Large Deviations for Stochastic Processes

LMS/EPSRC Short Course: Stochastic Stability, Large Deviations and Coupling Methods, Heriot-Watt University, Edinburgh,
4-9 September 2006

Anatolii A. Puhalskii

1 LDP in metric spaces

The purpose of these lectures is to introduce you to the basics of large deviation theory. The emphasis will be on the use of compactness ideas (more extensive results are in Puhalskii [13]). Other approaches to large deviation theory are considered in Dembo and Zeitouni [3], den Hollander [4], Deuschel and Stroock [6], Dupuis and Ellis [7], Freidlin and Wentzell [9], Kallenberg [12], Shwartz and Weiss [14], Varadhan [16].

To give you an idea of what kind of problems we are up against let me consider a simple but representative example. Consider a standard Brownian motion $(B(t),\, t\in\mathbb{R}_+)$. Let $X_n(t) = B(t)/\sqrt{n}$, where n is a large parameter. Since $X_n(t)-X_n(s)$ are Gaussian with mean 0 and variance $(t-s)/n$, the sequence of these r.v. obeys the LDP in $\mathbb{R}_+$ (for rate n) with rate function $x^2/\bigl(2(t-s)\bigr)$. Let us now consider a vector $(X_n(t_i),\, i=1,2,\dots,k)$, where $t_1<t_2<\ldots<t_k$. By the property of independence of increments, the k-dimensional vector $(X_n(t_i)-X_n(t_{i-1}),\, i=1,2,\dots,k)$, where $t_0=0$, obeys the LDP (for rate n) with rate function $1/2\sum_{i=1}^k x_i^2/(t_i-t_{i-1})$. Therefore, by the continuous mapping (contraction) principle, the vectors $(X_n(t_i),\, i=1,2,\dots,k)$ obey the LDP with $1/2\sum_{i=1}^k (x_i-x_{i-1})^2/(t_i-t_{i-1})$. This means that for small enough $\epsilon>0$ and large n

$$\frac{1}{n}\log\mathbf{P}\bigl(|X_n(t_i)-x_i|\le\epsilon,\, i=1,2,\dots,k\bigr)\approx -\frac{1}{2}\sum_{i=1}^k\Bigl(\frac{x_i-x_{i-1}}{t_i-t_{i-1}}\Bigr)^2(t_i-t_{i-1}). \tag{1.1}$$

Therefore, if $\mathbf{x}=(\mathbf{x}(t)\,,t\in[0,1])$ is an absolutely continuous function, then by considering finer and finer subdivisions of the $[0,1]$ interval one should expect that

$$\frac{1}{n}\log\mathbf{P}\bigl(\sup_{t\in[0,1]}|X_n(t)-\mathbf{x}(t)|\le\epsilon\bigr)\approx -\frac{1}{2}\int_0^1\dot{\mathbf{x}}(t)^2\,dt,$$

so that the rate function should be $(1/2)\int_0^1\dot{\mathbf{x}}(t)^2\,dt$ for absolutely continuous functions starting at zero. The right-hand side of (1.1) tends to $-\infty$ if $\mathbf{x}$ is either not absolutely continuous or if it doesn't start at zero rendering the action functional equal infinity at these functions. This is the kind of problems we will be looking at.

We start with general properties of the LDP. Consideration will be restriced to the setting of metric spaces as being sufficient for most of the applications. Generalisations, though, can be obtained along similar lines. Let $\mathbb{U}$ be a metric space. A function $f:\mathbb{U}\to[-\infty,\infty]$ is said to be lower semicontinuous if the sets $\{z:\ f(z)\le x\}$ are closed. If the sets $\{z:\ f(z)\ge x\}$ are closed, then f is said to be upper semicontinuous.

Exercise 1.1. *1. Show that the indicator function of an open (respectively, closed) set is a lower semicontinuous (respectively, upper semicontinuous) function.*

2. Show that a function f is lower semicontinuous (respectively, upper semicontinuous) if and only if $\liminf_{n\to\infty} f(z_n) \geq f(z)$ (respectively, $\limsup_{n\to\infty} f(z_n) \leq f(z)$) whenever $z_n \to z$.

3. Show that a function f is lower semicontinuous if and only if for each z we have that $\lim_{\epsilon\to 0} \inf_{y\in B_\epsilon(z)} f(y) = f(z)$, where $B_\epsilon(z)$ denotes the ϵ-ball about z.

4. Show that a function f is lower semicontinuous if and only if for each decreasing sequence K_n of compact sets we have that $\lim_{n\to\infty} \inf_{z\in K_n} f(z) = \inf_{z\in\cap_n K_n} f(z)$. (This property extends to directed families of sets.)

5. Show that lower semicontinuous functions attain infima on compact sets.

6. Show that if f is lower semicontinuous (respectively, upper semicontinuous) and $g : \mathbb{R} \to \mathbb{R}$ is continuous and non-decreasing, then $g \circ f$ is lower semicontinuous (respectively, upper semicontinuous)

Let us say that f is lower compact if the sets $\{z : f(z) \leq x\}$ are compact. It is upper compact if the sets $\{z : f(z) \geq x\}$ are compact.

Exercise 1.2. *1. Show that a function f is lower compact if and only if for each decreasing sequence F_n of closed sets $\lim_{n\to\infty} \inf_{z\in F_n} f(z) = \inf_{z\in\cap_n F_n} f(z)$. (We assume that $\inf_\emptyset = \infty$.) If f is lower compact, then the latter equality holds for decreasing directed families of closed sets.*

2. Show that lower compact functions attain minima on closed sets.

Let $\{\mathbf{P}_n,\, n \in \mathbb{N}\}$ be a sequence of probability measures on $(\mathbb{U}, \mathcal{B}(\mathbb{U}))$, where $\mathcal{B}(\mathbb{U})$ denotes the Borel σ-algebra on $\mathbb{U}$. Let $\mathbf{I} : \mathbb{U} \to [0,\infty]$ be lower compact and $r_n \to \infty$ as $n \to \infty$. We will say that this sequence obeys a large deviation principle (LDP) for rate r_n with rate function (or action functional, or deviation function) $\mathbf{I}$ if the following two asymptotic bounds hold

$$\limsup_{n\to\infty} \frac{1}{r_n} \log \mathbf{P}_n(F) \leq -\inf_{z\in F} \mathbf{I}(z) \text{ for all closed sets } F \tag{1.2}$$

and

$$\liminf_{n\to\infty} \frac{1}{r_n} \log \mathbf{P}_n(G) \geq -\inf_{z\in G} \mathbf{I}(z) \text{ for all open sets } G \tag{1.3}$$

An equivalent definition: for arbitrary measurable H

$$\limsup_{n\to\infty} \frac{1}{r_n} \log \mathbf{P}_n(H) \leq -\inf_{z\in \mathrm{cl}\, H} \mathbf{I}(z)$$

and

$$\liminf_{n\to\infty} \frac{1}{r_n} \log \mathbf{P}_n(H) \geq -\inf_{z\in\ \mathrm{int}\, H} \mathbf{I}(z)$$

It can be used to define the LDP for non Borel σ-algebras.

Exercise 1.3. *Show the equivalence.*

We will also say that a sequence X_n of random variables with values in $\mathbb{U}$ obeys the LDP with $\mathbf{I}$ if so does the sequence of the laws of the X_n.

Remark 1.1. A terminology note.

1. Varadhan [16] uses the name “rate function” to refer to a lower compact function. Now, the name *good rate function* seems to be in wider use, while the name “rate function” is reserved for lower semicontinuous functions. It was Deuschel and Stroock [5] who thought these functions were “good”. A better name, in my opinion, is “a tight rate function” since the definition implies that given arbitrary $A > 0$ there exists a compact K such that $-\inf_{z\in K^c} \mathbf{I}(z) \le -A$.

2. If (1.2) holds for compact sets only, one speaks of *weak* LDP and the definition above is referred to as *full* LDP.

3. If (1.2) and (1.3) hold with some function $\mathbf{I}$, then it can always be assumed lower semicontinuous by considering instead its lower semicontinuous regularisation defined by $\underline{\mathbf{I}}(z) = \liminf_{y\to z} \mathbf{I}(y)$. Then $\inf_{z\in F} \underline{\mathbf{I}}(z) \le \inf_{z\in F} \mathbf{I}(z)$ and $\inf_{z\in G} \underline{\mathbf{I}}(z) = \inf_{z\in G} \mathbf{I}(z)$.

4. Another form of the LDP uses a small “continuous” parameter ϵ which tends to zero in place of $1/r_n$, the probability measures being indexed by ϵ accordingly.

Exercise 1.4. *Verify the claim in part 3.*

It is convenient to denote $\mathbf{I}(\Gamma) = \inf_{z\in\Gamma} \mathbf{I}(z)$. Let also $B_\epsilon(z)$ denote the closed ϵ-ball about z.

Lemma 1.1. *Let* (1.2) *and* (1.3) *hold for some lower semicontinuous* $\mathbf{I}$*. Then*

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{r_n} \log \mathbf{P}_n(B_\epsilon(z)) = \lim_{\epsilon\to 0} \liminf_{n\to\infty} \frac{1}{r_n} \log \mathbf{P}_n(B_\epsilon(z)) = -\mathbf{I}(z)$$

In particular, $\mathbf{I}$ *is specified uniquely.*

Proof. The assertion follows by the inequalities

$$\limsup_{n\to\infty} \frac{1}{r_n} \log \mathbf{P}_n(B_\epsilon(z)) \le -\mathbf{I}(B_\epsilon(z)),$$
$$\liminf_{n\to\infty} \frac{1}{r_n} \log \mathbf{P}_n(B_\epsilon(z)) \ge -\mathbf{I}(B_{\epsilon/2}(z))$$

□

We call a set $H \subset \mathbb{U}$ $\mathbf{I}$–continuous if $\mathbf{I}\big(\operatorname{int} H\big) = \mathbf{I}\big(\operatorname{cl} H\big)$. If the LDP with $\mathbf{I}$ holds, then for every $\mathbf{I}$-continuous set H

$$\lim_{n\to\infty} \frac{1}{r_n} \log \mathbf{P}_n(H) = -\mathbf{I}(H)$$

Exercise 1.5. *Check the claim. What about the converse assertion ?*

Exercise 1.6 (den Hollander)**.** *Determine whether the LDP is satisfied for rate* n*.*

1. $\mathbb{U} = \mathbb{R}$*, the* $\mathbf{P}_n$ *are uniform on* $[-n, n]$*,*

2. $\mathbb{U} = \mathbb{R}$*, the* $\mathbf{P}_n$ *are uniform on* $[-1/n, 1/n]$*,*

3. $\mathbb{U} = [-1, 1]$, *the* $\mathbf{P}_n$ *are uniform on* $[-1, 1]$.

Exercise 1.7 (O'Brien)**.** *Let* Z_n *be binomially distributed random variables with parameters* n *and* p_n *and* $\mathbf{P}_n$ *be the laws of* $Z_n/(np_n)$*. Suppose that* $\lim_{n\to\infty} p_n = 0$ *and* $\lim_{n\to\infty} np_n = \infty$*. Show that the* $\mathbf{P}_n$ *satisfy an LDP on* $\mathbb{R}$ *for rate* np_n *and with rate function* $\mathbf{I}(z) = z \log z - z + 1$, $z \geq 0$, *and* $\mathbf{I}(z) = \infty$, $z < 0$.

Here is the first property.

Theorem 1.1 (Contraction (continuous mapping) principle I)**.** *Let* $\mathbb{U}'$ *be a metric space and* $f : \mathbb{U} \to \mathbb{U}'$ *be a continuous function. If the* $\mathbf{P}_n$ *obey the LDP with* $\mathbf{I}$*, then the image-measures* $\mathbf{P}_n \circ f^{-1}$ *obey the LDP on* $\mathbb{U}'$ *with* $\mathbf{I}'$*, where* $\mathbf{I}'(z') = \inf_{z \in f^{-1}(z')} \mathbf{I}(z')$.

Proof. For $\Gamma \subset \mathbb{U}'$, we have that $\mathbf{I}'(\Gamma') = \mathbf{I}(f^{-1}(\Gamma'))$, which implies that $\mathbf{I}'$ is lower compact by exercise 1.2(1). The bounds in the definition follow immediately. □

In what follows, we will denote $\mathbf{I}'$ in the above proof as $\mathbf{I} \circ f^{-1}$ and call it the image of $\mathbf{I}$ under f.

Exercise 1.8. *Fill in the details. What does the assertion look like if stated in terms of random elements rather than probability distributions ?*

Exercise 1.9. *Let* $\{\xi_n(i),\, n \in \mathbb{N}\}$*, where* $i = 1, 2, \ldots, k$*, be independent i.i.d. with finite exponential moments of some order. Derive the LDP for* $(X_n(1), \ldots, X_n(k))$ *where*

$$X_n(i) = \sum_{j=1}^{i} \frac{1}{n} \sum_{l=1}^{n} \xi_l(j).$$

Exercise 1.10. *(Koffman, Dembo-Zeitouni) Let*

$$X_n = \frac{1}{n}\Big(\frac{Y_1^2}{2} - \big(\frac{Y_2}{\sqrt{2}} + \sqrt{cn}\big)\Big),$$

where Y_1 *and* Y_2 *are independent standard normal. This is the signal generated by an optimal receiver for the detection of orthogonal signals in Gaussian white noise.* $c > 0$ *is the signal and the decision "signal is present" is taken if* $X_n \leq 0$*. Show that the* X_n *obey the LDP with* $\mathbf{I}(x) = (\sqrt{c} - \sqrt{-x})^2$ *if* $x \leq -c/4$ *and* $\mathbf{I}(x) = x + c/2$ *if* $x > -c/4$.

One can also use the LDP to obtain conditional laws of large numbers. We state it in the form of convergence of random variables. Let ρ denote the metric on $\mathbb{U}$. For $z \in \mathbb{U}$ and $\Gamma \subset \mathbb{U}$, $\rho(z, \Gamma)$ has the usual meaning: $\rho(z, \Gamma) = \inf_{z' \in \Gamma} \rho(z, z')$.

