^r f(X_{s-}^\varepsilon)\,ds\right)\right), \qquad t \in [0,1].$$

Note that the exponent is a martingale; as a stochastic integral with respect to a martingale. Since $\mathbb{E}^x M_t^{2\alpha,\varepsilon} < \infty$, condition (S3) implies $N_t^{\alpha,\varepsilon} \in L^2(\mathbb{P}^x)$. Therefore, by Jensen's inequality, $(N_t^{\alpha,\varepsilon})_{t\in[0,1]}$ is a submartingale. Applying Doob's maximal inequality yields

$$\begin{aligned}\mathbb{E}\left(\sup_{t\in[0,1]}(M_t^{\alpha,\varepsilon})^2\right) &= \mathbb{E}\left[\sup_{t\in[0,1]}(N_t^{\alpha,\varepsilon})^2\exp\left(\frac{2}{\varepsilon}\int_0^t \big(f(X_{r-}^\varepsilon) - Q(X_r^\varepsilon,\alpha(r,X_r^\varepsilon))\big)\,dr\right)\right]\\ &\le 4\exp\left(\frac{2}{\varepsilon}\|f\|_\infty + \frac{2}{\varepsilon}\sup_{|\xi|\le\|\alpha\|_\infty}\bar{Q}(\xi)\right)\mathbb{E}\big((N_1^{\alpha,\varepsilon})^2\big) < \infty.\end{aligned}$$

This proves (i). Now let $\alpha(t,y) := \frac{\partial}{\partial\beta}Q^*(y,\varphi'(t))$. For brevity, we suppress the index ε and set $M_t := M_t^\alpha$. First of all, by (i), $\mathbb{Q}^x$ is indeed a probability measure. Recall that for any measure of the form $d\mathbb{Q} = \beta\,d\mathbb{P}$ where $\beta > 0$ is a random variable, we have

$$\mathbb{E}_\mathbb{Q}(X \mid \mathcal{F}) = \frac{\mathbb{E}(X\beta \mid \mathcal{F})}{\mathbb{E}(\beta \mid \mathcal{F})},$$

see e. g. [38, pp. 255-6]. Hence, since $(M_t)_{t\in[0,1]}$ is a $\mathbb{P}^x$-martingale, we get

$$\mathbb{E}_{\mathbb{Q}^x}(X_t - \varphi_t \mid \mathcal{F}_s) = \frac{\mathbb{E}^x((X_t-\varphi_t)M_1 \mid \mathcal{F}_s)}{\mathbb{E}^x(M_1 \mid \mathcal{F}_s)} = \frac{\mathbb{E}^x((X_t-\varphi_t)M_t \mid \mathcal{F}_s)}{M_s} \qquad \text{for all} \;\; s \le t \le 1.$$

This means that it suffices to show that $((X_t - \varphi_t)M_t)_{t\in[0,1]}$ is a $\mathbb{P}^x$-martingale. Using (7.7) and

$$d(X_t - \varphi_t) = (f(X_{s-}) - \varphi'(s))\,ds + f(X_{s-})\,dB_s^\varepsilon + f(X_{s-})\,dJ_s^\varepsilon,$$

cf. (7.4), we conclude from Itô's formula

$$\begin{aligned}
&(X_t - \varphi_t)M_t - (x - \varphi_0) \\
&= \int_0^t f(X_{s-})M_{s-}\, dB_s^\varepsilon + \frac{1}{\varepsilon}\int_0^t (X_{s-} - \varphi_s)M_{s-} f(X_{s-})\alpha(s-, X_{s-})\, dB_s^\varepsilon \\
&+ \int_0^t \int \Big[(X_{s-} - \varphi_{s-} + z f(X_{s-}))M_{s-} e^{\varepsilon^{-1}\alpha(s-,X_{s-})f(X_{s-})z} - (X_{s-} - \varphi_{s-})M_{s-} \Big] d\tilde{N}_\varepsilon(dz, ds) \\
&+ \int_0^t M_s \Big[\underbrace{f(X_s) + f(X_s)^2\alpha(s, X_s) + \int_{\mathbb{R}\setminus\{0\}} f(X_s) y (e^{\alpha(s,X_s)f(X_s)y} - 1)\, \nu(dy)}_{\overset{(7.5)}{=} \frac{\partial}{\partial\xi} Q(X_s, \alpha(s, X_s))} - \varphi'(s) \Big] ds.
\end{aligned}$$

By (A.4), we have $\frac{\partial}{\partial\xi} Q(X_s, \alpha(s, X_s)) - \varphi'(s) = 0$. This shows that $((X_t - \varphi_t)M_t)_{t\in[0,1]}$ is a local $\mathbb{P}^x$-martingale. In fact, it is a martingale; we do not dwell on this technical issue.[1] A similar calculation yields the desired expression for the variance. $\square$

Proof of the lower bound. We have to show that the large deviation lower bound (L1') holds for each $\delta > 0$ and $\varphi \in D[0, 1]$ for which $I(\varphi) < \infty$. For any such φ there exists a sequence of polygons $(\ell_k)_{k\in\mathbb{N}}$ such that $\ell_k \to \varphi$ and $\limsup_{k\to\infty} I(\ell_k) \le I(\varphi)$, cf. Lemma A.8(ii). Therefore, it suffices to prove (L1') for piecewise differentiable functions. Let $\varphi \in C[0, 1]$ be piecewise differentiable, and set $\alpha(t, y) := \frac{\partial}{\partial\beta} Q^*(y, \varphi'(t))$. For constants $C, D > 0$ we define

$$A^\varepsilon := \left\{ \|X^\varepsilon - \varphi\|_\infty \le 4C^{\frac{1}{2}}\varepsilon^{\frac{1}{2}} \right\}, \qquad B^\varepsilon := \left\{ \left| \int_0^1 \alpha(t-, X_{t-}^\varepsilon)\, d(X_t^\varepsilon - \varphi_t) \right| \le 4D^{\frac{1}{2}}\varepsilon^{\frac{1}{2}} \right\}.$$

Clearly, $A^\varepsilon \cap B^\varepsilon \subseteq \{\|X^\varepsilon - \varphi\|_\infty < \delta\}$ for ε sufficiently small. We choose

$$C := \sup_{t,y} \frac{\partial^2}{\partial\xi^2} Q(y, \alpha(t, y)) \qquad\qquad D := \sup_{t,y} \frac{\partial^2}{\partial\xi^2} Q(y, \alpha(t, y))\alpha^2(t, y).$$

Let $M_t^{\alpha,\varepsilon}$ and $\mathbb{Q}^{x,\varepsilon}$ be as in Lemma 7.2. By Lemma 7.2(ii), $(X_t^\varepsilon - \varphi_t)_{t\in[0,1]}$ is a $\mathbb{Q}^{x,\varepsilon}$-martingale. Therefore, we obtain by applying Markov's inequality and Doob's maximal inequality

$$\mathbb{Q}^{x,\varepsilon}((A^\varepsilon)^c) \le \frac{1}{16\varepsilon C}\mathbb{E}_{\mathbb{Q}^{x,\varepsilon}}\left(\sup_{0\le t\le 1} |X_t^\varepsilon - \varphi_t|^2 \right) \le \frac{1}{4\varepsilon C}\mathbb{E}_{\mathbb{Q}^{x,\varepsilon}}(|X_1^\varepsilon - \varphi_1|^2) \overset{(7.6)}{\le} \frac{1}{4}.$$

Similarly, we find $\mathbb{Q}^{x,\varepsilon}((B^\varepsilon)^c) \le \frac{1}{4}$. Hence, $\mathbb{Q}^{x,\varepsilon}(A^\varepsilon \cap B^\varepsilon) \ge \frac{1}{2}$. Since $\mathbb{Q}^{x,\varepsilon}/\mathbb{P}^x$ has the strictly positive density $M_1^{\alpha,\varepsilon}$, we get

$$\mathbb{P}^x(\|X^\varepsilon - \varphi\|_\infty < \delta) \ge \mathbb{E}_{\mathbb{Q}^{x,\varepsilon}}\left(\mathbb{1}_{A^\varepsilon\cap B^\varepsilon}\frac{1}{M_1^{\alpha,\varepsilon}} \right) \ge \frac{1}{2}\inf_{\omega\in A^\varepsilon\cap B^\varepsilon}\frac{1}{M_1^{\alpha,\varepsilon}(\omega)}. \tag{7.8}$$

By virtue of our choice of α, Lemma A.7 shows

$$\begin{aligned}
\frac{1}{M_1^{\alpha,\varepsilon}} &= \exp\left[-\frac{1}{\varepsilon}\left(\int_0^1 \alpha(t-, X_{t-}^\varepsilon)\, d(X_t^\varepsilon - \varphi_t) + \int_0^1 (\alpha(t, X_t^\varepsilon)\varphi'(t) - Q(X_t^\varepsilon, \alpha(t, X_t^\varepsilon)))\, dt \right) \right] \\
&\overset{(A.3)}{=} \exp\left[-\frac{1}{\varepsilon}\left(\int_0^1 \alpha(t-, X_{t-}^\varepsilon)\, d(X_t^\varepsilon - \varphi_t) + \int_0^1 Q^*(X_t^\varepsilon, \varphi'(t))\, dt \right) \right].
\end{aligned} \tag{7.9}$$

[1] Apply Doob's maximal inequality (for the martingale M_t) and the Burkholder-Davis-Gundy inequality (for X_t) in order to show that the integrands are properly integrable.

On $A^\varepsilon \cap B^\varepsilon$ we have

$$\left|\int_0^1 \alpha(t-, X_{t-}^\varepsilon)\, d(X_t^\varepsilon - \varphi_t)\right| \le 4D^{\frac{1}{2}}\varepsilon^{\frac{1}{2}} \tag{7.10}$$

and

$$\begin{aligned}\int_0^1 Q^*(X_t^\varepsilon, \varphi'(t))\, dt &= \int_0^1 Q^*(X_t^\varepsilon, \varphi'(t)) - Q^*(\varphi(t), \varphi'(t))\, dt + \int_0^1 Q^*(\varphi(t), \varphi'(t))\, dt \\ &\le \Delta Q^*(4C^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}) + (1 + \Delta Q^*(4C^{\frac{1}{2}}\varepsilon^{\frac{1}{2}})) \int_0^1 Q^*(\varphi(t), \varphi'(t))\, dt. \quad (7.11)\end{aligned}$$

Recall that $\Delta Q^*(4C^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}) \to 0$ as $\varepsilon \to 0$, cf. (S2). Therefore, combining (7.8), (7.9), (7.10) and (7.11) yields

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}^x(\|X^\varepsilon - \varphi\|_\infty < \delta) \ge -\int_0^1 Q^*(\varphi(t), \varphi'(t))\, dt = -I(\varphi). \qquad \square$$

It remains to prove the large deviation upper bound. The following lemma comes in handy; it shows in particular how to approximate the Legendre transform $Q^*(x, \cdot)$ by polygons.

7.3 Lemma *(i). $Q^*(x,\cdot)|_{(-\infty, f(x))}$ is strictly decreasing, $Q^*(x,\cdot)|_{(f(x),\infty)}$ is strictly increasing, and $Q^*(x, f(x)) = 0$ for all $x \in \mathbb{R}$.*

(ii). For any $\varepsilon, \delta > 0$ and $x \in \mathbb{R}$ there exist $\xi_1, \dots, \xi_m$, $m = m(x) \le M$, such that the polygon

$$P(x,\beta) := \max_{i=1,\dots,m} (\xi_i \beta - Q(x, \xi_i))$$

satisfies $0 < Q^*(x,\beta) - P(x,\beta) < \varepsilon$ for $|\beta| < \delta$.

(iii). For each $\delta > 0$ and $\xi \in \mathbb{R}$,

$$Q(y, (1 + \Delta Q^*(\delta))^{-1}\xi) - (1 + \Delta Q^*(\delta))^{-1} Q(x,\xi) \le \frac{\Delta Q^*(\delta)}{1 + \Delta Q^*(\delta)} \qquad \text{for all } |x - y| < \delta.$$

Proof. (i). Fix $x \in \mathbb{R}$. By definition, we have $Q(x, 0) = 0$ and therefore $Q^*(x,\beta) \ge 0$ for any $\beta \in \mathbb{R}$. Thus, by (7.5), $Q^*(x, f(x)) = 0$, i. e. $Q^*(x,\cdot)$ attains its minimum at $\beta = f(x)$. Moreover, by (S1), $\frac{\partial^2}{\partial\beta^2} Q^*(x,\beta) > 0$; this implies that $Q^*(x,\cdot)$ is strictly convex. This proves the claimed monotonicity.