Theorem 1.2. *Let random variables* X_n *with values in* $\mathbb{U}$ *obey the LDP with* $\mathbf{I}$*. Let* H *be an* $\mathbf{I}$*-continuity subset of* $\mathbb{U}$ *with* $\mathbf{I}(H) < \infty$ *and let* F *denote the subset of* $cl\, H$ *where* $\inf_{z \in cl\, H} \mathbf{I}(z)$ *is attained. Then for arbitrary* $\epsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}\big(\rho(X_n, F) > \epsilon | X_n \in H\big) = 0.$$

Proof. By the definition of the LDP and continuity of $z \to \rho(z, F)$,

$$\lim_{n\to\infty} \mathbf{P}(X_n \in H)^{1/r_n} = \exp\big(-\mathbf{I}(H)\big),$$

$$\limsup_{n\to\infty} \mathbf{P}\big(\rho(X_n, F) > \epsilon,\ X_n \in H\big)^{1/r_n} \leq \exp\big(-\inf_{z \in \mathrm{cl}\, H:\, \rho(z,F) \geq \epsilon} \mathbf{I}(z)\big) < \exp\big(-\mathbf{I}(H)\big).$$

Therefore,

$$\limsup_{n\to\infty} \mathbf{P}\big(\rho(X_n,F)>\epsilon | X_n\in H\big)^{1/r_n} = \frac{\limsup_{n\to\infty} \mathbf{P}\big(\rho(X_n,F)>\epsilon,\ X_n\in H\big)^{1/r_n}}{\lim_{n\to\infty} \mathbf{P}(X_n\in H)^{1/r_n}} < 1.$$

The claim follows. □

The definition of the LDP is tailored to applications as it is the probability decay rate one is usually interested in. However, for theoretical developments the exponential form is more convenient as the above proof shows. The definition of the LDP can be written equivalently as

$$\limsup_{n\to\infty} \mathbf{P}_n(F)^{1/r_n} \le \sup_{z\in F} e^{-\mathbf{I}(z)} \text{ for all closed sets } F \tag{1.4}$$

and

$$\liminf_{n\to\infty} \mathbf{P}_n(G)^{1/r_n} \ge \sup_{z\in G} e^{-\mathbf{I}(z)} \text{ for all open sets } G. \tag{1.5}$$

Then the set function

$$\mathbf{\Pi}(\Gamma) = \exp(-\mathbf{I}(\Gamma)) \tag{1.6}$$

may be interpreted as a counterpart of a probability measure as it has the following properties.

1. $\mathbf{\Pi}(\emptyset)=0$, $\mathbf{\Pi}(\mathbb{U})=1$,
2. $\mathbf{\Pi}(\cup\Gamma_\alpha)=\sup_\alpha \mathbf{\Pi}(\Gamma_\alpha)$,
3. $\mathbf{\Pi}(F_n)\downarrow \mathbf{\Pi}(F)$ if $F_n\downarrow F$ and F_n are closed.

It is actually characterised by these properties, so we call such a function a deviability.

Exercise 1.11. *Check the claim.*

Besides, property 3 holds for decreasing directed families of closed sets too. In that form it is an analogue of the property of τ-smoothness of probability measures, cf., Vakhania, Tarieladze and Chobanyan [15].

We now want to obtain equivalent characterisations of the LDP. Let $\mathbb{C}_b^+(\mathbb{U})$, $\overline{\mathbb{C}}_b^+(\mathbb{U})$, and $\underline{\mathbb{C}}_b^+(\mathbb{U})$ denote the respective sets of $\mathbb{R}_+$-valued bounded continuous functions on $\mathbb{U}$, $\mathbb{R}_+$-valued bounded upper semi-continuous functions on $\mathbb{U}$, and $\mathbb{R}_+$-valued bounded lower semi-continuous functions on $\mathbb{U}$, respectively. For a function $f:\mathbb{U}\to\mathbb{R}_+$, let $\|f\|_n = \big(\int_{\mathbb{U}} f(z)^{r_n}\,\mathbf{P}_n(dz)\big)^{1/r_n}$. We call a set $H\subset\mathbb{U}$ $\mathbf{\Pi}$–continuous if $\mathbf{\Pi}(\operatorname{int} H)=\mathbf{\Pi}(\operatorname{cl} H)$, i.e., if it is $\mathbf{I}$-continuous.

Exercise 1.12. *Prove that if $\mathbf{\Pi}(\partial H)=0$, then H is $\mathbf{\Pi}$-continuous. Show that the converse is not true.*

Lemma 1.2. *If $f_n\downarrow f$ and the f_n are upper semicontinuous, then $\sup_{z\in\mathbb{U}} f_n(z)\mathbf{\Pi}(z)\downarrow \sup_{z\in\mathbb{U}} f(z)\mathbf{\Pi}(z)$.*

Theorem 1.3 (Portmanteau theorem)**.** *The following conditions are equivalent.*

1. *The $\mathbf{P}_n$ obey the LDP with $\mathbf{I}$*

1′. (i) $\limsup_{n\to\infty} \mathbf{P}_n(H)^{1/r_n} \le \mathbf{\Pi}(cl\ H)$

(ii) $\liminf_{n\to\infty} \mathbf{P}_n(H)^{1/r_n} \ge \mathbf{\Pi}(int\ H)$

for all Borel sets H

2. $\lim_{n\to\infty} \mathbf{P}_n(H)^{1/r_n} = \mathbf{\Pi}(H)$ *for all* $\mathbf{\Pi}$*–continuous Borel sets H*

3. $\lim_{n\to\infty} \|h\|_n = \sup_{z\in\mathbb{U}} h(z)\,\mathbf{\Pi}(z)$ *for all* $h \in \mathbb{C}_b^+(\mathbb{U})$

4. (i) $\limsup_{n\to\infty} \|f\|_n \le \sup_{z\in\mathbb{U}} f(z)\,\mathbf{\Pi}(z)$ *for all* $f \in \overline{\mathbb{C}}_b^+(\mathbb{U})$

(ii) $\liminf_{n\to\infty} \|g\|_n \ge \sup_{z\in\mathbb{U}} g(z)\,\mathbf{\Pi}(z)$ *for all* $g \in \underline{\mathbb{C}}_b^+(\mathbb{U})$

5. $\lim_{n\to\infty} \|h\|_n = \sup_{z\in\mathbb{U}} h(z)\,\mathbf{\Pi}(z)$ *for all bounded uniformly continuous functions* $h:\ \mathbb{U} \to \mathbb{R}_+$

Proof. Clearly, $1 \Leftrightarrow 1'$, $1' \Rightarrow 2$, $3 \Rightarrow 5$, $4 \Rightarrow 3$, $4 \Rightarrow 1$, and $4 \Rightarrow 5$.

We prove the implication $3 \Rightarrow 1$. To prove the lower bound, we note that, since G is open, $\mathbf{1}(G) = \sup h$ over $h \in \mathbb{C}_b^+(\mathbb{U})$ such that $h \le \mathbf{1}(G)$. Therefore, $\mathbf{\Pi}(G) = \sup_h \sup_{z\in\mathbb{U}} h(z)\,\mathbf{\Pi}(z)$, so that if $h_\epsilon \le \mathbf{1}(G)$ is such that $\mathbf{\Pi}(G) \le \sup_{z\in\mathbb{U}} h_\epsilon(z)\,\mathbf{\Pi}(z) + \epsilon$, then

$$\liminf_{n\to\infty} \mathbf{P}_n(G)^{1/r_n} \ge \lim_{n\to\infty} \|h_\epsilon\|_n = \sup_{z\in\mathbb{U}} h_\epsilon(z)\,\mathbf{\Pi}(z) \ge \mathbf{\Pi}(G) - \epsilon.$$

The proof of the upper bound is analogous if we note that $\mathbf{1}_F = \inf h$ over $h \in \mathbb{C}_b^+(\mathbb{U})$ such that $h \ge \mathbf{1}(F)$ so that $\mathbf{\Pi}(F) = \inf_h \sup_{z\in\mathbb{U}} h(z)\,\mathbf{\Pi}(z)$ (use Lemma 1.2).

We prove the implication $1 \Rightarrow 4$ by proving that (1.4) $\Rightarrow$ 4(i) and (1.5) $\Rightarrow$ 4(ii). For $f \in \overline{\mathbb{C}}_b^+(\mathbb{U})$ such that $\sup_{z\in\mathbb{U}} f(z) = 1$ let

$$f_k(z) = \max_{i=0,\dots,k-1} \Bigl[\frac{i+1}{k}\,\mathbf{1}\Bigl(f(z) \ge \frac{i}{k}\Bigr)\Bigr],\ k \in \mathbb{N}.$$

Since the sets $\{z:\ f(z) \ge x\}$ are closed by the upper semicontinuity of f, (1.4) yields

$$\begin{aligned}\limsup_{n\to\infty}\|f\|_n &\le \limsup_{n\to\infty}\|f_k\|_n \le \max_{i=0,\dots,k-1}\ \limsup_{n\to\infty}\Bigl[\frac{i+1}{k}\,\mathbf{P}_n\Bigl(f(z) \ge \frac{i+1}{k}\Bigr)^{1/r_n}\Bigr]\\ &\le \max_{i=0,\dots,k-1}\Bigl[\frac{i+1}{k}\,\mathbf{\Pi}\Bigl(f(z) \ge \frac{i+1}{k}\Bigr)\Bigr] = \sup_{z\in\mathbb{U}} f_k(z)\,\mathbf{\Pi}(z) \le \sup_{z\in\mathbb{U}} f(z)\,\mathbf{\Pi}(z) + \frac{1}{k}.\end{aligned}$$

The proof of (1.5) $\Rightarrow$ 4(ii) is similar if we consider $g_k(z) = \max_{i=0,\dots,k-1}\bigl[i/k\,\mathbf{1}(g(z) > i/k)\bigr]$.

Now we prove $2 \Rightarrow 1$. Let G be open and $\delta > 0$. Let h be a function from $\mathbb{C}_b^+(\mathbb{U})$ such that $h \le \mathbf{1}(G)$ and $\sup_{z\in\mathbb{U}} h(z)\,\mathbf{\Pi}(z) \ge \mathbf{\Pi}(G) - \delta$. Let $H_u = \{z \in \mathbb{U}:\ h(z) \ge u\}$, $u \in [0,1]$. Then the function $\mathbf{\Pi}(H_u)$ increases as $u \downarrow 0$. Therefore, it has at most countably many jumps. Also $\mathbf{\Pi}(H_u) \ge \sup_{z\in\mathbb{U}} h(z)\,\mathbf{\Pi}(z) - u$, so $\mathbf{\Pi}(H_u) \ge \mathbf{\Pi}(G) - 2\delta$ for u small enough. Thus, there exists $\epsilon > 0$ such that $\mathbf{\Pi}(H_\epsilon) \ge \mathbf{\Pi}(G) - 2\delta$ and $\mathbf{\Pi}(H_u)$ is continuous at ϵ. The latter is equivalent to H_ϵ being continuous relative to $\mathbf{\Pi}$, so we conclude that

$$\liminf_{n\to\infty} \mathbf{P}_n(G)^{1/r_n} \ge \lim_{n\to\infty} \mathbf{P}_n(H_\epsilon)^{1/r_n} = \mathbf{\Pi}(H_\epsilon) \ge \mathbf{\Pi}(G) - 2\delta.$$

The proof of the upper bound is similar.

We prove that $5 \Rightarrow 1$. Let us establish the upper bound (1.4). Let F be a closed subset of $\mathbb{U}$. Then $\mathbf{1}(F) = \inf_{\epsilon>0}(1-\rho(z,F)/\epsilon)^+$. The functions $(1-\rho(z,F)/\epsilon)^+$ are bounded and uniformly continuous so that

$$\limsup_{n\to\infty} \mathbf{P}_n(F)^{1/r_n} \le \inf_{\epsilon>0} \lim_{n\to\infty} \|\big(1-\rho(z,F)/\epsilon\big)^+\|_n = \inf_{\epsilon>0}\sup_{z\in\mathbb{U}}(1-\rho(z,F)/\epsilon)^+\,\mathbf{\Pi}(z) = \mathbf{\Pi}(F)$$

Lower bound (1.5) is proved in an analogous manner. □

Remark 1.2. As the proof shows, if convergences in part 5 hold for all $\mathbb{R}_+$-valued Lipshitz continuous functions, then the LDP holds.

Corollary 1.1 (Varadhan's lemma). *Let the $\mathbf{P}_n$ obey the LDP with $\mathbf{I}$. If a function $F : \mathbb{U} \to [-\infty,\infty)$ is continuous and bounded above, then*

$$\lim_{n\to\infty}\frac{1}{r_n}\log\int_{\mathbb{U}}\exp\big(r_nF(z)\big)\,\mathbf{P}_n(dz) = \sup_{z\in\mathbb{U}}\big(F(z)-\mathbf{I}(z)\big)$$

If F is an arbitrary continuous function, then the above convergence holds if and only if

$$\lim_{A\to\infty}\limsup_{n\to\infty}\frac{1}{r_n}\log\int_{\mathbb{U}}\exp\big(r_nF(z)\big)\mathbf{1}\big(F(z)>A\big)\,\mathbf{P}_n(dz) = -\infty.$$

Exercise 1.13. *Prove the corollary*

Exercise 1.14 (Dembo and Zeitouni). *Let $\epsilon\log p_\epsilon \to -\infty$ as $\epsilon \to 0$. Consider a family of r.v. Z_ϵ with $\mathbf{P}(Z_\epsilon = 0) = 1-2p_\epsilon$, $\mathbf{P}(Z_\epsilon = -\epsilon\log = \mathbf{P}(Z_\epsilon = \epsilon\log p_\epsilon) = p_\epsilon$. Show that the Z_ϵ obey the LDP with $\mathbf{I}(0)=0$ and $\mathbf{I}(z)=\infty$ for $z\ne 0$, so that $\Lambda^*(u)=0$. However, $\lim_{\epsilon\to}\epsilon\log\mathbf{E}\exp(\lambda Z_\epsilon/\epsilon)$ equals 0 if $|\lambda|\le 1$ and equals ∞ otherwise.*

Corollary 1.2. *Let $\mathbb{U}$ be separable. Let X_n and Y_n with values in $\mathbb{U}$ be such that $\mathbf{P}(\rho(X_n,Y_n)>\epsilon)^{1/r_n}\to 0$ as $n\to\infty$ for arbitrary $\epsilon>0$. Then the X_n obey the LDP with $\mathbf{I}$ if and only if the Y_n obey the LDP with $\mathbf{I}$.*

For a function $h : \mathbb{U} \to \mathbb{R}_+$, let $\overline{h}$ and $\underline{h}$ denote the respective upper semi-continuous and lower semi-continuous envelopes of h defined by

$$\overline{h} = \inf_{\substack{f\in\overline{\mathbb{C}}_b^+(\mathbb{U}):\\ f\ge h}} f \quad\text{and}\quad \underline{h} = \sup_{\substack{g\in\underline{\mathbb{C}}_b^+(\mathbb{U}):\\ g\le h}} g.$$

Exercise 1.15. *Check that $\overline{h}(z) = \limsup_{y\to z} h(y)$ and $\underline{h}(z) = \liminf_{y\to z} h(y)$.*

We say that h is $\mathbf{\Pi}$–continuous if $\sup_{z\in\mathbb{U}}\overline{h}(z)\,\mathbf{\Pi}(z) = \sup_{z\in\mathbb{U}}\underline{h}(z)\,\mathbf{\Pi}(z)$. We say that h is $\mathbf{\Pi}$ – upper-semi-continuous (respectively, $\mathbf{\Pi}$ – lower-semi-continuous) if $\sup_{z\in\mathbb{U}}\overline{h}(z)\,\mathbf{\Pi}(z) = \sup_{z\in\mathbb{U}} h(z)\,\mathbf{\Pi}(z)$ (respectively, $\sup_{z\in\mathbb{U}}\underline{h}(z)\,\mathbf{\Pi}(z) = \sup_{z\in\mathbb{U}} h(z)\,\mathbf{\Pi}(z)$). We also call a set $H\subset\mathbb{U}$ $\mathbf{\Pi}$ – closed (respectively, $\mathbf{\Pi}$ – open) if $\mathbf{\Pi}(H) = \mathbf{\Pi}(\mathrm{cl}\,H)$ (respectively, $\mathbf{\Pi}(H) = \mathbf{\Pi}(\mathrm{int}\,H)$).

Exercise 1.16. *A set is $\mathbf{\Pi}$ – continuous ($\mathbf{\Pi}$ – closed or $\mathbf{\Pi}$ – open, respectively) if and only if its indicator function is $\mathbf{\Pi}$ – continuous ($\mathbf{\Pi}$ – upper-semi-continuous or $\mathbf{\Pi}$ – lower-semi-continuous, respectively).*

The next theorem contains a number of other conditions equivalent to the LDP.

Theorem 1.4. *The $\mathbf{P}_n$ obey the LDP with $\mathbf{I}$ if and only if any of the following conditions hold.*

4′. *The inequalities of part 4 hold for all $\mathbf{\Pi}$ – upper-semi-continuous bounded measurable functions $f : \mathbb{U} \to \mathbb{R}_+$ and all $\mathbf{\Pi}$ – lower-semi-continuous bounded measurable functions $g : \mathbb{U} \to \mathbb{R}_+$, respectively*

3′. *The inequalities of part 3 hold for all $\mathbf{\Pi}$ – closed measurable sets F and $\mathbf{\Pi}$ – open measurable sets G, respectively.*

6. $\lim_{n\to\infty} \|h\|_n = \sup_{z\in\mathbb{U}} h(z)\,\mathbf{\Pi}(z)$ *for all $\mathbf{\Pi}$ – continuous measurable functions $h : \mathbb{U} \to \mathbb{R}_+$*

Theorem 1.5 (Contraction (continuous mapping) principle II)**.** *The assertion of Theorem 1.1 still holds if the function f is continuous at each z with $\mathbf{I}(z) < \infty$.*

Proof. Let $\mathbf{\Pi}' = \exp(-\mathbf{I}') = \mathbf{\Pi} \circ f^{-1}$. We prove $\mathbf{\Pi}'$ is a deviability. Let F_n be decreasing sequence of closed sets. For $a \in (0, 1]$, let $K_a = \{z \in \mathbb{U} : \mathbf{\Pi}(z) \geq a\}$ and $f_a : K_a \to \mathbb{U}'$. Then

$$\mathbf{\Pi}'(F_n) = \mathbf{\Pi} \circ f^{-1}(F_n) = \mathbf{\Pi}\big(f^{-1}(F_n)\big) \leq \mathbf{\Pi}\big(f^{-1}(F_n) \cap K_a\big) \vee a = \mathbf{\Pi}\big(f_a^{-1}(F_n)\big) \vee a$$

Since f_a is a continuous function, the sets $f_a^{-1}(F_n)$ are closed so that $\lim_{n\to\infty} \mathbf{\Pi}\big(f_a^{-1}(F_n)\big) = \mathbf{\Pi}\big(f_a^{-1}(\cap_n F_n)\big) \leq \mathbf{\Pi}\big(f^{-1}(\cap_n F_n)\big)$. Hence, $\lim_{n\to\infty} \mathbf{\Pi}'(F_n) \leq \mathbf{\Pi}'(\cap_n F_n) \vee a$. Since a is arbitrarily small, the required property follows.

Let $h' : \mathbb{U}' \to \mathbb{R}_+$ be a bounded continuous function. Then the function $h' \circ f$, being continuous $\mathbf{\Pi}$-a.e., is $\mathbf{\Pi}$-continuous, so

$$\lim_{n\to\infty} \Big(\int_{\mathbb{U}'} h'(z')^{r_n}\, \mathbf{P}_n \circ f^{-1}(dz')\Big)^{1/r_n} = \lim_{n\to\infty} \Big(\int_{\mathbb{U}} (h' \circ f)(z)^{r_n}\, \mathbf{P}_n(dz)\Big)^{1/r_n}$$
$$= \sup_{z\in\mathbb{U}} h' \circ f(z)\,\mathbf{\Pi}(z) = \sup_{z'\in\mathbb{U}'} h'(z')\,\mathbf{\Pi} \circ f^{-1}(z') = \sup_{z'\in\mathbb{U}'} h'(z')\,\mathbf{\Pi}'(z').$$

□

We now move on to developing compactness approaches. Let us say that the sequence $\mathbf{P}_n$ is exponentially tight (on order r_n) if for arbitrary $\epsilon > 0$ there exists compact K such that $\limsup_{n\to\infty} \mathbf{P}_n(\mathbb{U} \setminus K_n)^{1/r_n} < \epsilon$.

Exercise 1.17. *Check that weak LDP and exponential tightness imply full LDP.*

Let us say that the sequence $\mathbf{P}_n$ is LD relatively compact if every subsequence of $\mathbf{P}_n$ contains a further subsequence that obeys the LDP.

Theorem 1.6 (LD relative compactness criterion)**.** *1. An exponentially tight sequence of probability measures on a metric space is LD relatively compact.*

2. An LD relatively compact sequence on a Polish space is exponentially tight.

Proof. Sufficiency. Let compacts K_m be such that $\limsup_{n\to\infty} \mathbf{P}_n(K_m^c)^{1/r_n} < 1/m$. The metric spaces $\mathbb{C}_b^+(K_m)$ being separable, let $\{f_{mk},\, k \in \mathbb{N}\}$ be countable dense subsets for the uniform norm, which we denote $\|\cdot\|$. Let (n') be a subsequence such that $\lim_{n'\to\infty} \|f_{mk}\|_{n'}$ exist for all m and k. Given $f \in \mathbb{C}_b^+(\mathbb{U})$ and $\epsilon > 0$, let k and m be such that $\|f\|/m < \epsilon$ and $\|f\mathbf{1}_{K_m} - f_{mk}\| < \epsilon$. Since

$$\begin{aligned}
|\,\|f\|_n - \|f\|_l| &\leq |\,\|f\|_n - \|f\mathbf{1}_{K_m}\|_n| + |\,\|f\|_l - \|f\mathbf{1}_{K_m}\|_l| \\
&\quad + |\,\|f\mathbf{1}_{K_m}\|_n - \|f_{mk}\|_n| + |\,\|f\mathbf{1}_{K_m}\|_l - \|f_{mk}\|_l| + |\,\|f_{mk}\|_n - \|f_{mk}\|_l| \\
&\leq \|f\|\mathbf{P}_n(K_m^c)^{1/r_n} + \|f\|\mathbf{P}_l(K_m^c)^{1/r_l} + 2\|f\mathbf{1}_{K_m} - f_{mk}\| + |\,\|f_{mk}\|_n - \|f_{mk}\|_l|,
\end{aligned}$$

it follows that

$$\limsup_{n',l'\to\infty} \big|\,\|f\|_{n'} - \|f\|_{l'}\big| \le 4\epsilon,$$

so the sequence $\|f\|_{n'}$ is Cauchy, hence, it converges. We denote $\mathbf{S}f = \lim_{n'\to\infty}\|f\|_{n'}$.

Let $\mathbf{\Pi}(z) = \inf \mathbf{S}f$, the infimum being taken over f with $f(z) = 1$. The function $\mathbf{\Pi}$ is upper compact and $\mathbf{S}f = \sup_{z\in\mathbb{U}} f(z)\,\mathbf{\Pi}(z)$. For a proof of the latter, it is easy to see that $\mathbf{S}f \ge \sup_{z\in\mathbb{U}} f(z)\,\mathbf{\Pi}(z)$. For the converse, let us note that the above argument also shows existence of the limits $\mathbf{S}f\mathbf{1}_{K_m} = \lim_{n'\to\infty}\|f\mathbf{1}_{K_m}\|_{n'}$. Compactness considerations can be used to show that $\mathbf{S}f\mathbf{1}_{K_m} = \sup_{z\in K_m} f(z)\mathbf{\Pi}(z)$. Letting $m\to\infty$ completes the proof.

Necessity. Since $\mathbb{U}$ is separable, it is expressed as a countable union of open balls O_i of radius δ for arbitrary $\delta > 0$. Let us show that under the hypotheses there exist finitely many open balls O_i, $i = 1, 2, \ldots, k$ of radius δ such that

$$\sup_{n\in\mathbb{N}} \mathbf{P}_n\big(\mathbb{U}\setminus\bigcup_{i=1}^{k} O_i\big)^{1/r_n} < \epsilon.$$

Since each probability measure is tight by Ulam's theorem, the required property is equivalent to the inequality

$$\limsup_{n\to\infty} \mathbf{P}_n\big(\mathbb{U}\setminus\bigcup_{i=1}^{k} O_i\big)^{1/r_n} < \epsilon. \tag{1.7}$$

Choose subsequences n_l and m_l such that

$$\lim_{l\to\infty} \mathbf{P}_{n_l}\big(\mathbb{U}\setminus\bigcup_{i=1}^{m_l} O_i\big)^{1/r_{n_l}} = \limsup_{m\to\infty}\limsup_{n\to\infty} \mathbf{P}_n\big(\mathbb{U}\setminus\bigcup_{i=1}^{m} O_i\big)^{1/r_n}.$$

Since $\mathbf{P}_{n_l}$ contains a subsequence obeying the LDP with some $\mathbf{I}'$, we may assume, by replacing n_l, if necessary, with this subsequence, that the $\mathbf{P}_{n_l}$ obey the LDP for rate r_{n_l} with $\mathbf{I}'$. Therefore, introducing $\mathbf{\Pi}' = \exp(-\mathbf{I}')$, for arbitrary k

$$\lim_{l\to\infty} \mathbf{P}_{n_l}\big(\mathbb{U}\setminus\bigcup_{i=1}^{m_l} O_i\big)^{1/r_{n_l}} \le \lim_{l\to\infty} \mathbf{P}_{n_l}\big(\mathbb{U}\setminus\bigcup_{i=1}^{k} O_i\big)^{1/r_{n_l}} \le \mathbf{\Pi}'\big(\mathbb{U}\setminus\bigcup_{i=1}^{k} O_i\big).$$

By the facts that $\cup_{i=1}^{\infty} O_i = \mathbb{U}$ and the sets $\mathbb{U}\setminus\cup_{i=1}^{k} O_i$ are closed, we deduce that $\lim_{k\to\infty}\mathbf{\Pi}'\big(\mathbb{U}\setminus\cup_{i=1}^{k} O_i\big) = 0$. Hence,

$$\limsup_{m\to\infty}\limsup_{n\to\infty} \mathbf{P}_n\big(\mathbb{U}\setminus\bigcup_{i=1}^{m} O_i\big)^{1/r_n} = 0,$$

which implies (1.7).

Let $O_{l,1},\ldots,O_{l,k_l}$, for $l = 1, 2, \ldots$, be open balls of radius $1/l$ such that

$$\sup_{n\in\mathbb{N}} \mathbf{P}_n\big(\mathbb{U}\setminus\bigcup_{i=1}^{k_l} O_{l,i}\big)^{1/r_n} < \frac{\epsilon}{2^l}.$$

The set $B = \cap_{l=1}^{\infty}\cup_{i=1}^{k_l} O_{k,i}$ is totally bounded, hence, relatively compact as $\mathbb{U}$ is complete. Besides, for all n

$$\mathbf{P}_n\big(\mathbb{U}\setminus B\big)^{1/r_n} \le \sum_{l=1}^{\infty} \mathbf{P}_n\big(\mathbb{U}\setminus\bigcup_{i=1}^{k_l} O_{l,i}\big)^{1/r_n} < \epsilon. \tag{1.8}$$

□

We will say that $\mathbf{\Pi}$ is an LD accumulation point of $\mathbf{P}_n$ if there exists a subsequence that obeys the LDP with $\mathbf{I} = -\log \mathbf{\Pi}$. Let us consider some applications.

Theorem 1.7 (Gärtner et al.)**.** *Let X_n be $\mathbb{R}^k$–valued random variables such that for each $\lambda \in \mathbb{R}^k$*

$$\lim_{n\to\infty} \frac{1}{r_n} \log \mathbf{E}_n \exp\bigl(r_n \lambda \cdot X_n\bigr) = G(\lambda),$$

where $G(\lambda)$ is an $\overline{\mathbb{R}}$-valued lower semicontinuous and essentially smooth convex function such that $0 \in$ int (dom G). Then the X_n obey the LDP for rate r_n with $\mathbf{I}(x) = \sup_{\lambda\in\mathbb{R}^k}\bigl(\lambda \cdot x - G(\lambda)\bigr)$.