(ii). Let $k \in \mathbb{N}$ such that $Q^*(x,\cdot)([-\delta,\delta]) \subseteq [-k\varepsilon, k\varepsilon]$. By (i), there exist (at most) two points β_j, β_{-j} such that $|\beta_{\pm j}| < \delta$ and $Q^*(x, \beta_{\pm j}) = j\varepsilon$, $j = 1, \dots, k$; for $j = 0$, there exists (at most) one point $\beta_0 \in \mathbb{R}$, $|\beta_0| < \delta$, such that $Q^*(x, 0) = 0$, see Figure 7.1. Define

$$P(x,\beta) := \max_{j=-k,\dots,k} (Q^*(x,\beta_j) + \frac{\partial}{\partial\beta} Q^*(x,\beta_j)(\beta - \beta_j)).$$

Because of the monotonicity and convexity of $Q^*(x,\cdot)$, it is not difficult to see that $0 < Q^*(x,\beta) - P(x,\beta) < \varepsilon$. If we set $\xi_j := \frac{\partial}{\partial\beta} Q^*(x,\beta_j)$, $j = -k, \dots, k$, then by Lemma A.7

$$Q^*(x,\beta_j) + \frac{\partial}{\partial\beta} Q^*(x,\beta_j)(\beta - \beta_j) = \xi_j\beta - Q(x,\xi_j).$$

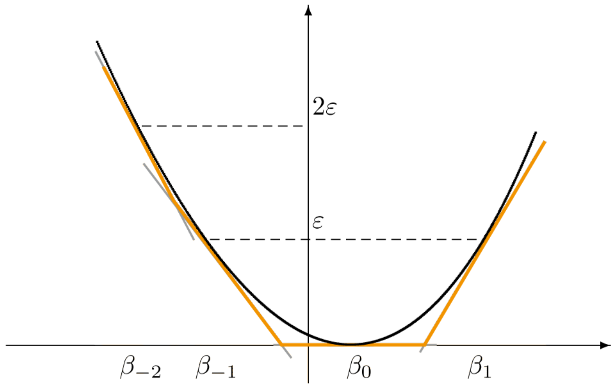

Figure 7.1: The Legendre transform $Q^*(x,\cdot)$ and the approximating polygon $P(x,\cdot)$

This finishes the proof. Note that we need at most

$$M := \left\lfloor 2\varepsilon^{-1} \sup_{x\in\mathbb{R}} \sup_{|\beta|\leq\delta} Q^*(x,\beta) \right\rfloor < \infty$$

supporting points β_j for *any* $x \in \mathbb{R}$.

(iii). For brevity, we write $\Delta^* := \Delta Q^*(\delta)$. By the definition of Δ^*, cf. (S2),

$$Q^*(x,\beta) - (1+\Delta^*)Q^*(y,\beta) \leq \Delta^* \qquad \text{for all} \quad |x-y| < \delta.$$

Hence,

$$\inf_{\alpha\in\mathbb{R}} (\xi(\beta-\alpha) + Q^*(x,\alpha) - (1+\Delta^*)Q^*(y,\beta)) \leq \Delta^*$$

for fixed $\beta \in \mathbb{R}$. Applying the duality lemma, cf. Theorem A.3, yields

$$\begin{aligned} &Q(y,(1+\Delta^*)^{-1}\xi) - (1+\Delta^*)^{-1}Q(x,\xi) \\ &\quad = \sup_{\beta\in\mathbb{R}} \inf_{\alpha\in\mathbb{R}} \big((1+\Delta^*)^{-1}\xi(\beta-\alpha) - Q^*(y,\beta) + (1+\Delta^*)^{-1}Q^*(x,\alpha)\big) \leq \frac{\Delta^*}{1+\Delta^*}. \square \end{aligned}$$

Proof of the upper bound. By Lemma 2.3, it suffices to show[2]

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}^x(d(X^\varepsilon, \Phi(r)) \geq \delta) \leq -r \qquad \text{for all} \quad r \geq 0,$$

where $\Phi(r)$ is the sublevel set of the rate function I. For $n \in \mathbb{N}$ we denote by $\Pi_n(X)(\cdot,\omega)$ the linear interpolation of $X(\cdot,\omega)$ on the grid $t_i := i\Delta t := i/n$, $i = 0,\ldots,n$, $n \in \mathbb{N}$. Fix $r \geq 0$ and $\delta, \chi > 0$. Choose $\varrho < \delta/2$ sufficiently small such that $\Delta Q^*(\varrho) < \chi$. Obviously,

$$\mathbb{P}^x(\|X^\varepsilon - \Phi(r)\|_\infty \geq \delta) \leq \mathbb{P}^x\left(\bigcup_{i=0}^{n-1} (A_i^\varepsilon)^c\right) + \mathbb{P}^x\left(\bigcap_{i=0}^{n-1} A_i^\varepsilon \cap \{\Pi_n(X^\varepsilon) \notin \Phi(r)\}\right) =: I_1 + I_2$$

for

$$A_i^\varepsilon := \left\{ \sup_{0\leq t\leq\Delta t} |X_{t_i+t}^\varepsilon - X_{t_i}^\varepsilon| < \varrho \right\}.$$

[2] Revised version: Corrected misprint.

We estimate the terms separately. Using the Markov property and a well-known inequality for Markov processes, see e. g. [9, Lemma 5.1] or [12, Lemma 6.3], we find[3]

$$\begin{aligned}I_1 &\leq n \sup_{y\in\mathbb{R}} \mathbb{P}^y \left(\sup_{0\leq t\leq \Delta t} |X_t^\varepsilon - y| \geq \varrho \right) \leq 2n \sup_{y\in\mathbb{R}} \sup_{0\leq t\leq\Delta t} \mathbb{P}^y \left(|X_t^\varepsilon - y| \geq \frac{\varrho}{2} \right) \\ &\leq 2n \exp\left(-\frac{C}{\varepsilon}\frac{\varrho}{2} \right) \sup_{y\in\mathbb{R}} \sup_{t\leq \Delta t} \mathbb{E}^y \left[\exp\left(\frac{C}{\varepsilon}(X_t^\varepsilon - y) \right) + \exp\left(-\frac{C}{\varepsilon}(X_t^\varepsilon - y) \right) \right]. \qquad (7.12)\end{aligned}$$

By Lemma 7.2(i) (applied to the constant function $\alpha = \pm C$) and (S3),

$$\begin{aligned}\mathbb{E}^y e^{\pm \frac{C}{\varepsilon}(X_t^\varepsilon - y)} &= \mathbb{E}^y \left[\exp\left(\pm \frac{C}{\varepsilon}(X_t^\varepsilon - y) - \frac{1}{\varepsilon}\int_0^t Q(X_s^\varepsilon, \pm C)\, ds \right) \exp\left(\frac{1}{\varepsilon}\int_0^t Q(X_s^\varepsilon, \pm C)\, ds \right) \right] \\ &\leq \exp\left(\varepsilon^{-1} \Delta t (\overline{Q}(C) \vee \overline{Q}(-C)) \right). \qquad (7.13)\end{aligned}$$

for any constant $C > 0$ and $t \leq \Delta t$. Plugging this estimate into (7.12) yields[4]

$$I_1 \leq 4n \exp\left(-\frac{\varrho}{2}\frac{C}{\varepsilon} + \frac{\Delta t}{\varepsilon}(\overline{Q}(C) \vee \overline{Q}(-C)) \right).$$

For $C = 4r/\varrho$ und $n \in \mathbb{N}$ sufficiently large such that $\Delta t(\overline{Q}(C) \vee \overline{Q}(-C)) \leq r$, we conclude

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}^x \left(\bigcup_{i=0}^{n-1} (A_i^\varepsilon)^c \right) \leq -r.$$

It remains to estimate I_2. Since $\Pi_n(X^\varepsilon)$ is a piecewise linear function, we can calculate $I(\Pi_n(X^\varepsilon))$ explicitely. As $\Delta Q^*(\varrho) < \chi$, cf. (S2), we obtain

$$\begin{aligned}I(\Pi_n(X^\varepsilon)) &= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} Q^* \left(\Pi_n(X^\varepsilon)(t), \frac{X_{t_{i+1}}^\varepsilon - X_{t_i}^\varepsilon}{\Delta t} \right) dt \\ &\leq \chi + \sum_{i=0}^{n-1} \Delta t(1+\chi) Q^* \left(\Pi_n(X^\varepsilon)(t_i), \frac{X_{t_{i+1}}^\varepsilon - X_{t_i}^\varepsilon}{\Delta t} \right) \\ &= \chi + \sum_{i=0}^{n-1} \Delta t(1+\chi) Q^* \left(X_{t_i}^\varepsilon, \frac{X_{t_{i+1}}^\varepsilon - X_{t_i}^\varepsilon}{\Delta t} \right) \qquad (7.14)\end{aligned}$$

on $\bigcap_{i=0}^{n-1} A_i^\varepsilon$. Applying again Markov's inequality, we find

$$\begin{aligned}I_2 &= \mathbb{P}^x \left(\bigcap_{i=0}^{n-1} A_i^\varepsilon \cap \{ I(\Pi_n(X^\varepsilon)) \geq r \} \right) \\ &\leq \exp\left(-\varepsilon^{-1}(1+\chi)^{-2} r \right) \mathbb{E}^x \left[\exp\left(\varepsilon^{-1}(1+\chi)^{-2} I(\Pi_n(X^\varepsilon)) \right) \prod_{i=0}^{n-1} \mathbb{1}_{A_i^\varepsilon} \right] \\ &\overset{(7.14)}{\leq} \exp\left(-\varepsilon^{-1}(1+\chi)^{-2}(r-\chi) \right) \mathbb{E}^x \left[\prod_{i=0}^{n-1} \mathbb{1}_{A_i^\varepsilon} \exp\left(\varepsilon^{-1}(1+\chi)^{-1} \Delta t Q^* \left(X_{t_i}^\varepsilon, \frac{X_{t_{i+1}}^\varepsilon - X_{t_i}^\varepsilon}{\Delta t} \right) \right) \right].\end{aligned}$$

Since $(X_t^\varepsilon)_{t\geq 0}$ is a Markov process, it is not difficult to see that

$$I_2 \leq \exp\left(-\varepsilon^{-1}(1+\chi)^{-2}(r-\chi) \right) \sup_{y\in\mathbb{R}} I_3(\varepsilon, y)^n \qquad (7.15)$$

[3]Revised version: Corrected misprints.

[4]Revised version: Corrected misprint.

where

$$I_3 := I_3(\varepsilon, y) := \mathbb{E}^y\left[\mathbb{1}_{A_0^\varepsilon} \exp\left(\varepsilon^{-1}(1+\chi)^{-1}\Delta t Q^*\left(y, \frac{X_{\Delta t}^\varepsilon - y}{\Delta t}\right)\right)\right].$$

Now let $\xi_i = \xi_i(y)$, $1 \leq i \leq m(y) \leq M$, as in Lemma 7.3(ii) (applied for $\varepsilon \hat{=} \chi$, $\delta \hat{=} \varrho/\Delta t$). Set $\xi_i(y) = 0$ for $m(y) < i \leq M$. Then,

$$\begin{aligned} I_3 &\leq \mathbb{E}^y\left\{\mathbb{1}_{A_0^\varepsilon} \exp\left[\varepsilon^{-1}(1+\chi)^{-1}\Delta t\left(\chi + \max_{i=1,\dots,M}\left(\xi_i \frac{X_{\Delta t}^\varepsilon - y}{\Delta t} - Q(y,\xi_i)\right)\right)\right]\right\} \\ &\leq e^{\varepsilon^{-1}(1+\chi)^{-1}\chi\Delta t} \sum_{i=1}^M \mathbb{E}^y\left\{\mathbb{1}_{A_0^\varepsilon} \exp\left[\varepsilon^{-1}(1+\chi)^{-1}\left(\xi_i(X_{\Delta t}^\varepsilon - y) - \Delta t Q(y,\xi_i)\right)\right]\right\}. \end{aligned}$$

Using a similar argumentation as in (7.13) and applying Lemma 7.3(iii), we get

$$\begin{aligned} I_3 &\leq e^{\varepsilon^{-1}(1+\chi)^{-1}\chi\Delta t} \sum_{i=1}^M \sup_{\omega \in A_0^\varepsilon} \exp\left(\varepsilon^{-1}\int_0^{\Delta t} Q(X_s^\varepsilon(\omega), (1+\chi)^{-1}\xi_i)\, ds - \varepsilon^{-1}(1+\chi)^{-1}\Delta t Q(y,\xi_i)\right) \\ &\leq M \exp\left(2\varepsilon^{-1}(1+\chi)^{-1}\chi\Delta t\right). \end{aligned}$$

Combining the estimates gives

$$\log I_2 \leq n \log M - \varepsilon^{-1}(1+\chi)^{-1}\left((r-\chi) + 2\chi\right).$$

Finally, we conclude

$$\limsup_{\varepsilon \to 0} \varepsilon \log I_2(\varepsilon) \leq -(1+\chi)^{-1}\left((r-\chi)+2\chi\right) \xrightarrow{\chi \to 0} -r. \qquad \square$$

Let us mention the following two important corollaries.

7.4 Corollary *Let $b, \sigma, \eta : \mathbb{R} \to \mathbb{R}$ be bounded, locally Lipschitz continuous functions and $(L_t)_{t \geq 0}$ a Lévy process with Lévy triplet $(0,0,\nu)$ such that $\mathbb{E}e^{\lambda|L_1|} < \infty$ for all $\lambda \geq 0$. If*

$$Q(x,\xi) = b(x)\xi + \frac{1}{2}\sigma^2(x)\xi^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{y\eta(x)\xi} - 1 - y\eta(x)\xi\mathbb{1}_{|y|\leq 1}\right) d\nu(y)$$

satisfies (S1) and (S2), then the family $(X^\varepsilon)_{\varepsilon > 0}$ of solutions on $(\bar{\Omega}, \bar{A}, \mathbb{P}^x)$,

$$dX_t^\varepsilon = b(X_{t-}^\varepsilon)\, dt + \sqrt{\varepsilon}\sigma(X_{t-}^\varepsilon)\, dB_t + \eta(X_{t-}^\varepsilon)\, dL_t^\varepsilon, \tag{7.16}$$

obeys a large deviation principle in $(D[0,1], \|\cdot\|_\infty)$ with good rate function I defined in (7.3).

Proof. Since $(L_t)_{t\geq 0}$ is a non-Gaussian Lévy process, the processes $(B_t)_{t \geq 0}$ and $(L_t)_{t\geq 0}$ are independent, see e. g. [24, Theorem II.6.3]. Therefore, the claim follows by applying Theorem 7.1 to $(t, B_t, L_t)_{t \geq 0}$. (As indicated in Example 5.10, we may replace B_t^ε by $\sqrt{\varepsilon}B_t$.) $\square$

In general, the Legendre transform $Q^*(x,\cdot)$ cannot be calculated explicitly[5], and consequently it is difficult to formulate sufficient conditions for (S1) and (S2) in terms of the coefficients of the SDE. Corollary 7.5 gives a sufficient condition for SDEs driven by Brownian motion.

[5] Revised version: Corrected misprint.

7.5 Corollary *Let $b, \sigma : \mathbb{R} \to \mathbb{R}$ be bounded, locally Lipschitz continuous functions such that $\inf_{x \in \mathbb{R}} \sigma(x) > 0$. Assume that $(X_t^\varepsilon)_{t \in [0,1]}$ is a solution of the SDE*

$$dX_t^\varepsilon = b(X_t^\varepsilon)\,dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)\,dB_t, \qquad X_0^\varepsilon = x.$$

Then $(X^\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle in $(C[0,1], \|\cdot\|_\infty)$ with good rate function

$$I(f) = \begin{cases} \frac{1}{2}\int_0^1 \frac{1}{\sigma(f(t))}\big(f'(t) - b(f(t))\big)^2\,dt, & f \in AC[0,1], f(0) = x, \\ \infty, & \textit{otherwise.} \end{cases}$$

In the next section we will show that a large deviation principle holds whenever the coefficients b, σ, η are globally Lipschitz continuous.

7.2 Large Deviations by Exponential Approximations

Dembo and Zeitouni [10] obtained a large deviation principle for solutions of SDEs driven by Brownian motion, i. e. SDEs of the form

$$dX_t = b(X_t)\,dt + \sqrt{\varepsilon}\sigma(X_t)\,dB_t, \qquad X_0 = x,$$

by applying a generalization of the contraction principle, cf. Theorem 4.4. The key point is the approximation of the drift and diffusion coefficient by simple functions:

$$dX_t^m = b(X^m_{\frac{\lfloor tm \rfloor}{m}})\,dt + \sqrt{\varepsilon}\sigma(X^m_{\frac{\lfloor tm \rfloor}{m}})\,dB_t, \qquad m \in \mathbb{N}.$$

Doing so, the stochastic integral can be evaluated pathwise, and therefore, similar to Example 2.9, we can define continuous mappings F^m such that $F^m(\sqrt{\varepsilon}B) = X^m$. Using the large deviation principle for the scaled Brownian motion $(\sqrt{\varepsilon}B)_{\varepsilon>0}$, Theorem 4.4 yields the desired large deviation principle. In this section we will show that the approach remains valid if we consider solutions of SDEs driven by Lévy processes, i. e.

$$dX_t = b(X_{t-})\,dt + \sqrt{\varepsilon}\sigma(X_{t-})\,dB_t + \eta(X_{t-})\,dL_t^\varepsilon \qquad X_0 = x, \tag{7.17}$$

where $L_t^\varepsilon := \varepsilon L_{t/\varepsilon}$ is a scaled Lévy process with finite exponential moments. Let us remark that the result we are going to prove is a special case of large deviation results presented in [9] and [19] for SDEs driven by Brownian motion and random measures. Throughout this section we denote by

$$\mathcal{F}_t^\varepsilon := \sigma(B_s, L_s^\varepsilon, \mathcal{N}; s \leq t)$$

the canonical filtration augmented by the $\mathbb{P}$-nullsets $\mathcal{N}$. We remind the reader that this filtration satisfies the *usual hypotheses*, i. e. it is a right-continuous complete filtration, cf. [32, Theorem I.31], and is admissible for both $(B_t)_{t\geq 0}$ and $(L_t^\varepsilon)_{t\geq 0}$. We call a process $(X_t)_{t\geq 0}$ $\mathcal{F}^\varepsilon$-previsible if it is measurable with respect to the σ-algebra generated by the left-continuous $\mathcal{F}^\varepsilon$-adapted processes.

7.6 Theorem *Let $(L_t)_{t\geq 0}$ be a Lévy process with Lévy triplet $(\gamma, 0, \nu)$ and symbol ψ such that $\mathbb{E}e^{\lambda |L_1|} < \infty$ for all $\lambda \geq 0$. Let $b, \sigma, \eta : \mathbb{R} \to \mathbb{R}$ be bounded globally Lipschitz continuous functions. In particular, there exists $L > 0$ such that*

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| + |\eta(x) - \eta(y)| \leq L|x - y| \qquad \text{for all} \;\; x, y \in \mathbb{R}.$$

Then the family $(X^\varepsilon)_{\varepsilon>0}$ of solutions of (7.17) satisfies a large deviation principle in $(D[0,1], \|\cdot\|_\infty)$ as $\varepsilon \to 0$ with good rate function

$$I(f) := I_x(f) := \begin{cases} \inf\left(\frac{1}{2}\int_0^1 |g'(t)|^2 \, dt + \int_0^1 \Psi^*(h'(t)) \, dt\right), & f \in AC[0,1], f(0) = x, \\ \infty, & \text{otherwise}, \end{cases} \tag{7.18}$$

where the infimum is taken over all functions $g, h \in AC[0,1]$, $g(0) = h(0) = 0$, such that

$$f(t) = x + \int_0^t b(f(s)) \, ds + \int_0^t \sigma(f(s)) g'(s) \, ds + \int_0^t \eta(f(s)) h'(s) \, ds$$

and

$$\Psi(w) := \psi(-\imath w) = \gamma w + \int_{\mathbb{R}\setminus\{0\}} \left(e^{yw} - 1 - yw \mathbb{1}_{|y|\leq 1}\right) \nu(dy), \qquad w \in \mathbb{R}$$

denotes the logarithmic moment generating function of L_1.

Remarks (i). Comparing Theorem 7.6 and Theorem 7.1, it seems reasonable to claim that the rate functions (7.18) and (7.3) coincide. In case both theorems are applicable, this follows from the uniqueness of the rate function. Unfortunately, we did not succeed in proving the equality directly except for the special cases $\eta = 0$ (continuous SDE) and $\sigma = 0$ (jump-only SDE).

(ii). The assumption on the boundedness of the coefficients can be weakened. Feng-Kurtz [17, Theorem 10.17, Remark 10.18] have shown that a large deviation principle holds if b, σ are of *linear growth*, i. e.

$$|b(x)| + |\sigma(x)| \leq M(1 + |x|), \qquad x \in \mathbb{R}$$

for some constant $M > 0$, and η satisfies a certain integrability condition.

We split the proof of Theorem 7.6 into several parts:

(i). Show that $(\sqrt{\varepsilon}B, L^\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle in $D[0,1] \times D[0,1]$ as $\varepsilon \to 0$, cf. Lemma 7.7.

(ii). Define continuous mappings $F^m : D[0,1] \times D[0,1] \to D[0,1]$, $m \in \mathbb{N}$, such that $(F^m(\sqrt{\varepsilon}B, L^\varepsilon))_{m\in\mathbb{N},\varepsilon>0}$ is an exponentially good approximation of $(X^\varepsilon)_{\varepsilon>0}$, cf. Lemma 7.9.

(iii). Apply Theorem 4.4 in order to obtain the large deviation principle for $(X^\varepsilon)_{\varepsilon>0}$.

Throughout the remaining part of this section the processes $(L_t)_{t\geq 0}$, $(B_t)_{t\geq 0}$, and the coefficients b, σ, η are supposed to meet the assumptions of Theorem 7.6. Without loss of generality we may assume $\mathbb{E}L_t = t\mathbb{E}L_1 = 0$ (otherwise consider the Lévy process $L_t - t\mathbb{E}L_1$). Then, by the Lévy-Itô decomposition,

$$L_t^\varepsilon = \int_0^t \int z \, d\tilde{N}_\varepsilon(dz, ds) \tag{7.19}$$

where $\tilde{N}_\varepsilon$ denotes the compensated jump counting measure of the Lévy process L^ε. It follows from the Lévy-Khinchine formula that the compensator $\hat{N}_\varepsilon(dz, ds)$ of the jump counting measure N_ε equals $\varepsilon^{-1} ds \, \nu(\varepsilon^{-1} dz)$.

7.7 Lemma *$(\sqrt{\varepsilon}B, L^\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle in $D[0,1]\times D[0,1]$ endowed with the norm*

$$\|(f,g)\| := \|f\|_\infty + \|g\|_\infty, \qquad f, g \in D[0,1]$$

as $\varepsilon \to 0$ with good rate function I_0,

$$I_0(g,h) := \begin{cases} \frac{1}{2}\int_0^1 |g'(t)|^2 \, dt + \int_0^1 \Psi^*(h'(t)) \, dt, & g, h \in AC[0,1], g(0) = h(0) = 0, \\ \infty, & \textit{otherwise.} \end{cases} \tag{7.20}$$

Proof. In accordance with Chapter 5 we denote by $Z^B_{\lfloor 1/\varepsilon \rfloor} / \lfloor 1/\varepsilon \rfloor$ and $Z^L_{\lfloor 1/\varepsilon \rfloor} / \lfloor 1/\varepsilon \rfloor$ the approximations of $(B_t^\varepsilon)_{t\in[0,1]}$ and $(L_t^\varepsilon)_{t\in[0,1]}$, respectively. Since $(B_t)_{t\geq 0}$ and $(L_t)_{t\geq 0}$ are independent, see e. g. [24, Theorem II.6.3], we find by combining Lemma 5.3, Theorem 5.5 and Lemma 2.7 that $(Z^B_{\lfloor 1/\varepsilon \rfloor} / \lfloor 1/\varepsilon \rfloor, Z^L_{\lfloor 1/\varepsilon \rfloor} / \lfloor 1/\varepsilon \rfloor)_{\varepsilon>0}$ satisfies a large deviation principle in $(D[0,1] \times D[0,1], \|\cdot\|)$ as $\varepsilon \to 0$ with good rate function I_0. Using the scaling property, cf. Example 5.10, and Lemma 5.6, it is not difficult to see that this family of processes is exponentially equivalent to $(\sqrt{\varepsilon}B, L^\varepsilon)_{\varepsilon>0}$. Therefore, the claim follows from Corollary 4.3 and Theorem 5.7. ☐

Remark By Theorem 5.1, $(\sqrt{\varepsilon}B)_{\varepsilon>0}$ as well as $(L^\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle in $(D[0,1], \|\cdot\|_\infty)$ with a good rate function but, since $(D[0,1], \|\cdot\|_\infty)$ is not a Polish space, this does not imply exponential tightness, cf. Lemma 2.5. Therefore, we cannot apply Lemma 2.7 to $(\sqrt{\varepsilon}B, L^\varepsilon)_{\varepsilon>0}$ directly.

The remaining part of the proof is based on the idea that the solutions $(X_t^{\varepsilon,m})_{t\in[0,1]}$ of the stochastic differential equation

$$dX_t^{\varepsilon,m} = b(X^{\varepsilon,m}_{\frac{\lfloor mt \rfloor}{m}-}) \, dt + \sqrt{\varepsilon}\sigma(X^{\varepsilon,m}_{\frac{\lfloor mt \rfloor}{m}-}) \, dB_t + \eta(X^{\varepsilon,m}_{\frac{\lfloor mt \rfloor}{m}-}) \, dL_t^\varepsilon, \qquad X_0^{\varepsilon,m} = x. \tag{7.21}$$

are an exponentially good approximation of $(X^\varepsilon)_{\varepsilon>0}$. In order to prove this we need the following technical lemma.