Proof. We first note that the sequence X_n is exponentially tight. To see this, we write by Chebyshev's inequality, for $A > 0$ and $\eta > 0$, denoting by e_i, $i = 1, \ldots, 2k$, the $2k$-vector, whose $\lfloor (k+1)/2 \rfloor$th entry is 1 if k is odd, -1 if k is even, and the rest of the entries are equal to 0,

$$\begin{aligned}\mathbf{P}_n(|X_n| > A)^{1/r_n} &\le \max_{i=1,\ldots,2k} P_n(e_i \cdot X^n > A/k)^{1/r_n} \\ &\le \exp(-\eta A/k) \max_{i=1,\ldots,2k} \bigl(\mathbf{E}_n \exp(r_n \eta e_i \cdot X_n)\bigr)^{1/r_n}.\end{aligned}$$

Exponential tightness follows since by hypotheses

$$\lim_n \bigl(\mathbf{E}_n \exp(r_n \eta e_i \cdot X_n)\bigr)^{1/r_n} = \exp(G(\eta e_i)),$$

where the right-hand side is finite if η is small enough by the fact that $G(\lambda)$ is finite in a neighbourhood of the origin.

Therefore, there exists a subsequence $\{X_{n'}\}$ that obeys the LDP with some $\mathbf{I}'$. Next, it follows from exponential Markov's inequality and Corollary 1.1 that if $\lambda \in \text{int}(\text{dom } G)$, then

$$\lim_{n'} \frac{1}{r_{n'}} \log\bigl(\mathbf{E}\exp(r_{n'}\lambda \cdot X_{n'})\bigr)^{1/r_{n'}} = \sup_{x\in\mathbb{R}^k} \left(\lambda \cdot x - \mathbf{I}'(x)\right), \ \lambda \in \text{int}(\text{dom } G).$$

Thus,

$$\sup_{x\in\mathbb{R}^k} \left(\lambda \cdot x - \mathbf{I}'(x)\right) = G(\lambda)$$

for all $\lambda \in \text{int}(\text{dom } G)$, which implies by convexity analysis arguments that $\mathbf{I} = \mathbf{I}'$. □

The following fact holds in greater generality. However, the form we give suffices for applications below and also nicely illustrates compactness methods.

Corollary 1.3. *Let independent sequences X_n and X'_n with values in Polish spaces $\mathbb{U}$ and $\mathbb{U}'$ obey the LDP with $\mathbf{I}$ and $\mathbf{I}'$, respectively. Then the sequence (X_n, X'_n) obeys the LDP in $\mathbb{U} \times \mathbb{U}'$ with $\hat{\mathbf{I}}(z, z') = \mathbf{I}(z) + \mathbf{I}'(z')$.*

Proof. Since $\mathbb{U}$ and $\mathbb{U}'$ are separable, the (X_n, X'_n) are well defined random variables. If K is a compact in $\mathbb{U}$ and K' is a compact in $\mathbb{U}'$, then $K \times K'$ is a compact in $\mathbb{U} \times \mathbb{U}'$. Since $\mathbf{P}_n\bigl((X_n, X'_n) \notin K \times K'\bigr)^{1/r_n} \le \mathbf{P}_n(X_n \notin K)^{1/r_n} + \mathbf{P}_n(X'_n \notin K')^{1/r_n}$ and each of the sequences X_n and X'_n is exponentially tight, the sequence (X_n, X'_n) is exponentially tight. Let $\mathbf{\Pi}$ be an LD accumulation point of the laws of (X_n, X'_n). Then, given continuous bounded $\mathbb{R}_+$-valued functions f and f' on $\mathbb{U}$ and $\mathbb{U}'$, respectively, we have that for a suitable subsequence n'

$$\lim_{n'\to\infty} \bigl(\mathbf{E}_n f(X_{n'})^{r_{n'}} f(X'_{n'})^{r_{n'}}\bigr)^{1/r_{n'}} = \sup_{(z,z')} f(z) f'(z') \mathbf{\Pi}(z, z').$$

On the other hand,

$$\lim_{n\to\infty}\bigl(\mathbf{E}_n f(X_n)^{r_n}\bigr)^{1/r_n} = \sup_z f(z)\exp(-\mathbf{I}(z)),$$
$$\lim_{n\to\infty}\bigl(\mathbf{E}_n f(X'_n)^{r_n}\bigr)^{1/r_n} = \sup_{z'} f(z')\exp(-\mathbf{I}'(z')).$$

Thus,

$$\sup_{(z,z')} f(z)f'(z')\mathbf{\Pi}(z,z') = \sup_z f(z)\exp(-\mathbf{I}(z))\sup_{z'} f(z')\exp(-\mathbf{I}'(z')).$$

Choosing $f(z) = (1-\rho(z,\hat{z})/\epsilon)^+$ and $f(z') = (1-\rho(z',\hat{z}')/\epsilon)^+$, where $\hat{z}$ and $\hat{z}'$ are some fixed elements of $\mathbb{U}$ and $\mathbb{U}'$, respectively, and ρ' is the metric on $\mathbb{U}'$, we obtain on letting $\epsilon \to 0$ by Lemma 1.2 that

$$\sup_{(z,z')} \mathbf{1}(z=\hat{z})\mathbf{1}(z'=\hat{z}')\mathbf{\Pi}(z,z') = \sup_z \mathbf{1}(z=\hat{z})\exp(-\mathbf{I}(z))\sup_{z'}\mathbf{1}(z'=\hat{z}')\exp(-\mathbf{I}'(z')),$$

i.e., $\mathbf{\Pi}(\hat{z},\hat{z}') = \exp(-\mathbf{I}(\hat{z}))\exp(-\mathbf{I}'(\hat{z}'))$. □

Corollary 1.4. *If the sequence $\{\mathbf{P}_n,\, n\in\mathbb{N}\}$ is exponentially tight and*

$$\limsup_{n\to\infty}\limsup_{\epsilon\to 0}\frac{1}{r_n}\log\mathbf{P}_n(B_\epsilon(z)) = \liminf_{n\to\infty}\liminf_{\epsilon\to 0}\frac{1}{r_n}\log\mathbf{P}_n(B_\epsilon(z)),$$

then the LDP holds. The rate function is equal to the common value of the above limits taken with a minus sign.

Note that we do not need to check the lower compactness property in the definition of the rate function.

Exercise 1.18. *Prove the corollary.*

Exercise 1.19. *Prove that if a lower compact function $\mathbf{I}$ is such that any subsequence $\mathbf{P}_{n'}$ of $\mathbf{P}_n$ contains a further subsequence $\mathbf{P}_{n''}$ that obeys the LDP with $\mathbf{I}$ for rate $r_{n''}$, then $\mathbf{P}_n$ obeys the LDP with $\mathbf{I}$ for rate r_n.*

Let $\mathbb{U}_0 \subset \mathbb{U}$. We will say that the sequence $\mathbf{P}_n$ is $\mathbb{U}_0$-exponentially tight if it is exponentially tight and any LD accumulation point $\mathbf{\Pi}$ is such that $\mathbf{\Pi}(\mathbb{U}\setminus\mathbb{U}_0) = 0$. Similarly we say that a function $f:\mathbb{U}\to\mathbb{U}'$ is $\mathbb{U}_0$-continuous if f is continuous at each $z\in\mathbb{U}_0$. We recall that sets Γ_α are said to make up a decreasing directed family if for arbitrary α' and α'' there exists α''' such that $\Gamma_{\alpha'}\cap\Gamma_{\alpha''}\supset\Gamma_{\alpha'''}$. The next theorem, which can be loosely referred to as the method of finite-dimensional distributions, is crucial for the developments below.

Theorem 1.8. *Let f_α be functions from $\mathbb{U}$ into metric spaces $\mathbb{U}_\alpha$ such that the sets $f_\alpha^{-1}\bigl(f_\alpha(z)\bigr)$ make up a decreasing directed family and $\{z\} = \cap_\alpha f_\alpha^{-1}\bigl(f_\alpha(z)\bigr)$ for $z\in\mathbb{U}_0$. If the sequence $\mathbf{P}_n$ is $\mathbb{U}_0$-exponentially tight, the functions f_α are $\mathbb{U}_0$-continuous, and the $\mathbf{P}_n\circ f_\alpha^{-1}$ obey the LDP with $\mathbf{I}_\alpha$, then the $\mathbf{P}_n$ obey the LDP with $\mathbf{I}(z) = \sup_\alpha \mathbf{I}_\alpha(f_\alpha(z))$ if $z\in\mathbb{U}_0$ and $\mathbf{I}(z)=\infty$ otherwise.*

Proof. Suppose that a subsequence $\mathbf{P}_{n'}$ obeys the LDP for rate $r_{n'}$ with $\mathbf{I}'$. It suffices to prove that $\mathbf{I}'=\mathbf{I}$. By the $\mathbb{U}_0$-exponential tightness, $\mathbf{I}'(z)=\infty$ if $z\notin\mathbb{U}_0$. By the continuous mapping principle (II), the $\mathbf{P}_{n'}\circ f_\alpha^{-1}$ obey the LDP for rate $r_{n'}$ with $\mathbf{I}'\circ f_\alpha^{-1}$. Hence, $\mathbf{I}'\circ f_\alpha^{-1} = \mathbf{I}_\alpha$. Let $z\in\mathbb{U}_0$. Since the f_α are $\mathbb{U}_0$-continuous, these maps are continuous when restricted to $\mathbb{U}_0$ with subspace topology. Therefore, the sets $f_\alpha^{-1}\bigl(f_\alpha(z)\bigr)\cap\mathbb{U}_0$ are closed in $\mathbb{U}_0$, so they can be written as $F_\alpha \cap \mathbb{U}_0$, where the F_α are closed in $\mathbb{U}$ and form a decreasing directed family (check). Since $\{z\} = \cap_\alpha\big(f_\alpha^{-1}\big(f_\alpha(z)\big) \cap \mathbb{U}_0\big)$, it follows that

$$\begin{aligned}\mathbf{I}'(z) &= \mathbf{I}'\big(\cap_\alpha\big(f_\alpha^{-1}\big(f_\alpha(z)\big) \cap \mathbb{U}_0\big)\big) = \mathbf{I}'\big(\cap_\alpha\big(F_\alpha \cap \mathbb{U}_0\big)\big) = \sup_\alpha \mathbf{I}'\big(F_\alpha \cap \mathbb{U}_0\big) \\ &= \sup_\alpha \mathbf{I}'\big(f_\alpha^{-1}\big(f_\alpha(z)\big) \cap \mathbb{U}_0\big) = \sup_\alpha \mathbf{I}'\big(f_\alpha^{-1}\big(f_\alpha(z)\big)\big) = \sup_\alpha \mathbf{I}' \circ f_\alpha^{-1}\big(f_\alpha(z)\big) = \sup_\alpha \mathbf{I}_\alpha(f_\alpha(z)).\end{aligned}$$

□

Remark 1.3. For the metric space case, the above theorem encompasses both the inverse contraction principle and the projection limit theorem, see, e.g., Dembo and Zeitouni [3].

2 LDP for stochastic processes

2.1 Empirical measure

In this part we will consider results in the theme of Sanov's theorem. Let consider i.i.d. $\mathbb{R}^d$-valued r.v. $\xi_1, \xi_2, \ldots$. We introduce the empirical measure

$$\nu_n = \frac{1}{n}\sum_{i=1}^n \delta_{\xi_i}, \tag{2.1}$$

where δ_x is the Dirac measure at x. Equivalently, for $\Gamma \subset \mathbb{R}^d$,

$$\nu_n(\Gamma) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(\xi_i \in \Gamma).$$

ν_n is considered as a random element of space $\mathcal{M}(\mathbb{R}^d)$ of probability measures on $\mathbb{R}^d$. The latter space is endowed with the topology of weak convergence, so it is a metric space (a Polish space, in fact). Recall that by Prohorov's theorem a subset of $\mathcal{M}$ is relatively compact if and only if it is tight. We develop an exponential tightness criterion.

Theorem 2.1. *Let μ_n be a sequence of random elements of $\mathcal{M}(\mathbb{R}^d)$ defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. This sequence is exponentially tight on order r_n if and only if for arbitrary $\epsilon > 0$ and $\delta > 0$ there exists $M > 0$ such that*

$$\limsup_{n\to\infty} \mathbf{P}\big(\mu_n(x \in \mathbb{R}^d : |x| > M) > \delta\big)^{1/r_n} < \epsilon.$$

Proof. Sufficiency. Since $\mathcal{M}(\mathbb{R}^d)$ is a Polish space, the distributions of the μ_n are tight, so we may assume that M is such that

$$\sup_n \mathbf{P}\big(\mu_n(x \in \mathbb{R}^d : |x| > M) > \delta\big)^{1/r_n} < \epsilon.$$

Given arbitrary $\epsilon > 0$, let M_k, where $k = 1, 2, \ldots$, be such that

$$\sup_n \mathbf{P}\big(\mu_n(x \in \mathbb{R}^d : |x| > M_k) > \frac{1}{k}\big)^{1/r_n} < \frac{\epsilon}{2^k}.$$

Introduce the set $K = \cap_{k=1}^\infty\{\mu : \mu(|x| > M_k) \le 1/k\}$. This set is weakly relatively compact by Prohorov's theorem and

$$\mathbf{P}(\mu_n \notin K)^{1/r_n} \le \sum_{k=1}^\infty \mathbf{P}\big(\mu_n(|x| > M_k) > \frac{1}{k}\big)^{1/r_n} \le \epsilon,$$

so the laws of the μ_n are exponentially tight.

Necessity. Suppose the laws of the μ_n are exponentially tight, i.e., for arbitrary $\epsilon > 0$ there exists a compact $K \subset \mathcal{M}(\mathbb{R}^d)$ with $\limsup_{n\to\infty} \mathbf{P}(\mu_n \notin K)^{1/r_n} < \epsilon$. Since by Prohorov's theorem for arbitrary δ there exists M such that $\sup_{\mu\in K} \mu\big(x : |x| > M\big) \le \delta$, for this M

$$\limsup_{n\to\infty} \mathbf{P}\big(\mu_n(x : |x| > M) > \delta\big)^{1/r_n} \le \limsup_{n\to\infty} \mathbf{P}(\mu_n \notin K)^{1/r_n} < \epsilon.$$

□

We are now ready to state and prove Sanov's theorem. For $\mu \in \mathcal{M}(\mathbb{R}^d)$ and $\nu \in \mathcal{M}(\mathbb{R}^d)$, let the relative entropy $H(\nu||\mu)$ be defined by

$$H(\nu||\mu) = \int_{\mathbb{R}^d} \frac{d\nu}{d\mu}(x) \log \frac{d\nu}{d\mu}(x)\, \mu(dx)$$

if ν is absolutely continuous with respect to μ, and $H(\nu||\mu) = \infty$ otherwise.