7.8 Lemma *Let $b, \sigma, \eta : [0,\infty) \times \Omega \to \mathbb{R}$ be $\mathcal{F}^\varepsilon$-previsible bounded processes, and set*

$$Z_t := z_0 + \int_0^t b_s \, ds + \sqrt{\varepsilon} \int_0^t \sigma_s \, dB_s + \int_0^t \eta_s \, dL_s^\varepsilon.$$

Let τ be a $\mathcal{F}_t^\varepsilon$-stopping time and $\varrho > 0$, $M > 0$ such that

$$|b_t| + |\sigma_t| + |\eta_t| \leq M(\varrho^2 + |Z_{t-}|^2)^{\frac{1}{2}} \qquad \textit{for all} \ \ t \in [0,\tau]. \tag{7.22}$$

Then, for fixed $\delta > 0$,

$$\varepsilon \log \mathbb{P}\left(\sup_{t\in[0,\tau\wedge 1]} |Z_t| > \delta \right) \leq C + \log\left(\frac{\varrho^2 + z_0^2}{\varrho^2 + \delta^2} \right) \qquad \textit{for all} \ \ \varepsilon \leq 1$$

for a constant $C > 0$ which does not depend on $\delta, \varrho, \varepsilon$.

Proof. Set $\varphi(y) := (\varrho^2 + y^2)^{\frac{1}{\varepsilon}}$. By Itô's formula and (7.19),

$$\begin{aligned}
\varphi(Z_t) - \varphi(Z_0) &= \int_0^t \varphi'(Z_{s-}) b_s \, ds + \sqrt{\varepsilon} \int_0^t \varphi'(Z_{s-}) \sigma_s \, dB_s \\
&\quad + \frac{\varepsilon}{2} \int_0^t \varphi''(Z_{s-}) \sigma_s^2 \, ds + \int_0^t \!\! \int \big(\varphi(Z_{s-} + \eta_s z) - \varphi(Z_{s-}) \big) \, d\tilde{N}_\varepsilon(dz, ds) \\
&\quad + \int_0^t \!\! \int \big(\varphi(Z_{s-} + \eta_s z) - \varphi(Z_{s-}) - \varphi'(Z_{s-}) \eta_s \, z \big) \, d\hat{N}_\varepsilon(dz, ds).
\end{aligned} \tag{7.23}$$

Obviously,

$$\varphi'(y) = \frac{2}{\varepsilon} \frac{\varphi(y)}{\varrho^2 + y^2} y \qquad\qquad \varphi''(y) = \frac{2}{\varepsilon} \frac{\varphi(y)}{\varrho^2 + y^2} \left(1 + 2\left(\frac{1}{\varepsilon} - 1 \right) \frac{y^2}{\varrho^2 + y^2} \right). \tag{7.24}$$

Define a stopping time $\tau_1 := \inf\{t \geq 0; |Z_t| > \delta\} \wedge \tau \wedge 1$. Using the boundedness of σ, η and $\int_{\mathbb{R}\setminus\{0\}} |y|^n \, \nu(dy) < \infty$, $n \geq 2$, we find

$$\begin{aligned}
\mathbb{E}\left(\int_0^1 \varphi'(Z_{s-})^2 \sigma_s^2 \mathbb{1}_{[0,\tau_1]}(s) \, ds \right) &< \infty, \\
\mathbb{E}\left(\int_0^1 \!\! \int_{\mathbb{R}\setminus\{0\}} \big(\varphi(Z_{s-} + \eta_s z) - \varphi(Z_{s-}) \big)^2 \mathbb{1}_{[0,\tau_1]}(s) d\hat{N}_\varepsilon(dz, ds) \right) &< \infty.
\end{aligned}$$

This means that the corresponding stochastic integrals in (7.23) are martingales, and therefore we obtain

$$\mathbb{E}\left(\int_0^{t\wedge\tau_1} \varphi'(Z_{s-}) \sigma_s \, dB_s \right) + \mathbb{E}\left(\int_0^{t\wedge\tau_1} \!\! \int \big(\varphi(Z_{s-} + \eta_s z) - \varphi(Z_{s-}) \big) \, d\tilde{N}_\varepsilon(dz, ds) \right) = 0.$$

Moreover, because of the growth condition (7.22), it is not difficult to see that

$$\mathbb{E}\left(\int_0^{t\wedge\tau_1} \varphi'(Z_{s-}) b_s \, ds \right) + \mathbb{E}\left(\frac{\varepsilon}{2} \int_0^{t\wedge\tau_1} \varphi''(Z_{s-}) \sigma_s^2 \, ds \right) \leq \frac{C_1}{\varepsilon} \mathbb{E}\left(\int_0^{t\wedge\tau_1} \varphi(Z_{s-}) \, ds \right)$$

for some constant C_1 which does not depend on $\varrho, \delta, \varepsilon$. In order to apply Gronwall's lemma we have to estimate the remaining term in (7.23). By Taylor's formula and the definition of L_t^ε, cf. (7.19),

$$\begin{aligned}
I &:= \int_0^{t\wedge\tau_1} \!\! \int \big(\varphi(Z_{s-} + \eta_s z) - \varphi(Z_{s-}) - \varphi'(Z_{s-}) \eta_s \, z \big) \, d\hat{N}_\varepsilon(dz, ds) \\
&= \frac{1}{\varepsilon} \int_0^{t\wedge\tau_1} \!\! \int_{\mathbb{R}\setminus\{0\}} \big(\varphi(Z_{s-} + \varepsilon\eta_s y) - \varphi(Z_{s-}) - \varphi'(Z_{s-}) \varepsilon\eta_s y \big) \, \nu(dy) \, ds \\
&= \frac{1}{\varepsilon} \int_0^{t\wedge\tau_1} \!\! \int_{\mathbb{R}\setminus\{0\}} \varphi''(Z_{s-} + \varepsilon\Theta\eta_s y)(\varepsilon\eta_s y)^2 \, \nu(dy) \, ds
\end{aligned}$$

for some intermediate value $\Theta = \Theta(\omega, s, y) \in (0,1)$.[6] By (7.22) and (7.24),

$$|I| \le 2M^2 \left(1 + \frac{1}{\varepsilon}\right) \int_0^{t\wedge\tau_1} \int_{\mathbb{R}\setminus\{0\}} y^2 (\varrho^2 + (Z_{s-} + \varepsilon\Theta\eta_s y)^2)^{\frac{1}{\varepsilon}-1} (\varrho^2 + Z_{s-}^2)\, \nu(dy)\, ds.$$

Note that

$$(1 + \varepsilon x)^{\frac{1}{\varepsilon}} \uparrow e^x \text{ as } \varepsilon \downarrow 0 \qquad \text{for all} \quad x \ge 0.$$

Therefore, we find

$$\begin{aligned}
(\varrho^2 + (Z_{s-} + \varepsilon\Theta\eta_s y)^2)^{\frac{1}{\varepsilon}-1} &\le \left(1 + 2\varepsilon \frac{|Z_{s-}\eta_s y|}{\varrho^2 + Z_{s-}^2} + \varepsilon^2 \frac{\eta_s^2 y^2}{\varrho^2 + Z_{s-}^2}\right)^{\frac{1}{\varepsilon}} (\varrho^2 + Z_{s-}^2)^{\frac{1}{\varepsilon}-1} \\
&\overset{(7.22)}{\le} (1 + \varepsilon M' |y|)^{\frac{2}{\varepsilon}} (\varrho^2 + Z_{s-}^2)^{\frac{1}{\varepsilon}-1} \\
&\le e^{2M'|y|} (\varrho^2 + Z_{s-}^2)^{\frac{1}{\varepsilon}-1}
\end{aligned}$$

for all $s \in [0, \tau_1]$ for a constant $M' > 0$. Thus,

$$|I| \le 2M^2 \left(1 + \frac{1}{\varepsilon}\right) \left(\int_{\mathbb{R}\setminus\{0\}} y^2 e^{2M'|y|}\, \nu(dy)\right) \left(\int_0^{t\wedge\tau_1} \varphi(Z_{s-})\, ds\right) =: C_2(\varepsilon) \int_0^{t\wedge\tau_1} \varphi(Z_{s-})\, ds.$$

Clearly,

$$\lim_{\varepsilon\to 0} \varepsilon C_2(\varepsilon) = 2M^2 \int_{\mathbb{R}\setminus\{0\}} y^2 e^{2M'|y|}\, \nu(dy). \tag{7.25}$$

Recall that L_1 has exponenial moments so that the integral on the right hand side is finite. Consequently, we have shown

$$\mathbb{E}\varphi(Z_{t\wedge\tau_1}) - \varphi(z_0) \le \left(\frac{C_1}{\varepsilon} + C_2(\varepsilon)\right) \int_0^t \mathbb{E}\varphi(Z_{s\wedge\tau_1})\, ds.$$

Applying Gronwall's lemma yields

$$\mathbb{E}\varphi(Z_{\tau_1}) \le \varphi(z_0) \exp\left(\frac{C_1}{\varepsilon} + C_2(\varepsilon)\right).$$

Finally, by the definition of τ_1 and Markov's inequality,

$$\begin{aligned}
\varepsilon \log \mathbb{P}\left(\sup_{t\in[0,\tau\wedge 1]} |Z_t| > \delta\right) &\le \varepsilon \log \mathbb{P}(|Z_{\tau_1}| \ge \delta) \le \varepsilon \log \mathbb{E}\varphi(Z_{\tau_1}) - \varepsilon \log \varphi(\delta) \\
&\le C_1 + \varepsilon\, C_2(\varepsilon) + \log\left(\frac{\varrho^2 + |z_0|^2}{\varrho^2 + \delta^2}\right). \qquad \square
\end{aligned}$$

Now we are ready to show that the family of solutions $(X^{\varepsilon,m})_{\varepsilon>0, m\in\mathbb{N}}$ given by (7.21) is indeed an exponentially good approximation.

7.9 Lemma *For $m \in \mathbb{N}$ define $F_m : D[0,1] \times D[0,1] \to D[0,1]$ via $f = F^m(g,h)$ where*

$$f(t) = f(t_k^m) + b(f(t_k^m-))(t - t_k^m) + \sigma(f(t_k^m-))(g(t) - g(t_k^m)) + \eta(f(t_k^m-))(h(t) - h(t_k^m))$$

for $t \in (t_k^m, t_{k+1}^m]$, $t_k^m := k/m$, $k = 0, \dots, m-1$, and $f(0) := f(0-) := x$. Then $X^{\varepsilon,m} := F^m(\sqrt{\varepsilon}B, L^\varepsilon)$ defines an exponentially good approximation of $(X^\varepsilon)_{\varepsilon>0}$.

[6] The measurability of the mapping $(s, \omega, y) \mapsto \varphi''(Z_{s-} + \varepsilon\Theta\eta_s y)$ follows from the integral representation of the remainder term in the Taylor formula.

Proof. Let $\delta,\varrho,\varepsilon>0$. Obviously, $(X_t^{\varepsilon,m})_{t\in[0,1]}$ is the (unique) solution to the SDE

$$dY_t = b(Y_{\frac{\lfloor mt\rfloor}{m}-})\,dt + \sqrt{\varepsilon}\sigma(Y_{\frac{\lfloor mt\rfloor}{m}-})\,dB_t + \eta(Y_{\frac{\lfloor mt\rfloor}{m}-})\,dL_t^{\varepsilon}, \qquad 0\le t\le 1, \qquad Y_0 = x.$$

Define a $\mathcal{F}_t^{\varepsilon}$-stopping time by

$$\tau := \tau(\varrho) := \inf\{t\ge 0; |X_t^{\varepsilon,m} - X_{\lfloor mt\rfloor/m}^{\varepsilon,m}| > \varrho\}\wedge 1,$$

and set

$$b_t := b(X_{\lfloor mt\rfloor/m-}^{\varepsilon,m}) - b(X_{t-}^{\varepsilon}) \quad \sigma_t := \sigma(X_{\lfloor mt\rfloor/m-}^{\varepsilon,m}) - \sigma(X_{t-}^{\varepsilon}) \quad \eta_t := \eta(X_{\lfloor mt\rfloor/m-}^{\varepsilon,m}) - \eta(X_{t-}^{\varepsilon}).$$

By the global Lipschitz continuity,

$$|b_t| + |\sigma_t| + |\eta_t| \le L|X_{\lfloor mt\rfloor/m-}^{\varepsilon,m} - X_{t-}^{\varepsilon}| \le \sqrt{2}L(\varrho^2 + |X_{t-}^{\varepsilon,m} - X_{t-}^{\varepsilon}|^2)^{\frac{1}{2}}$$

for any $t\in[0,\tau]$. Lemma 7.8 (applied to $Z_t := X_t - X_t^{\varepsilon,m}$) shows

$$\varepsilon\log\mathbb{P}\left(\sup_{t\in[0,\tau]}|X_t^{\varepsilon,m} - X_t^{\varepsilon}| > \delta\right) \le C + \log\left(\frac{\varrho^2}{\varrho^2+\delta^2}\right).$$

for some constant $C>0$ which does not depend on m,ε,ϱ. Hence,

$$\lim_{\varrho\to 0}\sup_{m\ge 1}\limsup_{\varepsilon\to 0}\varepsilon\log\mathbb{P}\left(\sup_{t\in[0,\tau]}|X_t^{\varepsilon,m} - X_t^{\varepsilon}| > \delta\right) = -\infty \qquad \text{for all } \ \delta>0.$$

As

$$\{\|X^{\varepsilon,m} - X^{\varepsilon}\|_\infty > \delta\} \subseteq \{\tau<1\}\cup\left\{\sup_{t\in[0,\tau]}|X_t^{\varepsilon,m} - X_t^{\varepsilon}| > \delta\right\}$$

it remains to show

$$\lim_{m\to\infty}\limsup_{\varepsilon\to 0}\varepsilon\log\mathbb{P}(\tau<1) = -\infty \qquad \text{for all } \ \varrho>0. \tag{7.26}$$

Since the coefficients are bounded, we find

$$|X_{\frac{k}{m}+s}^{\varepsilon,m} - X_{\frac{k}{m}}^{\varepsilon,m}| \le C\left(\frac{1}{m} + \sqrt{\varepsilon}\max_{0\le k\le m-1}\sup_{0\le s\le\frac{1}{m}}|B_{\frac{k}{m}+s} - B_{\frac{k}{m}}| + \max_{0\le k\le m-1}\sup_{0\le s\le\frac{1}{m}}|L_{\frac{k}{m}+s}^{\varepsilon} - L_{\frac{k}{m}}^{\varepsilon}|\right)$$

for $0\le s\le 1/m$ where $C := \max\{\|b\|_\infty, \|\eta\|_\infty, \|\sigma\|_\infty\}$. By the stationarity of the increments, we conclude

$$\begin{aligned}\mathbb{P}(\tau<1) &= \mathbb{P}\left(\bigcup_{k=0}^{m-1}\left\{\sup_{0\le s\le 1/m}|X_{\frac{k}{m}+s}^{\varepsilon,m} - X_{\frac{k}{m}}^{\varepsilon,m}| > \varrho\right\}\right)\\ &\le m\mathbb{P}\left(\sup_{0\le s\le\frac{1}{m}}|B_s| \ge \frac{\varrho - C/m}{2\sqrt{\varepsilon}C}\right) + m\mathbb{P}\left(\sup_{0\le s\le\frac{1}{m\varepsilon}}|L_s| \ge \frac{\varrho - C/m}{2\varepsilon C}\right)\end{aligned} \tag{7.27}$$

for all $m > C/\varrho$. Applying Etemadi's inequality, Markov's inequality and Lemma 5.9 gives

$$\mathbb{P}\left(\sup_{0\le s\le\frac{1}{m\varepsilon}}|L_s| \ge \frac{\varrho - C/m}{2\varepsilon C}\right) \le 6\exp\left(-\sqrt{m}\frac{\varrho - C/m}{2\varepsilon C} + \frac{C'}{\varepsilon}\right).$$

for a constant $C'>0$. An analogous estimate holds for the first term in (7.27). Combining these estimates proves (7.26). □

Summarizing, we have shown that $(\sqrt{\varepsilon}B, L^\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle with good rate function I_0, and that $(F^m(\sqrt{\varepsilon}B, L^\varepsilon))_{\varepsilon>0, m\in\mathbb{N}}$ defines an exponentially good approximation of $(X^\varepsilon)_{\varepsilon>0}$. The task is now to find a function $F: D[0,1]\times D[0,1] \to D[0,1]$ such that F^m converges uniformly on the sublevel sets of I_0 to F. Then, the claim follows from Theorem 4.4.

7.10 Lemma *(i). The mappings F^m, $m \in \mathbb{N}$, defined in Lemma 7.9 are continuous.*

(ii). For absolutely continuous functions $g, h \in D[0,1]$ and $x \in \mathbb{R}$ denote by $f := F(g,h)$ the unique solution of the integral equation

$$f(t) = x + \int_0^t b(f(s))\,ds + \int_0^t \sigma(f(s))g'(s)\,ds + \int_0^t \eta(f(s))h'(s)\,ds, \qquad t \in [0,1].$$

Then,

$$\lim_{m\to\infty} \sup_{(g,h)\in\Phi(r)} \|F^m(g,h) - F(g,h)\|_\infty = 0 \qquad \textit{for all } \; r \geq 0,$$

where $\Phi(r) := \{(g,h) \in D[0,1]\times D[0,1]; I_0(g,h) \leq r\}$ is the sublevel set of the good rate function I_0 defined in (7.20).

Proof. (i). For $(g_1,h_1),(g_2,h_2) \in D[0,1]\times D[0,1]$ we set $f_j(t) := F^m(g_j,h_j)(t)$, $j = 1,2$, and

$e(t) := |f_1(t) - f_2(t)|$. By the Lipschitz continuity and boundedness of σ, we have

$$|\sigma(f_1(t_k^m-))g_1(t) - \sigma(f_2(t_k^m-))g_2(t)| \leq \|\sigma\|_\infty \cdot \|g_1 - g_2\|_\infty + e(t_k^m-)\cdot\|g_2\|_\infty.$$

Using similar estimates for the other terms, we find

$$\sup_{t\in[t_k^m, t_{k+1}^m]} e(t) \leq C\left(e(t_k^m) + e(t_k^m-) + \|g_1-g_2\|_\infty + \|h_1-h_2\|_\infty\right), \qquad k = 0,\ldots,m-1,$$

for some constant $C > 0$ which does only depend on (g_2,h_2). Since $e(0) = e(0-) = 0$ the continuity of F^m at (g_1,h_1) follows by iterating the above estimate over k.

(ii). By assumption, b, σ, η are (globally) Lipschitz continuous, and therefore there exists a unique solution of the given integral equation. In order to prove the uniform convergence on the sublevel sets of I_0 we apply Gronwall's lemma. We claim that there exists a sequence $(\varepsilon_m)_{m\in\mathbb{N}}$ such that $\varepsilon_m \to 0$ as $m \to \infty$ and

$$\int_{\frac{\lfloor tm\rfloor}{m}}^t |h'(s)|\,ds \leq \varepsilon_m \qquad \text{for all } \; t\in[0,1], (g,h)\in\Phi(r). \tag{7.28}$$

Indeed: By the definition of the Legendre transform,

$$h'(s) \leq \frac{\Psi^*(h'(s)) + \Psi(\alpha)}{\alpha} \qquad \text{for all } \; \alpha > 0.$$

Hence,[7]

$$\int_{\frac{\lfloor tm\rfloor}{m}}^t h'(s)\mathbb{1}_{\{h'(s)\geq 0\}}\,ds \leq \frac{1}{\alpha_m}\underbrace{\int_0^1 \Psi^*(h'(s))\,ds}_{\leq r} + \frac{1}{m}\frac{\Psi(\alpha_m)}{\alpha_m} \tag{7.29}$$

[7]Revised version: Corrected misprint.

for any sequence $(\alpha_m)_{m\in\mathbb{N}} \subseteq (0,\infty)$. If $\Psi|_{(0,\infty)}$ is bounded, we choose $\alpha_m := m$; otherwise we set

$$\alpha_m := \sup\{x \geq 0; |\Psi(x)| \leq m\}, \qquad m \in \mathbb{N};$$

then clearly $\alpha_m \to \infty$ as $m \to \infty$. In both cases, (7.29) entails[8]

$$\int_0^1 h'(s)\mathbb{1}_{\{h'(s)\geq 0\}}\, ds \leq \varepsilon_m \xrightarrow{m\to\infty} 0$$

for a sequence $(\varepsilon_m)_{m\in\mathbb{N}}$ which does not depend on $(g,h)\in\Phi(r)$ and $t\in[0,1]$. A similar estimate holds for the integral $\int_{\frac{\lfloor t\, m\rfloor}{m}}^t(-h'(s))\mathbb{1}_{\{h'(s)<0\}}\, ds$[9]. This proves (7.28). An even simpler computation shows

$$\int_0^1 |h'(s)|\, ds \leq \int_0^1 \Psi^*(h'(s))\, ds + (\Psi(1)+\Psi(-1)). \tag{7.30}$$

Fix $(g,h)\in\Phi(r)$. Using the boundedness of the coefficients, we find by the Cauchy-Schwarz inequality

$$\begin{aligned}\sup_{0\leq t\leq 1}&\left|F^m(g,h)(t) - F^m(g,h)\left(\frac{\lfloor t\, m\rfloor}{m}\right)\right| \\ &\leq \frac{\|b\|_\infty}{m} + \|\sigma\|_\infty\sqrt{\frac{1}{m}}\underbrace{\sqrt{\int_0^1 g'(s)^2\, ds}}_{\leq\sqrt{2r}} + \|\eta\|_\infty\varepsilon_m =: \delta_m \xrightarrow{m\to\infty} 0.\end{aligned}$$

A similar calculation shows

$$\begin{aligned}e(t) &:= |F(g,h)(t) - F^m(g,h)(t)| \\ &\leq L\int_0^t (1+|g'(s)|+|h'(s)|)\left|F(g,h)(s) - F^m(g,h)\left(\frac{\lfloor m\, s\rfloor}{m}\right)\right|\, ds \\ &\leq LC\delta_m + L\int_0^t (1+|g'(s)|+|h'(s)|)e(s)\, ds\end{aligned}$$

where we used that

$$\begin{aligned}\int_0^t (1+|g'(s)|+|h'(s)|)\, ds &\overset{(7.30)}{\leq} 1 + \sqrt{\int_0^1 g'(s)^2\, ds} + \int_0^1 \Psi^*(h'(s))\, ds + \Psi(1)+\Psi(-1) \\ &\leq \ 1+\sqrt{2r}+r+\Psi(1)+\Psi(-1) =: C \end{aligned}\tag{7.31}$$

for any $t\in[0,1]$ as $(g,h)\in\Phi(r)$. From Gronwall's lemma we see

$$\begin{aligned}e(t) &\leq LC\delta_m\left[1 + L\int_0^t (1+|g'(s)|+|h'(s)|)\exp\left(L\int_s^t (1+|g'(r)|+|h'(r)|)\, dr\right)\, ds\right] \\ &\overset{(7.31)}{\leq} LC\delta_m\left(1+LCe^{LC}\right)\end{aligned}$$

Since the constants L, C, δ_m do not depend on t and (g,h), we conclude

$$\sup_{(g,h)\in\Phi(r)} \|F(g,h) - F^m(g,h)\|_\infty \leq LC\delta_m\left(1+LCe^{LC}\right) \xrightarrow{m\to\infty} 0. \qquad \square$$

[8] Revised version: Corrected misprint.

[9] Revised version: Corrected misprint.

Proof of Theorem 7.6. In Lemma 7.7 we have shown that $(\sqrt{\varepsilon}B, L^\varepsilon)_{\varepsilon>0}$ obeys a large deviation principle with good rate function I_0. Since $(F^m(\sqrt{\varepsilon}B, L^\varepsilon))_{\varepsilon>0, m\in\mathbb{N}}$ is an exponentially good approximation of $(X^\varepsilon)_{\varepsilon>0}$, cf. Lemma 7.9, and F^m converges uniformly on the compact sublevel sets of I_0 to F the assumptions of Theorem 4.4 are satisfied. Consequently, $(X^\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle with good rate function

$$I(f) = \begin{cases} \inf\left(\frac{1}{2}\int_0^1 |g'(t)|^2\,dt + \int_0^1 \Psi^*(h'(t))\,dt\right), & f \in AC[0,1], f(0) = x, \\ \infty, & \text{otherwise}, \end{cases}$$

where the infimum is taken over all functions $g, h \in AC[0,1]$, $g(0) = h(0) = 0$, such that

$$f(t) = F(g,h)(t) = x + \int_0^t b(f(s))\,ds + \int_0^t \sigma(f(s))g'(s)\,ds + \int_0^t \eta(f(s))h'(s)\,ds. \qquad \square$$

Remark It is not difficult to see that the proof of Theorem 7.6 carries over to solutions of non-homogeneous SDEs of the form

$$dX_t^\varepsilon = b(t, X_{t-})\,dt + \sqrt{\varepsilon}\sigma(t, X_{t-})\,dB_t + \eta(t, X_{t-})\,dL_t^\varepsilon$$

if the coefficients b, σ, η are bounded and (globally) Lipschitz continuous in both components, i. e. there exists a constant $L > 0$ such that

$$|b(s,x) - b(t,y)| + |\sigma(s,x) - \sigma(t,y)| + |\eta(s,x) - \eta(t,y)| \leq L(|t-s| + |y-x|)$$

for any $s, t \in [0,1]$, $x, y \in \mathbb{R}$.

The following result shows that the large deviation bounds in Theorem 7.6 can be strengthened; they hold uniformly (with respect to the initial condition $x \in \mathbb{R}$) on compact sets.

7.11 Corollary *Let $K \subseteq \mathbb{R}$ be compact. Denote by $(X_t^{\varepsilon,x})_{t\in[0,1]}$ the solution of* (7.17) *for the initial condition $X_0^{\varepsilon,x} = x \in \mathbb{R}$. Then,*

$$\limsup_{\varepsilon\to 0} \varepsilon \log \sup_{x\in K} \mathbb{P}(X^{\varepsilon,x} \in B) \leq -\inf_{x\in K}\inf_{f\in B} I_x(f) \tag{7.32}$$

for any closed set $B \subseteq D[0,1]$, and

$$\liminf_{\varepsilon\to 0} \varepsilon \log \inf_{x\in K} \mathbb{P}(X^{\varepsilon,x} \in A) \geq -\sup_{x\in K}\inf_{f\in A} I_x(f) \tag{7.33}$$

for any open set $A \subseteq D[0,1]$.