Exercise 2.1. *1. Check that $H(\nu||\mu)$ is nonnegative and lower compact in ν.*

2. Show that

$$H(\nu||\mu) = \sup_{\lambda(\cdot)} \Big(\int_{\mathbb{R}^d} \lambda(x)\, \nu(dx) - \log \int_{\mathbb{R}^d} e^{\lambda(x)}\, \mu(dx) \Big),$$

where the supremum is taken over all Borel functions $(\lambda(x),\, x \in \mathbb{R}^d)$ such that the integrals exist.

Theorem 2.2. *Let ξ_i be i.i.d. with law μ. Then the ν_n defined by (2.1) obey the LDP in $\mathcal{M}(\mathbb{R}^d)$ with $H(\cdot||\mu)$.*

Proof. Check exponential tightness first. For $\lambda > 0$, $\delta > 0$ and $M > 0$, by the exponential Markov inequality

$$\mathbf{P}\big(\mu_n(|x| > M) > \delta\big) = \mathbf{P}\Big(\frac{1}{n}\sum_{i=1}^n \mathbf{1}(|\xi_i| > M) > \delta\Big) \le \exp(-\lambda n \delta)\big(\mathbf{E}\exp(\lambda \mathbf{1}(|\xi_i| > M))\big)^n$$

Hence,

$$\limsup_{n\to\infty} \mathbf{P}\big(\mu_n(|x| > M) > \delta\big)^{1/n} \le \exp(-\lambda\delta)\mathbf{E}\exp(\lambda \mathbf{1}(|\xi_i| > M)),$$

so by dominated convergence

$$\lim_{M\to\infty} \limsup_{n\to\infty} \mathbf{P}\big(\mu_n(|x| > M) > \delta\big)^{1/n} \le \exp(-\lambda\delta),$$

which implies that the latter left-hand side is zero by λ being arbitrary. Exponential tightness follows by Theorem 2.1.

An application of Cramér's theorem shows that if $\Gamma_1, \dots, \Gamma_k$ are disjoint subsets of $\mathbb{R}^d$, then the sequence $(\mu_n(\Gamma_1), \dots, \mu_n(\Gamma_k))$ obeys the LDP with $\mathbf{I}_{\Gamma_1,\dots,\Gamma_k}$ given by

$$\begin{aligned}\mathbf{I}_{\Gamma_1,\dots,\Gamma_k}(x_1,\dots,x_k) &= \sup_{\lambda_1,\dots,\lambda_k} \Big(\sum_{i=1}^k \lambda_i x_i - \log \mathbf{E}\exp\Big(\sum_{i=1}^k \lambda_i \mathbf{1}(\xi_1 \in \Gamma_i)\Big) \\ &= \sup_{\lambda_1,\dots,\lambda_k} \Big(\sum_{i=1}^k \lambda_i x_i - \log \sum_{i=1}^k e^{\lambda_i} \mathbf{P}(\xi_1 \in \Gamma_i)\Big),\end{aligned}$$

where $x_i \in [0,1]$ and $\sum_{i=1}^k x_i \le 1$. Calculations show that

$$\mathbf{I}_{\Gamma_1,\dots,\Gamma_k}(x_1,\dots,x_k) = \sum_{i=1}^k \frac{x_i}{\mathbf{P}(\xi_1 \in \Gamma_i)} \log \frac{x_i}{\mathbf{P}(\xi_1 \in \Gamma_i)}\, \mathbf{P}(\xi_1 \in \Gamma_i)$$
$$+ \frac{1-\sum_{i=1}^k x_i}{\mathbf{P}(\xi_1 \notin \cup_{i=1}^k \Gamma_i)} \log \frac{x_i}{\mathbf{P}\big(\xi_1 \notin \cup_{i=1}^k \Gamma_i\big)}\, \mathbf{P}\Big(\sum_{i=1}^k \xi_1 \notin \cup_{i=1}^k \Gamma_i\Big),$$

where $0/0 = 0$ and $0 \log 0 = 0$. Since

$$\sup_{\Gamma_1,\dots,\Gamma_k} \mathbf{I}_{\Gamma_1,\dots,\Gamma_k}(\nu(\Gamma_1),\dots,\nu(\Gamma_k)) = H(\nu||\mu),$$

the proof is over by Theorem 1.8. □

Exercise 2.2. *Fill in the details.*

Exercise 2.3. *Consider a queueing network of K nodes in which routing decisions at the nodes are independent and identically distributed for a given node, i.e., a customer that has completed service in node i is routed to node j with probability p_{ij} and leaves the network with probability $1-\sum_{j=1}^K p_{ij}$. Let $R_{ij}(n)$ denote the number of customers routed from node i to node j out of the first n customers served at node i. Derive the LDP for the process $R(n)/n$, where $R(n) = (R_{ij}(n),\, i,j = 1,2,\dots,K)$.*

2.2 Space $\mathbb{C}$

We start by recalling basic properties of the space of continuous functions, for more detail see Billingsley [1], Jacod and Shiryaev [11], Ethier and Kurtz [8]. Space $\mathbb{C}(\mathbb{R}^d)$ is the space of all continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^d$. The metric is given by

$$\rho_{\mathbb{C}}(\mathbf{x},\mathbf{y}) = \sum_{k=1}^\infty 2^{-k} \sup_{t\in[0,k]} |\mathbf{x}(t)-\mathbf{y}(t)| \wedge 1.$$

Exercise 2.4. *Check that $\rho_{\mathbb{C}}$ is a metric and that $\mathbf{x}_n \to \mathbf{x}$ in $\mathbb{C}$ if and only if $\lim_{n\to\infty} \sup_{t\le T} |\mathbf{x}_n(t) - \mathbf{x}(t)| = 0$ for all $T > 0$.*

Remark 2.1. Equivalent metrics are given by the equalities

$$\rho'_{\mathbb{C}}(\mathbf{x},\mathbf{y}) = \sum_{k=1}^\infty 2^{-k} \frac{\sup_{t\in[0,k]}|\mathbf{x}(t)-\mathbf{y}(t)|}{1+\sup_{t\in[0,k]}|\mathbf{x}(t)-\mathbf{y}(t)|},$$
$$\rho''_{\mathbb{C}}(\mathbf{x},\mathbf{y}) = \sup_{t\in\mathbb{R}_+} \frac{|\mathbf{x}(t)-\mathbf{y}(t)|\wedge 1}{1+t},$$
$$\rho'''_{\mathbb{C}}(\mathbf{x},\mathbf{y}) = \int_0^\infty e^{-u} \sup_{t\in[0,u]} |\mathbf{x}(t)-\mathbf{y}(t)| \wedge 1\, du.$$

$\mathbb{C}(\mathbb{R}^d)$ is complete and separable (Polish). Modulus of continuity

$$w_T(\mathbf{x},\delta) = \sup_{\substack{s,t\in[0,T]:\\ |s-t|\le\delta}} |\mathbf{x}(t)-\mathbf{x}(s)|.$$

Exercise 2.5. *Prove that if* $\mathbf{x} \in \mathbb{C}(\mathbb{R}^d)$*, then* $w_T(\mathbf{x}, \delta) \to 0$ *as* $\delta \to 0$.

Theorem 2.3. *Set* $K \subset \mathbb{C}(\mathbb{R}^d)$ *is relatively compact if and only if the following hold:*

(i) $\sup_{\mathbf{x}\in K} |\mathbf{x}(0)| < \infty$,

(ii) $\lim_{\delta\to 0} \sup_{\mathbf{x}\in K} w_T(\mathbf{x}, \delta) = 0$ *for all* $T > 0$.

Theorem 2.4. *A sequence* $\mathbf{P}_n$ *of distributions on* $\mathbb{C}(\mathbb{R}^d)$ *is exponentially tight if and only if the following conditions hold*

(i) $\lim_{A\to\infty} \mathbf{P}_n(\mathbf{x} :\, |\mathbf{x}(0)| > A)^{1/r_n} = 0$,

(ii) $\lim_{\delta\to 0} \mathbf{P}_n\big(\mathbf{x} : w_T(\mathbf{x}, \delta) > \epsilon\big)^{1/r_n} = 0$ *for all* $T > 0$ *and* $\epsilon > 0$.

Exercise 2.6. *Deduce Theorem 2.4 from Theorem 2.3.*

Theorem 2.5. *A sequence* $\mathbf{P}_n$ *of distributions on* $\mathbb{C}(\mathbb{R}^d)$ *is exponentially tight if and only if the following conditions hold*

(i) $\lim_{A\to\infty} \mathbf{P}_n(\mathbf{x} : |\mathbf{x}(0)| > A)^{1/r_n} = 0$,

(ii) $\lim_{\delta\to 0} \sup_{t\in[0,T]} \mathbf{P}_n\big(\mathbf{x} :\, \sup_{s\in[t,t+\delta]} |\mathbf{x}(s) - \mathbf{x}(t)| > \epsilon\big)^{1/r_n} = 0$ *for all* $T > 0$ *and* $\epsilon > 0$.

Proof. It suffices to check that condition (ii) implies condition (ii) of Theorem 2.4. We have that

$$\mathbf{P}_n\big(\mathbf{x} :\, w_T(\mathbf{x}, \delta) > \epsilon\big) \le \mathbf{P}_n\Big(\bigcup_{i=1}^{\lfloor T/\delta\rfloor+1} \{\mathbf{x} :\, 3 \sup_{s\in[i\delta,(i+1)\delta]} |\mathbf{x}(s) - \mathbf{x}(i\delta)| > \epsilon\}\Big)$$
$$\le \sum_{i=1}^{\lfloor T/\delta\rfloor+1} \mathbf{P}_n\big(\mathbf{x} :\, 3 \sup_{s\in[i\delta,(i+1)\delta]} |\mathbf{x}(s)-\mathbf{x}(i\delta)| > \epsilon\big) \le \big(\frac{T}{\delta}+1\big) \sup_{t\in[0,T]} \mathbf{P}_n\big(\mathbf{x} :\, \sup_{s\in[t,t+\delta]} |\mathbf{x}(s)-\mathbf{x}(t)| > \epsilon/3\big).$$

The claim follows ∎

Exercise 2.7. *Finish the proof.*

By Theorem 1.8 we obtain the following result.

Theorem 2.6. *Let* $X_n = (X_n(t),\, t \in \mathbb{R}_+)$ *be a sequence of stochastic processes with trajectories in* $\mathbb{C}(\mathbb{R}^d)$*. Suppose the following conditions hold:*

1. *finite-dimensional projections* $(X_n(t_1), \dots, X_n(t_k))$ *obey LDPs in* $(\mathbb{R}^d)^k$ *with* $\mathbf{I}_{t_1,\dots,t_k}$*,*

2. *the sequence* X_n *is exponentially tight.*

Then the sequence X_n *obeys the LDP in* $\mathbb{C}(\mathbb{R}^d)$ *with* $\mathbf{I}(\mathbf{x}) = \sup_{t_1,\dots,t_k} \mathbf{I}_{t_1,\dots,t_k}(\mathbf{x}(t_1), \dots, \mathbf{x}(t_k))$.

Exercise 2.8.
1. *Fill in the details.*

2. *Prove that the function* $\mathbf{I}$ *in the statement is lower compact.*

3. *Deduce the statement of Theorem 2.6.*

Theorem 2.7 (Shilder)**.** *Let $B = (B(t),\, t \in \mathbb{R}_+)$ be a Brownian motion in $\mathbb{R}^d$. Then the sequence $\sqrt{\epsilon}B$ obeys the LDP for rate ϵ with*

$$\mathbf{I}(\mathbf{x}) = \frac{1}{2}\int_0^\infty |\dot{\mathbf{x}}(t)|^2\, dt$$

if $\mathbf{x}$ is absolutely continuous and $\mathbf{x}(0) = 0$, and $\mathbf{I}(\mathbf{x}) = \infty$, otherwise.

Let us precede the proof with a simple but useful bound: if $\xi = (\xi^{(1)}, \ldots, \xi^{(d)})$ is a d-dimensional random variable on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, then for $c > 0$

$$\mathbf{P}(|\xi| > c) \le 2d \max_{i=1,2\ldots,2d} \mathbf{P}\big(e_i \cdot \xi > \frac{c}{\sqrt{d}}\big), \tag{2.2}$$

where e_i for $i = 1, 2, \ldots, d$ is the i-th unit coordinate vector and $e_i = -e_{i-d}$ for $i = d+1, \ldots, 2d$. This bound follows by the inequalities

$$\mathbf{P}(|\xi| > c) \le \mathbf{P}\big(\sqrt{d} \max_{i=1,2,\ldots,d} |\xi^{(i)}| > c\big) \le \sum_{i=1}^d \mathbf{P}\big(|\xi^{(i)}| > \frac{c}{\sqrt{d}}\big) \le 2d \max_{i=1,2\ldots,2d} \mathbf{P}\big(e_i \cdot \xi > \frac{c}{\sqrt{d}}\big).$$

Proof of Theorem 2.7. We check exponential tightness. Using Doob's maximal inequality in view of (2.2) for $\mu > 0$

$$\begin{aligned}
\mathbf{P}_\epsilon\big(\sup_{s\in[t,t+\delta]} |\sqrt{\epsilon}B(s) - \sqrt{\epsilon}B(t)| > \eta\big) &\le 2d \max_{i=1,\ldots,2d} \mathbf{P}_\epsilon\big(\sqrt{\epsilon} \sup_{s\in[t,t+\delta]} e_i \cdot (B(s) - B(t)) > \frac{\eta}{\sqrt{d}}\big) \\
&= 2d \max_{i=1,\ldots,2d} \mathbf{P}_\epsilon\big(\sup_{s\in[t,t+\delta]} \exp\big(\frac{\mu}{\sqrt{\epsilon}}\, e_i \cdot (B(s) - B(t))\big) > \exp\big(\frac{\mu\eta}{\epsilon\sqrt{d}}\big)\big) \\
&\le \exp\big(-\frac{\mu\eta}{\epsilon\sqrt{d}}\big)\, 2d \max_{i=1,\ldots,2d} \mathbf{E}_\epsilon \exp\big(\frac{\mu}{\sqrt{\epsilon}}\, e_i \cdot (B(t+\delta) - B(t))\big) \\
&= 2d \exp\big(-\frac{\mu\eta}{\epsilon\sqrt{d}}\big)\, \exp\big(\frac{\mu^2\delta}{2\epsilon}\big).
\end{aligned} \tag{2.3}$$

Hence,

$$\limsup_{\delta\to 0} \limsup_{\epsilon\to 0} \mathbf{P}_\epsilon\big(\sup_{s\in[t,t+\delta]} |\sqrt{\epsilon}B_s - \sqrt{\epsilon}B_t| > \eta\big)^\epsilon \le \exp\big(-\frac{\mu\eta}{\sqrt{d}}\big)\,.$$

Since μ is arbitrary, condition (ii) of Theorem 2.5 is proved. Condition (i) is obvious.