Proof. Let $(x_\varepsilon)_{\varepsilon>0}$ such that $x_\varepsilon \to x \in \mathbb{R}$ as $\varepsilon \to 0$. As b, σ, η are (globally) Lipschitz continuous, Lemma 7.8 applied to $Y_t^\varepsilon := X_t^{\varepsilon,x_\varepsilon} - X_t^{\varepsilon,x}$ and the stopping time $\tau = 1$ yields

$$\varepsilon \log \mathbb{P}(\|X^{\varepsilon,x_\varepsilon} - X^{\varepsilon,x}\|_\infty > \delta) \leq C + \log\left(\frac{\varrho^2 + |x_\varepsilon - x|^2}{\varrho^2 + \delta^2}\right).$$

Letting $\varrho \to 0$ and $\varepsilon \to 0$, we find that $(X^{\varepsilon,x_\varepsilon})_{\varepsilon>0}$ and $(X^{\varepsilon,x})_{\varepsilon>0}$ are exponentially equivalent, and therefore $(X^{\varepsilon,x_\varepsilon})_{\varepsilon>0}$ obeys a large deviation principle in $D[0,1]$ with good rate function I_x, cf. Corollary 4.3. In particular,

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}(X^{\varepsilon,x_\varepsilon} \in B) \leq -\inf_{f\in B} I_x(f) \tag{7.34}$$

for any closed set $B \subseteq D[0,1]$. Fix $\delta > 0$. By (7.34), there exists for each $x \in K$ some $\varepsilon_x > 0$ such that

$$\varepsilon \log \sup_{y \in B(x,\varepsilon_x)} \mathbb{P}(X^{\varepsilon,y} \in B) \leq -\min\left\{\inf_{f \in B} I_x(f) - \delta, \frac{1}{\delta}\right\} \qquad \text{for all } \varepsilon \leq \varepsilon_x;$$

otherwise we could construct a sequence $(x_\varepsilon)_{\varepsilon>0}$, $x_\varepsilon \to x$, such that (7.34) is violated. Choosing a finite cover $\bigcup_{j=1}^n B(x_j, \varepsilon_{x_j})$ of K, we conclude

$$\varepsilon \log \sup_{y \in K} \mathbb{P}(X^{\varepsilon,y} \in B) \leq -\min\left\{\inf_{f \in B} I_x(f) - \delta, \frac{1}{\delta}\right\} \qquad \text{for all } \varepsilon \leq \min_{j=1,\ldots,n} \varepsilon_{x_j}.$$

As $\delta > 0$ is arbitrary, this proves (7.32); (7.33) follows in a similar way. □

So far we have used the extended contraction principle presented in Section 4.1 which relies on the (uniform) approximation of the function $(g,h) \mapsto f = F(g,h)$,

$$f(t) = x + \int_0^t b(f(s))\,ds + \int_0^t \sigma(f(s))g'(s)\,ds + \int_0^t \eta(f(s))h'(s)\,ds,$$

by continuous functions. In particular, F is continuous on the sublevel sets of the good rate function I_0. In a more general setting – for example if one wants to consider SDEs driven by semimartingales – we cannot expect continuity, and, consequently, our approach breaks down. Recently, Garcia [22] has shown a way to define a family of almost compact functions which gives rise to an exponential approximation of stochastic integrals with respect to semimartingales. Using the results of Section 4.2, Garcia proves the following

7.12 Theorem ([22, Theorem 8.2]) *Let $(Y^\varepsilon)_{\varepsilon>0}$ be a family of semimartingales satisfying a large deviation principle in $D[0,1]$ with rate function I_0. Suppose that $(Y^\varepsilon)_{\varepsilon>0}$ is* uniformly exponentially tight*, i. e. for any $\alpha > 0$ there exists $R > 0$ such that*

$$\limsup_{\varepsilon \to 0} \varepsilon \log \sup_{Z \in \mathcal{S}} \mathbb{P}\left[\sup_{0 \leq t \leq 1} \left|\int_0^t Z_{s-}\,dY_s^\varepsilon\right| \geq R\right] \leq -\alpha$$

where $\mathcal{S}$ denotes the collection of simple functions Z for which $\sup_{0 \leq t \leq 1} |Z(t)| \leq 1$. Let $b : \mathbb{R} \to \mathbb{R}$ be a bounded, Lipschitz continuous function, and assume that X^ε is a solution of the SDE

$$dX_t^\varepsilon = b(X_{t-}^\varepsilon)\,dY_t^\varepsilon, \qquad X_0^\varepsilon = x.$$

If $(X^\varepsilon, Y^\varepsilon)_{\varepsilon>0}$ is exponentially tight, then $(X^\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle with rate function

$$I(f) = \inf\left\{I_0(g); g \in \mathrm{BV}[0,1] : f(t) = x + \int_0^t b(f(s))\,dg(s), t \in [0,1]\right\}.$$

The result has been further generalized by Ganguly [19] to infinite-dimensional semimartingales.

8 Conclusion

The purpose of our work was to present large deviation results for Lévy processes and solutions of Lévy-driven SDEs. We did not make any attempt to state the results in their most general form; instead we intended to demonstrate typical large deviation techniques.

Starting from the Gärtner-Ellis approach we saw that exponentially good approximations play an important role in large deviation theory. In many cases they allow us to reduce our investigations to processes which are easier to handle, such as polygons or simple functions. Secondly, exponentially good approximations give rise to extended versions of the contraction principle. These extensions turn out to be crucial when considering large deviations for SDEs or – more generally – stochastic integrals. Moreover, the change-of-measure technique proved itself a fundamental tool for large deviation lower bounds.

Chapter 7 left us with some open questions. First of all, it was unsatisfying that we did not succeed in identifying the rate functions obtained in Section 7.1 and Section 7.2. Possibly, the proof would have revealed a connection between those totally different approaches. Furthermore, it would be interesting to see in which way the results of Section 7.2 can be extended:

- Does the approach remain valid if we consider ε-dependent coefficients, i. e. SDEs of the form
$$dX_t = b_\varepsilon(X_{t-})\,dt + \sqrt{\varepsilon}\sigma_\varepsilon(X_{t-})dB_t + \eta_\varepsilon(X_{t-})\,dL_t^\varepsilon,$$
where $b_\varepsilon \to b$, $\sigma_\varepsilon \to \sigma$, $\eta_\varepsilon \to \eta$ uniformly as $\varepsilon \to 0$?

- In Chapter 5 we have shown that scaled Lévy processes of the form $L(\varepsilon^{-1}\cdot)/S(\varepsilon^{-1})$ obey a large deviation principle as $\varepsilon \to 0$ if the scaling function S satisfies the growth conditions
$$\varepsilon S(\varepsilon^{-1}) \xrightarrow{\varepsilon\to 0} 0 \quad \text{and} \quad \sqrt{\varepsilon}S(\varepsilon^{-1}) \xrightarrow{\varepsilon\to 0} \infty,$$
and used this result in order to establish the law of iterated logarithm for Lévy processes. Can we modify the proof such that this large deviation principle carries over to SDEs driven by Lévy processes? And, secondly, does it give rise to a law of iterated logarithm? For SDEs driven by Brownian motion this question was discussed by Baldi [2].

A

Appendix

A.1 Lemma (Etemadi's inequality [16]) *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent random variables. Set $S_n := \sum_{j=1}^n X_j$. Then*

$$\mathbb{P}\left(\max_{1\leq j\leq n}|S_j| \geq 3r\right) \leq 3\max_{1\leq j\leq n}\mathbb{P}(|S_j|\geq r) \qquad \text{for all} \quad r\geq 0.$$

Proof. For

$$A_j := \left\{\max_{1\leq k<j}|S_k| < 3r, |S_j|\geq 3r\right\}, \qquad j=1,\dots,n$$

we have

$$\left\{\max_{1\leq j\leq n}|S_j|\geq 3r\right\} = \dot{\bigcup_{j=1}^n} A_j.$$

Consequently, by the independence of the random variables,

$$\begin{aligned}\mathbb{P}\left(\max_{1\leq j\leq n}|S_j|\geq 3r\right) &\leq \mathbb{P}(|S_n|\geq r) + \sum_{j=1}^{n-1}\mathbb{P}(A_j\cap\{|S_n|<r\})\\ &\leq \mathbb{P}(|S_n|\geq r) + \sum_{j=1}^{n-1}\mathbb{P}(A_j)\mathbb{P}(|S_n-S_j|>2r)\\ &\leq \mathbb{P}(|S_n|\geq r) + \max_{1\leq j\leq n}\mathbb{P}(|S_n-S_j|>2r)\\ &\leq 3\max_{1\leq j\leq n}\mathbb{P}(|S_j|\geq r).\end{aligned}$$

□

A.2 Corollary *Let $(X_t)_{t\geq 0}$ be a Lévy process. Then*

$$\mathbb{P}\left(\sup_{u\in[s,t]}|X_u-X_s|\geq 3r\right) \leq 3\sup_{u\in[s,t]}\mathbb{P}(|X_u-X_s|\geq r) \qquad \text{for all} \quad t>s\geq 0, r\geq 0.$$

Proof. Since X has càdlàg sample paths, we find

$$\sup_{u\in[s,t]}|X_u-X_s| = \sup_{u\in([s,t]\cap\mathbb{Q})\cup\{t\}}|X_u-X_s|.$$

Therefore, the claim follows readily from Lemma A.1 and the monotone convergence theorem. □

A.3 Theorem (Duality lemma [10, Lemma 4.5.8]) *Let (M,d) be a metric space, and let $f: M\to(-\infty,\infty]$ be a lower semicontinuous, convex function. For*

$$f^*(\lambda) := \sup_{x\in M}\big(\langle\lambda,x\rangle - f(x)\big), \qquad \lambda\in M^*$$

we have $f = (f^*)^*$, *i. e.*

$$f(x) = \sup_{\lambda\in M^*} \big(\langle\lambda, x\rangle - f^*(\lambda)\big), \qquad x \in M.$$

Proof. Obviously, the definition of f^* implies

$$\sup_{\lambda\in M^*} \big(\langle\lambda, x\rangle - f^*(\lambda)\big) \leq \sup_{\lambda\in M^*} \big(\langle\lambda, x\rangle - (\langle\lambda, x\rangle - f(x))\big) = f(x) \qquad \text{for all} \quad x \in M.$$

Consequently, $(f^*)^* \leq f$ holds in any case. The inequality '$\geq$' is basically a consequence of the Hahn-Banach theorem. For a detailed proof we refer the reader to [10, Lemma 4.5.8] or [35, Theorem 12.2]. □

A.4 Lemma *Let* (M, d) *be a metric space and* $f : M \to [-\infty, \infty]$ *a proper, convex function.*

(i). $\lambda \in \partial f(x) \Rightarrow x \in \partial f^*(\lambda)$

(ii). f *is lower semicontinuous at* $x \in M$ *if, and only if,* f *is weakly lower semicontinuous at* $x \in M$, *i. e. for any sequence* $(x_n)_{n\in\mathbb{N}}$ *in* M, $x_n \to x$ *in* $\sigma(M, M^*)$, *we have*

$$f(x) \leq \liminf_{n\to\infty} f(x_n).$$

(iii). If f *is Gâteaux differentiable at* x *with Gâteaux derivative* $D_x \in M^*$, *then* $\partial f(x) = \{D_x\}$.

Proof. (i). Let $\lambda \in \partial f(x)$. Then, by definition, $f(y) \geq f(x) + \langle\lambda, y - x\rangle$ for all $y \in M$. Hence,

$$f^*(\lambda) = \sup_{y\in M} \big(\langle\lambda, y\rangle - f(y)\big) \leq \sup_{y\in M} \big(\langle\lambda, y\rangle - (f(x) + \langle\lambda, y - x\rangle)\big) = \langle\lambda, x\rangle - f(x).$$

Thus,

$$f^*(\mu) - f^*(\lambda) \geq \langle\mu, x\rangle - f(x) - (\langle\lambda, x\rangle - f(x)) = \langle\mu - \lambda, x\rangle \qquad \text{for all} \quad \mu \in M^*.$$

This shows $x \in \partial f^*(\lambda)$.

(ii). Note that f is convex and (weakly) lower semicontinuous if, and only if, its sublevel sets

$$\Phi(r) := \{x \in M; f(x) \leq r\}, \qquad r \in \mathbb{R}$$

are (weakly) closed and convex. Therefore, the claim follows from the widely known fact that any convex set is closed if, and only if, it is weakly closed.

(iii). Fix $y \in M$. Since the mapping

$$(0, \infty) \ni t \mapsto \frac{f(x + ty) - f(x)}{t}$$

is increasing, we have

$$\langle D_x, y\rangle = \lim_{t\to 0} \frac{f(x+ty)-f(x)}{t} = \inf_{t>0} \frac{f(x+ty)-f(x)}{t}.$$

Clearly, this implies $D_x \in \partial f(x)$. On the other hand,

$$\langle D_x, y\rangle = \lim_{t\downarrow 0} \frac{f(x+ty)-f(x)}{t} \geq \lim_{t\downarrow 0} \frac{\langle \lambda, (x+ty)-x\rangle}{t} = \langle \lambda, y\rangle$$

for any $\lambda \in \partial f(x)$. Repeating the argumentation for $t \uparrow 0$, we find $D_x = \lambda$. This finishes the proof. □

A.5 Theorem ([6]) *Let $(M, \|\cdot\|)$ be a Banach space and $f : M \to \mathbb{R}$ a lower semicontinuous, proper, convex function. For any $x \in \operatorname{dom} f$ and $\varepsilon > 0$ there exists $y \in M$ such that $\partial f(y) \neq \varnothing$ and*

$$\|x-y\| \leq \varepsilon \qquad\qquad |f(x)-f(y)| \leq \varepsilon.$$

A.6 Lemma ([6]) *Let $(M, \|\cdot\|)$ be a Banach space and $f : M \to \mathbb{R}$ a lower semicontinuous, proper, convex function. Define*

$$\partial_\varepsilon f(x) := \{x^* \in M^*; \forall y \in M : f(y) \geq (f(x)-\varepsilon) + \langle y-x, x^*\rangle\}, \quad \varepsilon > 0, x \in M. \tag{A.1}$$

For $x^* \in \partial_\varepsilon f(x)$ and $\gamma > 0$ there exist $\bar{x} \in M$ and $\bar{x}^* \in M^*$ such that $\bar{x}^* \in \partial f(\bar{x})$,

$$\|x-\bar{x}\| \leq \gamma, \qquad\qquad \|x^*-\bar{x}^*\| \leq \frac{\varepsilon}{\gamma}.$$

Proof. We define a relation on $\operatorname{dom} f$ by

$$y \leq z :\Leftrightarrow \frac{\varepsilon}{\gamma}\|y-z\| \leq (f(y)-\langle x^*, y\rangle) - (f(z)-\langle x^*, z\rangle) \tag{A.2}$$

Obviously, $\leq$ is a partial order on $\operatorname{dom} f$. In order to apply Zorn's lemma, we show that any totally ordered subset $\{x_\alpha; \alpha \in I\} \subseteq \operatorname{dom} f$ has an upper bound. Since $x^* \in \partial_\varepsilon f(x)$,

$$f(x_\alpha) - \langle x^*, x_\alpha\rangle \overset{(A.2)}{\geq} f(x_\beta) - \langle x^*, x_\beta\rangle \overset{(A.1)}{\geq} f(x) - \langle x^*, x\rangle - \varepsilon$$

for any $\alpha \leq \beta$. Consequently, there exists $a > -\infty$ such that $f(x_\alpha) - \langle x_\alpha, x^*\rangle \downarrow a$ for $\alpha \uparrow$. In particular, we can choose $\alpha = \alpha(\delta)$ such that

$$f(x_\beta) - \langle x^*, x_\beta\rangle \leq a + \frac{\delta\varepsilon}{\gamma} \qquad \text{for all} \quad \beta \geq \alpha.$$

By (A.2), this shows that $(x_\alpha)_{\alpha\in I}$ is a Cauchy net in M. Since M is a Banach space, there exists $\bar{x} \in M$ such that $x_\alpha \to \bar{x}$. In particular, $x_\alpha \leq \bar{x}$ for all $\alpha \in I$.
Applying Zorn's lemma, we find that $\{y \in \operatorname{dom} f; x \leq y\}$ has at least one maximal element $\bar{x}$. Then, $x \leq \bar{x}$ entails

$$\frac{\varepsilon}{\gamma}\|x-\bar{x}\| \leq f(x) - f(\bar{x}) - \langle x^*, x-\bar{x}\rangle \leq \varepsilon$$

by the definition of ∂f_ε, i.e. $\|x - \bar{x}\| \le \gamma$. Moreover, since $\bar{x}$ is the maximal element, we have

$$\frac{\varepsilon}{\gamma}\|z - \bar{x}\| \ge (f(\bar{x}) - \langle \bar{x}, x^*\rangle) - (f(z) - \langle z, x^*\rangle) \qquad \text{for all} \quad z \ne \bar{x}.$$

Define

$$H_1 := \{(y, c) \in M \times \mathbb{R}; c \ge f(\bar{x} + y) - f(\bar{x}) - \langle x^*, y\rangle\},$$
$$H_2 := \left\{(y, c) \in M \times \mathbb{R}; c < -\frac{\varepsilon}{\gamma}\|y\|\right\},$$

then $H_1 \cap H_2 = \varnothing$. Note that H_1 is closed and convex as f is lower semicontinuous and convex. By Hahn-Banach's theorem, there exists $z^* \in M^*$ such that

$$-\frac{\varepsilon}{\gamma}\|y\| \le \langle z^*, y\rangle \le f(\bar{x} + y) - f(\bar{x}) - \langle x^*, y\rangle \qquad \text{for all} \quad y \in M.$$

Setting $\bar{x}^* := x^* + z^*$ finishes the proof. □

Proof of Theorem A.5. First of all, we note that $\partial f_\varepsilon(x) \ne \varnothing$ for any $\varepsilon > 0$. Indeed: By the duality lemma A.3, we have

$$f(x) = \sup_{x^* \in M^*} \big(\langle x^*, x\rangle - f^*(x^*)\big).$$

In particular, there exists $x^* = x^*(\varepsilon)$ such that

$$f(x) - \big(\langle x^*, x\rangle - f^*(x^*)\big) \le \varepsilon.$$

By the definition of the Legendre transform, $f^*(x^*) \ge \langle x^*, y\rangle - f(y)$ for any $y \in M$. Hence, $x^* \in \partial_\varepsilon f(x)$.
Pick $x^* \in \partial_{\varepsilon/2} f(x)$ and choose $\gamma > 0$ such that $\gamma < \varepsilon$ and $\gamma\|x^*\| < \varepsilon/2$. Let $\bar{x} \in M$ and $\bar{x}^* \in M^*$ as in Lemma A.6. Then,

$$f(\bar{x}) - f(x) \le -\langle x - \bar{x}, \bar{x}^*\rangle \le \|x - \bar{x}\|\|\bar{x}^*\| \le \gamma\left(\|x^*\| + \frac{\varepsilon}{2\gamma}\right) \le \varepsilon.$$

Thus, $\bar{x} \in \operatorname{dom} f$, $f(\bar{x}) \le f(x) + \varepsilon$. This finishes the proof. □

A.7 Lemma *Let $f : \mathbb{R} \to \mathbb{R}$ be a convex, continuously differentiable function such that $f^*(\lambda) < \infty$ for all $\lambda \in \mathbb{R}$. For any $\lambda \in \mathbb{R}$ there exists $x = x(\lambda)$ such that $f'(x(\lambda)) = \lambda$ and*

$$f^*(\lambda) = x(\lambda)\lambda - f(x(\lambda)). \tag{A.3}$$

If f^* is differentiable, then

$$f'\big((f^*)'(\lambda)\big) = \lambda. \tag{A.4}$$

In particular, (A.3) *holds for $x(\lambda) = (f^*)'(\lambda)$.*

Proof. We claim that for any $\lambda \in \mathbb{R}$ there exists $x \in \mathbb{R}$ such that $f'(x) \ge \lambda$. Indeed: Suppose not; then we have $f(x) \le 2\lambda x$ for $|x| \ge R$ sufficiently large. Hence,

$$f^*(4\lambda) = \sup_{x \in \mathbb{R}} \big(2\lambda x + (2\lambda x - f(x))\big) \ge \sup_{|x| \ge R} (2\lambda x) = \infty.$$

This contradicts $f^* < \infty$. Similarly, there exists $x \in \mathbb{R}$ such that $f'(x) \le \lambda$. As f' is continuous, the intermediate value theorem shows $f'(\mathbb{R}) = \mathbb{R}$. Moreover, we see that $\lambda x - f(x) \to -\infty$ as $|x| \to \infty$. This means that the mapping attains its maximum, i. e.

$$f^*(\lambda) = \max_{x \in \mathbb{R}} \big(\lambda x - f(x)\big) = \lambda x(\lambda) - f(x(\lambda))$$

where $x(\lambda)$ satisfies $f'(x(\lambda)) = \lambda$. (Note that $x \mapsto \lambda x - f(x)$ is concave; therefore any local maximum is a global maximum.) This proves the first claim. Now suppose that f^* is differentiable. By the duality lemma A.1 and Lemma A.4(i),(iii), we have

$$\lambda = f'(x) \Leftrightarrow x = (f^*)'(\lambda).$$

Thus,

$$f'\big((f^*)'(\lambda)\big) = f'(x) = \lambda. \qquad \square$$

A.8 Lemma *Let $(L_t)_{t \ge 0}$ be a Lévy process with symbol ψ such that $\mathbb{E}e^{\lambda |L_1|} < \infty$ for all $\lambda \ge 0$, and let $f : \mathbb{R} \to \mathbb{R}$ be bounded. Set $Q(x, \xi) := \psi(f(x)\xi)$ and*

$$I(\varphi) := \begin{cases} \int_0^1 Q^*(\varphi(t), \varphi'(t))\, dt, & \varphi \in AC[0,1], \varphi(0) = 0, \\ \infty, & \textit{otherwise.} \end{cases}$$

for $\varphi \in D[0,1]$. Suppose that (S1) and (S2) hold.

(i). I is a good rate function.

(ii). For $\varphi \in D[0,1]$, $I(\varphi) < \infty$, and $\varepsilon > 0$ there exists $\delta > 0$ such that for any partition $\Pi = \{0 = t_0 < \ldots < t_m = 1\}$ for which $|\Pi| < \delta$, the piecewise linear approximation $\Pi(\varphi)$ on the grid Π satisfies $\|\varphi - \Pi(\varphi)\|_\infty \le \varepsilon$ and $I(\Pi(\varphi)) \le I(\varphi) + \varepsilon$.

Proof. To prove (i), it suffices, by the Arzèla-Ascoli theorem, to show that the sublevel sets $\Phi(r)$, $r \ge 0$, are uniformly bounded, uniformly equicontinuous, and closed. Let $\varphi \in \Phi(r)$. By the definition of the Legendre transform,

$$Q^*(\varphi(t), \varphi'(t)) \ge \xi \varphi'(t) - Q(\varphi(t), \xi) \qquad \text{for all } \ \xi \in \mathbb{R}.$$

Hence,

$$\pm\varphi'(t) \le \frac{Q^*(\varphi(t), \varphi'(t))}{\xi} + \frac{Q(\varphi(t), \pm\xi)}{\xi} \overset{(S3)}{\le} \frac{Q^*(\varphi(t), \varphi'(t))}{\xi} + \frac{\overline{Q}(\xi) \vee \overline{Q}(-\xi)}{\xi}$$

for any $\xi > 0$. Consequently, we get

$$|\varphi(t) - \varphi(s)| \le \int_s^t |\varphi'(u)|\, du \le \frac{r}{\xi} + \frac{\overline{Q}(\xi) \vee \overline{Q}(-\xi)}{\xi} |t - s|$$

for all $s, t \in [0,1]$ and $\xi > 0$. This shows that $\Phi(r)$ is uniformly bounded and uniformly equicontinuous. It remains to show that $\Phi(r)$ is closed, i. e. that I is lower semicontinuous. To this end let $(\varphi_n)_{n \in \mathbb{N}} \subseteq D[0,1]$ such that $\varphi_n \to \varphi \in D[0,1]$ uniformly and fix

$\varepsilon > 0$. We may assume $\liminf_{n\to\infty} I(\varphi_n) < \infty$. Let $(\Pi_n)_{n\in\mathbb{N}}$ be a sequence of partitions $\Pi_n = \{0 = t_0 < \ldots < t_m = 1\}$ such that $|\Pi_n| := \max_{t_j\in\Pi_n} |t_{j+1} - t_j| \to 0$ as $n \to \infty$. From the first part of the proof we know that $\{\varphi_n; n \in \mathbb{N}\}$ is uniformly equicontinuous. Since $\varphi_n \to \varphi$ uniformly, it is therefore not difficult to see that $\varphi \in AC[0,1]$ and

$$\varphi'(t) = \lim_{n\to\infty} \sum_{t_j\in\Pi_n} \frac{\varphi_n(t_{j+1}) - \varphi_n(t_j)}{t_{j+1} - t_j} \mathbb{1}_{[t_j,t_{j+1})}(t) \qquad \text{a.s.}$$