We now evaluate the rate function. We have that

$$\mathbf{I}_{t_1,\ldots,t_k}(x_1, \ldots, x_k) = \frac{1}{2}\sum_{i=1}^k \frac{|x_i - x_{i-1}|^2}{t_i - t_{i-1}}.$$

Let $\mathbf{x}$ be absolutely continuous and $\mathbf{x}(0) = 0$. By Cauchy-Schwarz

$$\begin{aligned}
\mathbf{I}_{t_1,\ldots,t_k}(\mathbf{x}(t_1), \ldots, \mathbf{x}(t_k)) = \frac{1}{2}\sum_{i=1}^k \frac{|\mathbf{x}(t_i) - \mathbf{x}(t_{i-1})|^2}{t_i - t_{i-1}} &= \frac{1}{2}\sum_{i=1}^k \frac{1}{t_i - t_{i-1}} \Big|\int_{t_{i-1}}^{t_i} \dot{\mathbf{x}}(t)\, dt\Big|^2 \\
&\le \frac{1}{2}\sum_{i=1}^k \int_{t_{i-1}}^{t_i} |\dot{\mathbf{x}}(t)|^2\, dt \le \frac{1}{2}\int_0^\infty |\dot{\mathbf{x}}(t)|^2\, dt
\end{aligned}$$

On the other hand,

$$\mathbf{I}_{t_1,\dots,t_k}(\mathbf{x}(t_1),\dots,\mathbf{x}(t_k)) = \frac{1}{2}\sum_{i=1}^{k}\Big(\frac{|\mathbf{x}(t_i)-\mathbf{x}(t_{i-1})|}{t_i-t_{i-1}}\Big)^2(t_i-t_{i-1}),$$

so by Fatou, as $\sup_i(t_i-t_{i-1})\to 0$,

$$\liminf \mathbf{I}_{t_1,\dots,t_k}(\mathbf{x}(t_1),\dots,\mathbf{x}(t_k)) \ge \frac{1}{2}\int_0^\infty |\dot{\mathbf{x}}(t)|^2\,dt\,.$$

Thus,

$$\sup_{t_1,\dots,t_k}\mathbf{I}_{t_1,\dots,t_k}(\mathbf{x}(t_1),\dots,\mathbf{x}(t_k)) = \frac{1}{2}\int_0^\infty |\dot{\mathbf{x}}(t)|^2\,dt\,.$$

Suppose, $\mathbf{x}$ is not absolutely continuous on an interval $[0,T]$. It means that there exists $\epsilon>0$ such that for any $\delta>0$ there exist nonoverlapping intervals $(s_i,u_i)\in[0,T]$ with $\sum(u_i-s_i)<\delta$ and $\sum|\mathbf{x}(u_i)-\mathbf{x}(s_i)|>\epsilon$. By Cauchy-Schwarz,

$$\sup_{t_1,\dots,t_k}\mathbf{I}_{t_1,\dots,t_k}(\mathbf{x}(t_1),\dots,\mathbf{x}(t_k)) \ge \sum_{i=1}^{k}\Big(\frac{|\mathbf{x}(u_i)-\mathbf{x}(s_i)|}{u_i-s_i}\Big)^2(u_i-s_i) \ge \frac{\big(\sum_{i=1}^k|\mathbf{x}(u_i)-\mathbf{x}(s_i)|\big)^2}{\sum(u_i-s_i)} \ge \frac{\epsilon^2}{\delta}.$$

Since δ is arbitrarily small, we obtain the required conclusion. The case $\mathbf{x}(0)\neq 0$ is considered similarly. ☐

Exercise 2.9. *1. Prove that* $\mathbf{I}$ *in the hypotheses is lower compact.*

2. What about the LDP in $\mathbb{C}[0,1]$ *alluded to in the introductory example ?*

Exercise 2.10. *Find a function* $\mathbf{x}$ *such that, given* $\ell>0$, $\lim_{\epsilon\to0}\mathbf{P}(\sup_{t\in[0,1]}|\sqrt{\epsilon}B(t)-\mathbf{x}(t)|>\delta|\sup_{t\in[0,1]}|\sqrt{\epsilon}B(t)|>\ell)=0$ *for any* $\delta>0$.

Exercise 2.11. *Let* B^0 *be the Brownian bridge. Establish an LDP for* ϵB^0*. (Hint. Use the representation* $B^0(t)=B(t)-tB(1)$*, where* B *is a Brownian motion.)*

Theorem 2.8. *Let* X_ϵ *solve the equation*

$$X_\epsilon(t) = x + \int_0^t b(X_\epsilon(s))\,ds + \sqrt{\epsilon}B(t),$$

where $b:\mathbb{R}^d\to\mathbb{R}^d$ *is a Lipshitz-continuous function. Then the* X_ϵ *obey the LDP for rate* ϵ *with*

$$\mathbf{I}(\mathbf{x}) = \frac{1}{2}\int_0^\infty |\dot{\mathbf{x}}(s)-b(\mathbf{x}(s))|^2\,ds$$

if $\mathbf{x}$ *is absolutely continuous and* $\mathbf{x}(0)=x$*, and* $\mathbf{I}(\mathbf{x})=\infty$ *otherwise.*

Proof. Let us consider the mapping from $\mathbb{C}(\mathbb{R}^d)$ to $\mathbb{C}(\mathbb{R}^d)$ that associates with $\mathbf{x}$ the solution $\mathbf{y}$ of the equation

$$\mathbf{y}(t) = x + \mathbf{x}(t) + \int_0^t b(\mathbf{y}(s))\,ds.$$

Since b is Lipshitz, this equation has a unique solution. Moreover, if $\mathbf{y}'$ and $\mathbf{y}''$ are solutions associated with $\mathbf{x}'$ and $\mathbf{x}''$, respectively, then

$$|\mathbf{y}'(t) - \mathbf{y}''(t)| \le |\mathbf{x}'(t) - \mathbf{x}''(t)| + L\int_0^t |\mathbf{y}'(s) - \mathbf{y}''(s)|\,ds,$$

where L is a Lipshitz constant. Gronwall's inequality yields the bound

$$\sup_{t\in[0,T]} |\mathbf{y}'(t) - \mathbf{y}''(t)| \le \sup_{t\in[0,T]} |\mathbf{x}'(t) - \mathbf{x}''(t)| e^{LT}.$$

Hence, the mapping is continuous. An application of the continuous mapping principle completes the proof. □

Exercise 2.12. *Do the assignment in exercise 2.7 for the solution of*

$$X_\epsilon(t) = -\int_0^t X_\epsilon(s)\,ds + \sqrt{\epsilon}\,B(t).$$

Let us now consider the case of a variable diffusion.

Theorem 2.9. *Let X_ϵ be a weak solution of the equation*

$$X_\epsilon(t) = x + \int_0^t b(X_\epsilon(s))\,ds + \sqrt{\epsilon}\int_0^t \sigma(X_\epsilon(s))\,dB_\epsilon(s),$$

where B_ϵ is a standard Brownian motion, $b(\cdot) : \mathbb{R}^d \to \mathbb{R}^d$ is a bounded continuous function, $\sigma(\cdot)$ is a bounded continuous function with values in the set of symmetric nonnegative definite matrices, and the matrix $c(x) = \sigma(x)\sigma(x)^T$ is uniformly elliptic. Then the X_ϵ obey the LDP in $\mathbb{C}(\mathbb{R}^d)$ with

$$\mathbf{I}(\mathbf{x}) = \frac{1}{2}\int_0^\infty (\dot{\mathbf{x}}(t) - b(\mathbf{x}(t)))^T c(\mathbf{x}(t))^{-1}(\dot{\mathbf{x}}(t) - b(\mathbf{x}(t)))\,dt$$

if $\mathbf{x}$ is absolutely continuous and $\mathbf{x}(0) = x$, and $\mathbf{I}(\mathbf{x}) = \infty$ otherwise.

The following lemma will be useful both in the proof of this theorem, and below. Let Λ_0 denot the set of all $\mathbb{R}^d$–valued piecewise constant functions $(\lambda(t),\, t \in \mathbb{R}_+)$ of the form

$$\lambda(t) = \sum_{i=1}^k \lambda_i \mathbf{1}(t \in (t_{i-1}, t_i]),$$

where $0 \le t_0 < t_1 < \ldots < t_k,\ \lambda_i \in \mathbb{R}^d,\ i = 1, \ldots, k,\ k \in \mathbb{N}$.

Lemma 2.1. *Let $f(t,\lambda), t \in \mathbb{R}_+, \lambda \in \mathbb{R}^d$, be an $\mathbb{R}$-valued function, which is Lebesgue measurable in t, continuous in λ, and is such that $f(t,0)=0$ and $\int_0^T f(t,\lambda)dt$ is well defined for $T \in \mathbb{R}_+$ and $\lambda \in \mathbb{R}^d$. Then for $T \in \mathbb{R}_+$*

$$\int_0^T \sup_{\lambda \in \mathbb{R}^d} f(t,\lambda)\ dt = \sup_{(\lambda(t))\in\Lambda_0} \int_0^T f(t,\lambda(t))\, dt.$$

Proof. We denote $F(t) = \sup_{\lambda\in\mathbb{R}^d} f(t,\lambda)$. Since the supremum may be taken over the rational λ in view of continuity of $f(t,\lambda)$ in λ, the function $F(t)$ is Lebesgue measurable and non-negative, so that the integral on the left-hand side of the equality in the statement of the lemma is well defined.

Given arbitrary $\epsilon > 0$, we introduce the set

$$A_\epsilon = \{(t,\lambda) \in [0,T]\times\mathbb{R}^d : \frac{1}{\epsilon} \geq f(t,\lambda) \geq (F(t)-\epsilon)^+ \wedge \frac{1}{\epsilon}\}.$$

By a measurable selection theorem, see, e.g., Clarke [2], Ethier and Kurtz [8], there exists an $\mathbb{R}^d$–valued Lebesgue measurable function $\tilde{\lambda}_\epsilon(t)$ such that

$$\frac{1}{\epsilon} \geq f(t,\tilde{\lambda}_\epsilon(t)) \geq (F(t)-\epsilon)^+ \wedge \frac{1}{\epsilon},\ t\in[0,T].$$

By Luzin's theorem there exists a continuous function $\lambda_\epsilon(t)$ such that $\int_0^T \mathbf{1}\big(\tilde{\lambda}_\epsilon(t) \neq \lambda_\epsilon(t)\big)\, dt < \epsilon^2$. Then

$$\int_0^T f(t,\lambda_\epsilon(t)) \vee 0\, dt \geq \int_0^T f(t,\tilde{\lambda}_\epsilon(t))\, dt - \epsilon \geq \int_0^T (F(t)-\epsilon)^+ \wedge \frac{1}{\epsilon}\, dt - \epsilon.$$

Since $(\lambda_\epsilon(t))$ is continuous, it can be approximated by functions from Λ_0. Since $f(t,\lambda)$ is continuous in λ and $f(t,0)=0$, by Fatou's lemma there exists a function $\lambda_0 \in \Lambda_0$ such that

$$\int_0^T f(t,\lambda_0(t))\, dt \geq \int_0^T f(t,\lambda_\epsilon(t)) \vee 0\, dt - \epsilon.$$

Thus, since $\epsilon > 0$ is arbitrary,

$$\int_0^T \sup_{\lambda \in \mathbb{R}^d} f(t,\lambda)\ dt \leq \sup_{(\lambda(t))\in\Lambda_0} \int_0^T f(t,\lambda(t))\, dt.$$

The reverse inequality is obvious. □

Proof of Theorem 2.9. We apply Corollary 1.4 and start by checking exponential tightness. We assume that X_ϵ and B_ϵ are defined on a filtered probability space $(\mathbf{P}_\epsilon, \mathcal{F}_\epsilon, \mathbf{F}_\epsilon)$, where $\mathbf{F}_\epsilon = (\mathcal{F}_\epsilon(t),\ t \in \mathbb{R}_+)$ is a filtration. We have for $s < t$

$$X_\epsilon(t) - X_\epsilon(s) = \int_s^t b(X_\epsilon(u))\, du + \sqrt{\epsilon}\int_s^t \sigma(X_\epsilon(u))\, dB_\epsilon(u).$$

Therefore, for $\mu > 0$ and $\eta > 0$,

$$\mathbf{P}_\epsilon\big(\sup_{t\in[s,s+\delta]}|X_\epsilon(t)-X_\epsilon(s)|>\eta\big)\le \mathbf{P}_\epsilon\big(\int_s^{s+\delta}|b(X_\epsilon(u))|\,du>\frac{\eta}{2}\big)$$
$$+\mathbf{P}_\epsilon\big(\sqrt{\epsilon}\sup_{t\in[s,s+\delta]}\Big|\int_s^t\sigma(X_\epsilon(u))\,dB_\epsilon(u)\Big|>\frac{\eta}{2}\big).$$

Since the function $|b(\cdot)|$ is bounded, the first probability on the right side is equal to zero for δ small enough. The second probability is bounded in analogy with the proof of Theorem 2.7: by (2.2)

$$\mathbf{P}_\epsilon\big(\sqrt{\epsilon}\sup_{t\in[s,s+\delta]}\Big|\int_s^t\sigma(X_\epsilon(u))\,dB_\epsilon(u)\Big|>\frac{\eta}{2}\big)$$
$$\le \exp\big(-\frac{\mu\eta}{2\epsilon\sqrt{d}}\big)\,2d\max_{i=1,\ldots,2d}\mathbf{E}_\epsilon\exp\big(\frac{\mu}{\sqrt{\epsilon}}\,e_i\cdot\int_t^{t+\delta}\sigma(X_\epsilon(u))\,dB_\epsilon(u)\big)$$
$$=\exp\big(-\frac{\mu\eta}{2\epsilon\sqrt{d}}\big)\,2d\max_{i=1,\ldots,d}\mathbf{E}_\epsilon\exp\big(\frac{\mu^2}{2\epsilon}\int_t^{t+\delta}|\sigma(X_\epsilon(u))e_i|^2\,du\big),$$

where the latter equality holds as $\sigma(\cdot)$ is bounded. Using boundedness of σ once again, we conclude that

$$\limsup_{\delta\to0}\limsup_{\epsilon\to0}\mathbf{P}_\epsilon\big(\sqrt{\epsilon}\sup_{t\in[s,s+\delta]}\Big|\int_s^t\sigma(X_\epsilon(u))\,dB_\epsilon(u)\Big|>\frac{\eta}{2}\big)^\epsilon\le\exp\big(-\frac{\mu\eta}{2\sqrt{d}}\big)$$

Since μ is arbitrary, the left-hand side is equal to zero. The proof of exponential tightness is over.

We prove the upper bound in Corollary 1.4. Let $\lambda(t)$ be a function from Λ_0. We define for $\mathbf{x}\in\mathbb{C}(\mathbb{R}^d)$

$$\int_0^t\lambda(s)\cdot dX_\epsilon(s)=\sum_{i=1}^k\lambda_i\cdot(X_\epsilon(t\wedge t_i)-X_\epsilon(t\wedge t_{i-1})). \tag{2.4}$$

Since the functions b and σ are bounded, the stochastic process $Y_\epsilon=(Y_\epsilon(t),\,t\in\mathbb{R}_+)$ defined by

$$Y_\epsilon(t)=\exp\Big(\int_0^t\frac{\lambda(s)}{\epsilon}\cdot dX_\epsilon(s)-\int_0^t\frac{\lambda(s)}{\epsilon}\cdot b(X_\epsilon(s))\,ds-\frac{\epsilon}{2}\int_0^t|\sigma(X_\epsilon(s))\frac{\lambda(s)}{\epsilon}|^2\,ds\Big)$$

is a martingale on $(\mathbf{P}_\epsilon,\mathcal{F}_\epsilon,\mathbf{F}_\epsilon)$. By the equality $\mathbf{E}_\epsilon Y_\epsilon(t)=1$, it follows that for $\delta>0$

$$\mathbf{E}_\epsilon Y_\epsilon(t)\mathbf{1}(\rho_{\mathbb{C}}(X_\epsilon,\mathbf{x})<\delta)\le 1. \tag{2.5}$$

Let $\gamma>0$ be arbitrary. By (2.4)

$$|\int_0^t\lambda(s)\cdot dX_\epsilon(s)-\int_0^t\lambda(s)\cdot d\mathbf{x}(s)|\le 2\sup_{s\in[0,t]}|X_\epsilon(s)-\mathbf{x}(s)|\sum_{i=1}^k|\lambda_i|$$

Therefore, if $\delta > 0$ is small enough then

$$|\int_0^t \lambda(s)\cdot dX_\epsilon(s) - \int_0^t \lambda(s)\cdot d\mathbf{x}(s)| \le \gamma$$

on the set where $\rho_{\mathbb{C}}(X_\epsilon,\mathbf{x}) < \delta$. Similarly, by continuity of $b(\cdot)$ and $\sigma(\cdot)$ we may assume that

$$|\int_0^t \lambda(s)\cdot b(X_\epsilon(s))\,ds - \int_0^t \lambda(s)\cdot b(\mathbf{x}(s))\,ds| < \gamma,$$

$$|\frac{1}{2}\int_0^t |\sigma(X_\epsilon(s))\lambda(s)|^2\,ds - \frac{1}{2}\int_0^t |\sigma(\mathbf{x}(s))\lambda(s)|^2\,ds| < \gamma$$

on the set where $\rho_{\mathbb{C}}(X_\epsilon,\mathbf{x}) < \delta$. Hence, by (2.5) for δ small enough

$$e^{-3\gamma/\epsilon}\exp\frac{1}{\epsilon}\Big(\int_0^t \lambda(s)\cdot d\mathbf{x}(s) - \int_0^t \lambda(s)\cdot b(\mathbf{x}(s))\,ds - \frac{1}{2}\int_0^t |\sigma(\mathbf{x}(s))\lambda(s)|^2\,ds\Big)\mathbf{P}_\epsilon\big(\rho_{\mathbb{C}}(X_\epsilon,\mathbf{x}) < \delta\big) \le 1.$$

Hence,

$$\limsup_{\delta\to 0}\limsup_{\epsilon\to 0}\mathbf{P}_\epsilon\big(\rho_{\mathbb{C}}(X_\epsilon,\mathbf{x}) < \delta\big)^\epsilon \le e^{3\gamma}\exp -\Big(\int_0^t \lambda(s)\cdot d\mathbf{x}(s) - \int_0^t \lambda(s)\cdot b(\mathbf{x}(s))\,ds - \frac{1}{2}\int_0^t |\sigma(\mathbf{x}(s))\lambda(s)|^2\,ds\Big)$$

Since γ, $\lambda(\cdot)$ and t are arbitrary, we obtain that

$$\limsup_{\delta\to 0}\limsup_{\epsilon\to 0}\epsilon\log\mathbf{P}_\epsilon\big(\rho_{\mathbb{C}}(X_\epsilon,\mathbf{x}) < \delta\big) \le -\sup_{\lambda(\cdot)\in\Lambda_0}\Big(\int_0^\infty \lambda(s)\cdot d\mathbf{x}(s) - \int_0^\infty \lambda(s)\cdot b(\mathbf{x}(s))\,ds - \frac{1}{2}\int_0^\infty |\sigma(\mathbf{x}(s))\lambda(s)|^2\,ds\Big)$$

By Lemma 2.1, if $\mathbf{x}(\cdot)$ is absolutely continuous and starts at zero, the latter supremum equals

$$\int_0^\infty \sup_{\lambda\in\mathbb{R}^d}\big(\lambda\cdot\dot{\mathbf{x}}(s) - \lambda\cdot b(\mathbf{x}(s)) - \frac{1}{2}|\sigma(\mathbf{x}(s))\lambda|^2\big)\,ds = \mathbf{I}(\mathbf{x}).$$

If $\mathbf{x}$ either is not absolutely continuous or does not start at zero, then the supremum is equal to infinity (check). Thus,

$$\limsup_{\delta\to 0}\limsup_{\epsilon\to 0}\epsilon\log\mathbf{P}_\epsilon\big(\rho_{\mathbb{C}}(X_\epsilon,\mathbf{x}) < \delta\big) \le -\mathbf{I}(\mathbf{x}).$$

We now check that

$$\liminf_{\delta\to 0}\liminf_{\epsilon\to 0}\epsilon\log\mathbf{P}_\epsilon\big(\rho_{\mathbb{C}}(X_\epsilon,\mathbf{x}) < \delta\big) \ge -\mathbf{I}(\mathbf{x}). \tag{2.6}$$

Since the inequality is clearly true if $\mathbf{I}(\mathbf{x}) = \infty$, we assume that $\mathbf{x}(0) = 0$ and $\mathbf{x}$ is absolutely continuous. Let us introduce the stochastic process

$$\lambda_\epsilon(t) = c(X_\epsilon(t))^{-1}(\dot{\mathbf{x}}(t) - b(X_\epsilon(t))). \tag{2.7}$$

By the ellipticity condition on $c(\cdot)$, boundedness of $b(\cdot)$ and the fact that $\mathbf{I}(\mathbf{x}) < \infty$, we obtain that $\int_0^t |\lambda_\epsilon(s)|^2 \, ds < \infty$ a.s. Hence, the stochastic integral $\int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s) \, dB_\epsilon(s)$ and the Lebesgue integral $\int_0^t |\sigma(X_\epsilon(s))\lambda_\epsilon(s)|^2 \, ds$ are well defined. Let

$$Z_\epsilon(t) = \exp\Big(\frac{1}{\sqrt{\epsilon}} \int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s) \, dB_\epsilon(s) - \frac{1}{2\epsilon} \int_0^t |\sigma(X_\epsilon(s))\lambda_\epsilon(s)|^2 \, ds\Big). \tag{2.8}$$

It is a martingale (since $b(\cdot)$ and $\sigma(\cdot)$ are bounded, so we define a new probability measure $\mathbf{Q}_\epsilon$ by $d\mathbf{Q}_\epsilon|_{\mathcal{F}_t} = Z_\epsilon(t) d\mathbf{P}_\epsilon|_{\mathcal{F}_t}$. By Girsanov's theorem, see, e.g., Ikeda and Watanabe [10], $B_\epsilon(t) - (1/\sqrt{\epsilon}) \int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s) \, ds$ is a Brownian motion under measure $\mathbf{Q}_\epsilon$. We denote it $\tilde{B}_\epsilon$. Hence, under this measure

$$X_\epsilon(t) = x + \mathbf{x}(t) + \sqrt{\epsilon} \int_0^t \sigma(X_\epsilon(s)) \, d\tilde{B}_\epsilon(s). \tag{2.9}$$

We also note that

$$Z_\epsilon(t) = \exp\Big(\frac{1}{\sqrt{\epsilon}} \int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s) \, d\tilde{B}_\epsilon(s) + \frac{1}{2\epsilon} \int_0^t |\sigma(X_\epsilon(s))\lambda_\epsilon(s)|^2 \, ds\Big). \tag{2.10}$$

In addition, by the continuity and ellipticity assumptions and (2.7) given arbitrary $\gamma > 0$ there exists $\delta > 0$ such that on the set where $\sup_{t \in [0,T]} |X_\epsilon(t) - \mathbf{x}(t)| < \delta$ the bound holds

$$\Big|\int_0^t |\sigma(X_\epsilon(s))\lambda_\epsilon(s)|^2 \, ds - \int_0^t (\dot{\mathbf{x}}(s) - b(\mathbf{x}(s)))^T c(\mathbf{x}(t))^{-1} (\dot{\mathbf{x}}(s) - b(\mathbf{x}(s))) \, ds\Big| < \gamma.$$

We therefore have

$$\begin{aligned}
\mathbf{P}_\epsilon\big(\sup_{t \in [0,T]} |X_\epsilon(t) - \mathbf{x}(t)| < \delta\big) &= \int_{\Omega_\epsilon} \mathbf{1}\big(\sup_{t \in [0,T]} |X_\epsilon(t) - \mathbf{x}(t)| < \delta\big) Z_\epsilon(t)^{-1} \, d\mathbf{Q}_\epsilon \\
&\geq \exp\Big(-\frac{1}{2\epsilon} \int_0^t (\dot{\mathbf{x}}(s) - b(\mathbf{x}(s)))^T c(\mathbf{x}(t))^{-1} (\dot{\mathbf{x}}(s) - b(\mathbf{x}(s))) \, ds\Big) \exp\big(-\frac{\gamma}{2\epsilon}\big) \\
&\int_{\Omega_\epsilon} \mathbf{1}\big(\sup_{t \in [0,T]} |X_\epsilon(t) - \mathbf{x}(t)| < \delta\big) \exp\Big(-\frac{1}{\sqrt{\epsilon}} \int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s) \, d\tilde{B}_\epsilon(s)\Big) \, d\mathbf{Q}_\epsilon.
\end{aligned}$$

Next,

$$\begin{aligned}
&\int_{\Omega_\epsilon} \mathbf{1}\big(\sup_{t \in [0,T]} |X_\epsilon(t) - \mathbf{x}(t)| < \delta\big) \exp\Big(-\frac{1}{\sqrt{\epsilon}} \int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s) \, d\tilde{B}_\epsilon(s)\Big) \, d\mathbf{Q}_\epsilon \\
&\geq \exp\big(-\frac{\gamma}{\epsilon}\big) \mathbf{Q}_\epsilon\Big(\sup_{t \in [0,T]} |X_\epsilon(t) - \mathbf{x}(t)| < \delta, \ \int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s) \, d\tilde{B}_\epsilon(s) < \frac{\gamma}{\sqrt{\epsilon}}\Big).
\end{aligned}$$

Putting everything together yields for all δ small enough

$$\liminf_{\epsilon\to 0} \epsilon \log \mathbf{P}_\epsilon\big(\sup_{t\in[0,T]} |X_\epsilon(t) - \mathbf{x}(t)| < \delta\big) \geq -\mathbf{I}(\mathbf{x}) - 3\gamma$$

$$- \limsup_{\epsilon\to 0} \epsilon \log \mathbf{Q}_\epsilon\big(\sup_{t\in[0,T]} |X_\epsilon(t) - \mathbf{x}(t)| < \delta, \int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s)\, d\tilde{B}_\epsilon(s) < \frac{\gamma}{\sqrt{\epsilon}}\big) \quad (2.11)$$