Moreover,

$$\|\varphi_n - \varphi\|_\infty + \sup_{t_j\in\Pi_n} \sup_{t\in[t_j,t_{j+1}]} |\varphi_n(t_j) - \varphi_n(t)| \leq \varepsilon \tag{A.5}$$

for $n \in \mathbb{N}$ sufficiently large. Since $Q^* \geq 0$ we obtain from Fatou's lemma

$$\begin{aligned} I(\varphi) &\leq \liminf_{n\to\infty} \sum_{t_j\in\Pi_n} \int_{t_j}^{t_{j+1}} Q^*\left(\varphi(t), \frac{\varphi_n(t_{j+1}) - \varphi_n(t_j)}{t_{j+1} - t_j}\right) dt \\ &\leq \Delta Q^*(\varepsilon) + (1 + \Delta Q^*(\varepsilon)) \liminf_{n\to\infty} \sum_{t_j\in\Pi_n} (t_{j+1} - t_j) Q^*\left(\varphi_n(t_j), \frac{\varphi_n(t_{j+1}) - \varphi_n(t_j)}{t_{j+1} - t_j}\right), \end{aligned} \tag{A.6}$$

cf. (S2). Recall that $Q^*(x,\cdot)$ is, for fixed $x \in \mathbb{R}$, a convex function. Therefore, Jensen's inequality shows

$$\begin{aligned} Q^*\left(\varphi_n(t_j), \frac{\varphi_n(t_{j+1}) - \varphi_n(t_j)}{t_{j+1} - t_j}\right) &= Q^*\left(\varphi_n(t_j), \int_{t_j}^{t_{j+1}} \varphi_n'(t) \frac{dt}{t_{j+1} - t_j}\right) \\ &\leq \int_{t_j}^{t_{j+1}} Q^*(\varphi_n(t_j), \varphi_n'(t)) \frac{dt}{t_{j+1} - t_j}. \end{aligned} \tag{A.7}$$

Combining (A.5), (A.6) and (A.7) yields

$$\begin{aligned} I(\varphi) &\leq \Delta Q^*(\varepsilon) + (1 + \Delta Q^*(\varepsilon)) \liminf_{n\to\infty} \sum_{t_j\in\Pi_n} \int_{t_j}^{t_{j+1}} Q^*(\varphi_n(t_j), \varphi_n'(t))\, dt \\ &\leq \Delta Q^*(\varepsilon)(2 + \Delta Q^*(\varepsilon)) + (1 + \Delta Q^*(\varepsilon))^2 \liminf_{n\to\infty} \int_0^1 Q^*(\varphi_n(t), \varphi_n'(t))\, dt \\ &= \Delta Q^*(\varepsilon)(2 + \Delta Q^*(\varepsilon)) + (1 + \Delta Q^*(\varepsilon))^2 \liminf_{n\to\infty} I(\varphi_n). \end{aligned}$$

By assumption, $\Delta Q^*(\varepsilon) \to 0$ as $\varepsilon \to 0$; this proves the lower semicontinuity; hence, (i). Now let $\varphi \in D[0,1]$, $I(\varphi) < \infty$, and $\varepsilon > 0$. Obviously, $\varphi \in \operatorname{dom} I \subseteq AC[0,1]$. For fixed $\varrho < \varepsilon$ we can choose $\delta > 0$ such that $|\varphi(t) - \varphi(s)| \leq \varrho$ for $|t - s| < \delta$. In particular, $\|\varphi - \Pi(\varphi)\|_\infty \leq \varepsilon$ holds for any partition $\Pi = \{0 = t_0 < \ldots < t_m = 1\}$ of $[0,1]$ such that $|\Pi| < \delta$. Using similar arguments as in the first part of this proof, we get

$$\begin{aligned} I(\Pi(\varphi)) &= \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} Q^*\left(\Pi(\varphi)(t), \frac{\varphi(t_{j+1}) - \varphi(t_j)}{t_{j+1} - t_j}\right) dt \\ &\leq \Delta Q^*(\varrho) + (1 + \Delta Q^*(\varrho)) \sum_{j=0}^{m-1} (t_{j+1} - t_j) Q^*\left(\varphi(t_j), \frac{\varphi(t_{j+1}) - \varphi(t_j)}{t_{j+1} - t_j}\right) \\ &\leq \Delta Q^*(\varrho) + (1 + \Delta Q^*(\varrho)) \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} Q^*\left(\varphi(t_j), \varphi'(t)\right) dt \\ &\leq \Delta Q^*(\varrho)(2 + \Delta Q^*(\varrho)) + (1 + \Delta Q^*(\varrho))^2 I(\varphi). \end{aligned}$$

By (S2), $\Delta Q^*(\varrho) \to 0$ as $\varrho \to 0$; therefore, the claim follows by choosing ϱ sufficiently small. $\square$

A.9 Lemma (Integration by parts) *Let* $\alpha \in \mathrm{BV}[0,1] \cap D[0,1]$, $f \in L^1([0,1], \lambda|_{[0,1]})$, *and set* $F(t) := \int_0^t f(s)\,ds$ *for* $t \in [0,1]$. *Then*

$$\int_0^1 F\,d\alpha = \int_0^1 f(s)(\alpha(1) - \alpha(s))\,ds.$$

Proof. Set $t_j^n := j/n$, $j = 0, \dots, n$. By definition of the left-hand side and the dominated convergence theorem,

$$\begin{aligned}\int_0^1 F\,d\alpha &= \lim_{n\to\infty} \sum_{j=0}^{n-1} F\left(t_{j+1}^n\right)\left(\alpha(t_{j+1}^n) - \alpha(t_j^n)\right)\\ &= \lim_{n\to\infty} \sum_{k=0}^{n-1}\sum_{j=k}^{n-1} \left(F(t_{k+1}^n) - F(t_k^n)\right)\left(\alpha(t_{j+1}^n) - \alpha(t_j^n)\right)\\ &= \lim_{n\to\infty} \sum_{k=0}^{n} \int_{t_k^n}^{t_{k+1}^n} f(s)(\alpha(1) - \alpha(t_k^n))\,ds\\ &= \int_0^1 f(s)(\alpha(1) - \alpha(s))\,ds.\end{aligned}$$

Note that the dominated convergence theorem applies since $\alpha \in D[0,1]$ has at most countable discontinuity points. $\square$

A.10 Lemma *Let* $(M, \|\cdot\|)$ *be a normed space and* $\mathcal{B}$ *a* σ*-algebra on* M *such that (C1) and (C7) hold. Let* $K \subseteq M$ *be compact and* $\delta \geq 0$. *Then* $K + B[0,\delta] \in \mathcal{B}$. *In particular, (C2) holds.*

Proof. From (C1) and (C7) we see that $B(x,r) \in \mathcal{B}$ for any $x \in M$, $r > 0$. Since K is compact, there exists $F_n \subseteq M$ finite such that

$$K \subseteq \bigcup_{x\in F_n} B\left(x, \frac{1}{n}\right).$$

It is not difficult to show that

$$K + B[0,\delta] = \bigcap_{n\in\mathbb{N}} \bigcup_{x\in F_n} B\left(x, \frac{1}{n} + \delta\right).$$

Since the right side is contained in $\mathcal{B}$, the claim follows. $\square$

Bibliography

[1] Azencott, R.: Grandes deviations et applications. In: Hennequin, P. L. (ed.): *Ecole d'Eté de Probabilités de Saint-Flour VIII-1978*. Springer, Lecture Notes in Mathematics **774**, Berlin 1980, 1–176.

[2] Baldi, P.: Large Deviations and Functional Iterated Logarithm Law for Diffusion processes. *Probability Theory and Related Fields* **71** (1986), 435–453.

[3] Baldi, P., Chaleyat-Maurel, M.: An Extension of Ventsel-Freidlin estimates. In: Korezlioglu, H., Ustunel, A. S.: *Stochastic Analysis and Related Topics*. Springer, Lecture Notes in Mathematics **1316**, Berlin 1988, 305–327.

[4] Billingsley, P.: *Convergence of Probability Measures*. Wiley, New York 1999 (2nd edition).

[5] Böttcher, B., Schilling, R., Wang, J.: *Lévy Matters III. Lévy-Type Processes: Construction, Approximation and Sample Path Properties*. Springer, Lecture Notes in Mathematics **2099**, Cham 2013.

[6] Brønsted, A., Rockafellar, R. T.: On the Subdifferentiability of Convex Functions. *Proceedings of the American Mathematical Society* **16** (1965), 605–611.

[7] Cutland, N. J.: An extension of the Ventcel-Freidlin large deviation principle. *Stochastics* **24** (1988), 121–149.

[8] de Acosta, A.: Large deviations for vector-valued Lévy processes. *Stochastic Processes and their Applications* **41** (1993), 75–115.

[9] de Acosta, A.: A general non-convex large deviation result with applications to stochastic equations. *Probability Theory and Related Fields* **118** (2000), 483–521.

[10] Dembo, A., Zeitouni, O.: *Large Deviations Techniques and Applications*. Springer, Berlin 2010 (2nd edition).

[11] Dettweiler, E.: *Grundlagen der stochastischen Analysis*. 2006.

[12] Dynkin, E. B.: *Die Grundlagen der Theorie der Markoffschen Prozesse*. Springer, Berlin 1961.

[13] Eichelsbacher, P., Grunwald, M.: Exponential tightness can fail in the strong topology. *Statistics & Probability Letters* **41** (1999), 83–86.

[14] Ethier, S. N., Kurtz, T. G.: *Markov Processes. Characterization and Convergence*. Wiley, New York 1986.

[15] Ellis, R. S.: Large Deviations for a General Class of Random Vectors. *Annals of Probability* **12** (1984), 1–12.

[16] Etemadi, N.: On some Classical Results in Probability Theory. *Sankhya A* **47** (1985), 215–221.

[17] Feng, J., Kurtz, T. G.: *Large Deviations for Stochastic Processes.* American Mathematical Society, Providence (RI) 2006.

[18] Freidlin, M. I., Wentzell, A. D.: *Random Perturbations of Dynamical Systems.* Springer, New York 1998 (2nd edition). Russian original.

[19] Ganguly, A.: Large deviation principle for stochastic integrals and stochastic differential equations driven by infinite-dimensional semimartingales. (2011). – arXiv: 1101.5515.

[20] Gärtner, J.: On Large Deviations from the Invariant Measure. *Theory of Probability & Its Applications* **22** (1977), 24–39.

[21] Garcia, J.: An Extension of the Contraction Principle. *Journal of Theoretical Probability* **17** (2004), 403–434.

[22] Garcia, J.: A Large Deviation Principle for Stochastic Integrals. *Journal of Theoretical Probability* **21** (2008), 476–501.

[23] Hollander, F. den: *Large Deviations.* American Mathematical Society, Providence (RI) 2000.

[24] Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam 1989.

[25] Krylov, N. V.: *Introduction to the Theory of Random Processes.* American Mathematical Society, Providence (RI) 2002.

[26] Kulik, A. M., Soboleva, D. D.: Large deviations for one-dimensional SDE with discontinuous diffusion coefficient. *Theory of Stochastic Processes* **18** (2012), 101–110.

[27] Kurdila, A., Zabarankin, M.: *Convex Functional Analysis.* Birkhäuser, Basel 2005.

[28] Liptser, R. S., Puhalskii, A. A.: Limit theorems on large deviations for semimartingales. *Stochastics and Stochastic Reports* **38** (1992), 201–249.

[29] Lynch, J., Sethuraman, J.: Large Deviations for Processes with Independent Increments. *Annals of Probability* **15** (1987), 610–627.

[30] Millet, A., Nualart, D., Sanz, M.: Large Deviations for a Class of Anticipating Stochastic Differential Equations. *Annals of Probability* **20** (1992), 1902–1931.

[31] Mogulskii, A. A.: Large Deviations for Processes with Independent Increments. *Annals of Probability* **21** (1993), 202–215.

[32] Protter, P. E.: *Stochastic Integration and Differential Equations.* Springer, Berlin 2005 (2nd edition).

[33] Puhalskii, A. A.: On functional principle of large deviations. In: Sazonov, V., Shervashidze, T. (eds.): *New Trends in Probability and Statistics* Bd. 1. VSP/Moks'las, Zeist 1991, 198–218.

[34] Puhalskii, A. A.: *Large Deviations and Idempotent Probability.* Chapman & Hall, London 2001.

[35] Rockafellar, R. T.: *Convex Analysis.* Princeton University Press, Princeton (NJ) 1997.

[36] Sato, K.-I.: *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge 2005.

[37] Schilling, R. L.: *Measures, Integrals and Martingales.* Cambridge University Press, Cambridge 2007 (2nd edition).

[38] Schilling, R. L., Partzsch, L.: *Brownian Motion. An Introduction to Stochastic Processes.* De Gruyter, Berlin 2012.

[39] Schilling, R. L., Schnurr, A.: The Symbol Associated with the Solution of a Stochastic Differential Equation. *Electronic Journal of Probability* **15** (2010), 1369–1393.

[40] Wentzell, A. D.: Rough limit theorems on large deviations for Markov stochastic processes, I. *Probability Theory & Its Applications* **21** (1976), 227–242. Rough limit theorems on large deviations for Markov stochastic processes, II. *Probability Theory & Its Applications* **21** (1976), 499–512. Russian original.

ERKLÄRUNG

Hiermit erkläre ich, dass ich die am heutigen Tag eingereichte Diplomarbeit zum Thema "Large Deviations for Lévy(-Type) Processes" unter Betreuung von Prof. Dr. rer. nat. René L. Schilling selbstständig erarbeitet, verfasst und Zitate kenntlich gemacht habe. Andere als die angegebenen Hilfsmittel wurden von mir nicht benutzt.

Datum Unterschrift