By Chebyshev,

$$\mathbf{Q}_\epsilon\big(\int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s)\, d\tilde{B}_\epsilon(s) \geq \frac{\gamma}{\sqrt{\epsilon}}\big) \leq \frac{\epsilon}{\gamma^2} \int_{\Omega_\epsilon} |\int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s)\, d\tilde{B}_\epsilon(s)|^2 \, d\mathbf{Q}_\epsilon$$

$$\leq \frac{\epsilon}{\gamma^2} \int_{\Omega_\epsilon} \int_0^t \|\sigma(X_\epsilon(s))\|^2 |\lambda_\epsilon(s)|^2 \, ds\, d\mathbf{Q}_\epsilon,$$

which tends to zero as $\epsilon \to 0$ by the boundedness of $\sigma(\cdot)$ and integrability of $|\lambda_\epsilon(s)|^2$. Also, (2.9) obviously implies that $\lim_{\epsilon\to 0} \mathbf{Q}_\epsilon\big(\sup_{t\in[0,T]}|X_\epsilon(t) - \mathbf{x}(t)| < \delta\big) = 1$. Hence,

$$\lim_{\epsilon\to 0} \mathbf{Q}_\epsilon\big(\sup_{t\in[0,T]} |X_\epsilon(t) - \mathbf{x}(t)| < \delta,\ \int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s)\, d\tilde{B}_\epsilon(s) < \frac{\gamma}{\sqrt{\epsilon}}\big) = 1,$$

so

$$\lim_{\epsilon\to 0} \epsilon \log \mathbf{Q}_\epsilon\big(\sup_{t\in[0,T]} |X_\epsilon(t) - \mathbf{x}(t)| < \delta, \int_0^t \sigma(X_\epsilon(s))\lambda_\epsilon(s)\, d\tilde{B}_\epsilon(s) < \frac{\gamma}{\sqrt{\epsilon}}\big) = 0,$$

which implies by (2.11) the bound (2.6). □

Exercise 2.13. *Finish the proof.*

2.3 Space $\mathbb{D}$

Let $\mathbb{D}(\mathbb{R}^d)$ denote the space of $\mathbb{R}^d$-valued right-continuous functions on $\mathbb{R}_+$ that possess left-hand limits. We recall its basic properties, see Billingsley [1], Jacod and Shiryaev [11], Ethier and Kurtz [8] for more detail. Space $\mathbb{D}(\mathbb{R}^d)$ is turned into a metric space as follows. Let Λ denote the set of strictly increasing continuous functions $\lambda(t)$, $t \in \mathbb{R}_+$, with $\lambda(0) = 0$ and $\lambda(t) \to \infty$ as $t \to \infty$ such that

$$\gamma(\lambda) = \sup_{0\leq s<t} \log \big| \frac{\lambda(t) - \lambda(s)}{t - s} \big| < \infty.$$

We define

$$\rho_{\mathbb{D}}(\mathbf{x}, \mathbf{y}) = \inf_{\lambda\in\Lambda} \big(\gamma(\lambda) \vee \int_0^\infty e^{-u} \sup_{t\in[0,u]} |\mathbf{x}(t) - \mathbf{y}(\lambda(t))| \wedge 1\, du \big).$$

Space $\mathbb{D}(\mathbb{R}^d)$ equipped with metric $\rho_{\mathbb{D}}$ is complete and separable. It is also known that $\mathbf{x}_n \to \mathbf{x}$ in $(\mathbb{D}, \rho_{\mathbb{D}})$ if and only if there exists a sequence $\lambda_n \in \Lambda$ such that $\lim_{n\to\infty} \sup_{t\in\mathbb{R}_+} |\lambda_n(t) - t| = 0$ and $\lim_{n\to\infty} \sup_{t\in[0,T]} |\mathbf{x}(t) - \mathbf{x}_n(\lambda_n(t))| = 0$ for all T.

Compact sets in $\mathbb{D}(\mathbb{R}^d)$ are described as follows. For $\mathbf{x} \in \mathbb{D}$, $T > 0$ and $\delta > 0$, we define the modulus of continuity

$$w'_T(\mathbf{x}, \delta) = \inf_{(t_j)} \max_{j=1,\ldots,k} w_{\mathbf{x}}\big([t_{j-1}, t_j)\big),$$

where $w_{\mathbf{x}}\big([s,t)\big) = \sup_{u,v \in [s,t)} |\mathbf{x}_u - \mathbf{x}_v|$, $s < t$, and the infimum is taken over all collections (t_j) such that $0 = t_0 < t_1 < \ldots < t_k = T$ and $t_j - t_{j-1} > \delta$ for $j < k$. Note that $w'_T(\mathbf{x}, \delta) \le w_T(\mathbf{x}, 2\delta)$ for δ small enough.

Theorem 2.10. *Set $K \subset \mathbb{D}(\mathbb{R}^d)$ is relatively compact if and only if the following hold:*

(i) $\sup_{\mathbf{x} \in K} \sup_{t \in [0,T]} |\mathbf{x}(t)| < \infty$,

(ii) $\lim_{\delta \to 0} \sup_{\mathbf{x} \in K} w'_T(\mathbf{x}, \delta) = 0$ *for all* $T > 0$.

Theorem 2.11. *A sequence $\mathbf{P}_n$ of distributions on $\mathbb{D}(\mathbb{R}^d)$ is exponentially tight if and only if the following conditions hold*

(i) $\lim_{A \to \infty} \mathbf{P}_n(\mathbf{x} : \sup_{t \in [0,T]} |\mathbf{x}(t)| > A)^{1/r_n} = 0$,

(ii) $\lim_{\delta \to 0} \mathbf{P}_n\big(\mathbf{x} : w'_T(\mathbf{x}, \delta) > \epsilon\big)^{1/r_n} = 0$ *for all* $T > 0$ *and* $\epsilon > 0$.

In many cases, the limit rate functions are equal to infinity at discontinuous functions, so we introduce the following definition which adapts the definition of a $\mathbb{U}_0$-exponentially tight sequence to this particular setting. A sequence $\mathbf{P}_n$ of probability measures on $\mathbb{D}(\mathbb{R}^d)$ is said to be $\mathbb{C}$-exponentially tight if $\mathbf{\Pi}(\mathbb{D}(\mathbb{R}^d) \setminus \mathbb{C}(\mathbb{R}^d)) = 0$ for every LD accumulation point $\mathbf{\Pi}$ of $\mathbf{P}_n$.

Theorem 2.12. *A sequence $\mathbf{P}_n$ of distributions on $\mathbb{D}(\mathbb{R}^d)$ is $\mathbb{C}$-exponentially tight if and only if the following conditions hold*

(i) $\lim_{A \to \infty} \mathbf{P}_n(\mathbf{x} : \sup_{t \in [0,T]} |\mathbf{x}(t)| > A)^{1/r_n} = 0$,

(ii) $\lim_{\delta \to 0} \mathbf{P}_n\big(\mathbf{x} : w_T(\mathbf{x}, \delta) > \epsilon\big)^{1/r_n} = 0$ *for all* $T > 0$ *and* $\epsilon > 0$.

Theorem 2.13. *A sequence $\mathbf{P}_n$ of distributions on $\mathbb{D}(\mathbb{R}^d)$ is $\mathbb{C}$-exponentially tight if and only if the following conditions hold*

(i) $\lim_{A \to \infty} \mathbf{P}_n(\mathbf{x} : \sup_{t \in [0,T]} |\mathbf{x}(t)| > A)^{1/r_n} = 0$,

(ii) $\lim_{\delta \to 0} \sup_{t \in [0,T]} \mathbf{P}_n\big(\mathbf{x} : \sup_{s \in [t, t+\delta]} |\mathbf{x}(s) - \mathbf{x}(t)| > \epsilon\big)^{1/r_n} = 0$ *for all* $T > 0$ *and* $\epsilon > 0$.

Let us say that a sequence of stochastic processes is $\mathbb{C}$-exponentially tight if the sequence of their distributions is $\mathbb{C}$-exponentially tight.

Theorem 2.14. *Let $X_n = (X_n(t),\, t \in \mathbb{R}_+)$ be a sequence of stochastic processes with trajectories in $\mathbb{D}(\mathbb{R}^d)$. Suppose the following conditions hold:*

1. *finite-dimensional projections $(X_n(t_1), \ldots, X_n(t_k))$ obey the LDP in $(\mathbb{R}^d)^k$ with $\mathbf{I}_{t_1,\ldots,t_k}$,*
2. *the sequence X_n is $\mathbb{C}$-exponentially tight.*

Then the sequence X_n obeys the LDP in $\mathbb{D}(\mathbb{R}^d)$ with $\mathbf{I}(\mathbf{x}) = \sup_{t_1,\ldots,t_k} \mathbf{I}_{t_1,\ldots,t_k}(\mathbf{x}(t_1), \ldots, \mathbf{x}(t_k))$ if $\mathbf{x} \in \mathbb{C}$ and $\mathbf{I}(\mathbf{x}) = \infty$ otherwise.

Exercise 2.14. *Prove the stated theorems.*

We give applications to concrete stochastic processes.

Theorem 2.15 (Mogulskii)**.** *Let ξ_i be i.i.d. $\mathbb{R}^d$-valued r.v. with $\mathbf{E}\exp(\lambda\cdot\xi_1)<\infty$ for all $\lambda\in\mathbb{R}^d$. Let*

$$X_n(t)=\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor}\xi_i.$$

Then the processes X_n obey the LDP in $\mathbb{D}(\mathbb{R}^d)$ with

$$\mathbf{I}(\mathbf{x})=\int_0^\infty\sup_{\lambda\in\mathbb{R}^d}\big(\lambda\cdot\dot{\mathbf{x}}(t)-\log\mathbf{E}\exp(\lambda\cdot\xi_1)\big)\,dt$$

if $\mathbf{x}$ is absolutely continuous and $\mathbf{x}(0)=0$, and $\mathbf{I}(\mathbf{x})=\infty$ otherwise.

Proof. We start the proof by checking $\mathbb{C}$-exponential tightness as stipulated by Theorem 2.13. We have for $T>0$ and $A>0$ by Markov's inequality

$$\mathbf{P}\big(\sup_{t\in[0,T]}|X_n(t)|>A\big)\le\mathbf{P}\big(\sum_{i=1}^{\lfloor nT\rfloor}|\xi_i|>nA\big)\le\exp(-nA)\big(\mathbf{E}\exp(|\xi_1|)\big)^{nT}.$$

Hence,

$$\limsup_{n\to\infty}\mathbf{P}\big(\sup_{t\in[0,T]}|X_n(t)|>A\big)^{1/n}\le\exp(-A)\big(\mathbf{E}\exp(|\xi_1|)\big)^{T},$$

which tends to zero as $A\to\infty$. Similarly, for $\mu>0$, $t>0$ and $\delta>0$,

$$\mathbf{P}\big(\sup_{s\in[t,t+\delta]}|X_n(s)-X_n(t)|>\epsilon\big)\le\mathbf{P}\big(\sum_{i=\lfloor nt\rfloor+1}^{\lfloor n(t+\delta)\rfloor}\mu|\xi_i|>\mu n\epsilon\big)\le\exp(-\mu n\epsilon)\big(\mathbf{E}\exp(\mu|\xi_1|)\big)^{n\delta}.$$

Hence,

$$\limsup_{\delta\to0}\limsup_{n\to\infty}\sup_{t\in[0,T]}\mathbf{P}\big(\sup_{s\in[t,t+\delta]}|X_n(s)-X_n(t)|>\epsilon\big)^{1/n}\le\exp(-\mu\epsilon).$$

Since μ is arbitrarily large, the left-hand side is equal to zero. The proof of $\mathbb{C}$-exponential tightness is over.

Let $0\le t_1\le\ldots\le t_k$. By Cramér's theorem and independence of increments the $(X_n(t_1),\ldots,X_n(t_k))$ obey the LDP in $(\mathbb{R}^d)^k$ with

$$\mathbf{I}_{t_1,\ldots,t_k}(x_1,\ldots,x_k)=\sum_{i=1}^k\sup_{\lambda\in\mathbb{R}^d}(\lambda\cdot(x_i-x_{i-1})-(t_i-t_{i-1})\log\mathbf{E}\exp(\lambda\cdot\xi_1)),$$

where $t_0=0$. Hence, $\sup_{t_1,\ldots,t_k}\mathbf{I}_{t_1,\ldots,t_k}(\mathbf{x}(t_1),\ldots,\mathbf{x}(t_k))$ coincides with $\mathbf{I}(\mathbf{x})$ in the statement of the theorem (use Lemma 2.1). □

Exercise 2.15. *Fill in the details.*

Exercise 2.16. *Let $X_n(t) = N(nt)/n$, where $N(t)$ is a Poisson process of rate λ. Show that the sequence X_n obeys the LDP for rate n with*

$$\mathbf{I}(\mathbf{x}) = \int_0^\infty \big(\dot{\mathbf{x}}(t) \log \dot{\mathbf{x}}(t) - \dot{\mathbf{x}}(t) + 1\big)\, dt$$

if $\mathbf{x}$ is absolutely continuous, nondecreasing and $\mathbf{x}(0) = 0$, and $\mathbf{I}(\mathbf{x}) = \infty$ otherwise.

Exercise 2.17. *(Compound Poisson) Let $N(\cdot)$ be a Poisson process of rate γ and ξ_i be $\mathbb{R}$-valued i.i.d., independent of $N(\cdot)$, with $\mathbf{E} \exp \lambda\xi_1 < \infty$ for all $\lambda \in \mathbb{R}$. Derive the LDP for*

$$X_n(t) = \frac{1}{n} \sum_{i=1}^{N(nt)} \xi_i.$$

References

[1] P. Billingsley. *Convergence of Probability Measures.* Wiley, 1968.

[2] F.H. Clarke. *Optimization and Nonsmooth Analysis.* Wiley, 1983.

[3] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Springer, 2nd edition, 1998.

[4] F. den Hollander. *Large deviations*, volume 14 of *Fields Institute Monographs.* American Mathematical Society, Providence, RI, 2000.

[5] J. D. Deuschel and D. W. Stroock. *Large Deviations.* Academic Press, 1989.

[6] J. D. Deuschel and D. W. Stroock. *Large Deviations.* American Mathematical Society, second edition, 2001.

[7] P. Dupuis and R. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations.* Wiley, 1997.

[8] S. N. Ethier and T. G. Kurtz. *Markov Processes. Characterization and Convergence.* Wiley, 1986.

[9] M.I. Freidlin and A.D. Wentzell. *Random Perturbations of Dynamical Systems.* Springer, 2nd edition, 1998.

[10] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North Holland, 2nd edition, 1989.

[11] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, second edition, 2003.

[12] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[13] A. Puhalskii. *Large Deviations and Idempotent Probability.* Chapman & Hall/CRC, 2001.

[14] A. Shwartz and A. Weiss. *Large Deviations for Performance Analysis*. Chapman & Hall, 1995.

[15] N.N. Vakhania, V.I. Tarieladze, and S.A. Chobanyan. *Probability Distributions on Banach Spaces*. Nauka, 1985. (in Russian, English translation: Reidel, 1987).

[16] S. R. S. Varadhan. *Large Deviations and Applications*. SIAM, 1984.