Large deviations results of Gärtner-Ellis type with an application to insurance

Robin McCann
University of Helsinki
2022

Master's thesis

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty		Koulutusohjelma — Utbildningsprogram — Degree programme
Matemaattis-luonnontieteellinen		Matematiikan ja tilastotieteen maisteriohjelma
Tekijä — Författare — Author		
Robin McCann		
Työn nimi — Arbetets titel — Title		
Large deviations results of Gärtner-Ellis type with an application to insurance		
Oppiaine — Läroämne — Subject		
Matematiikka		
Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Pro gradu -tutkielma	Tammikuu 2022	61 s.

Tiivistelmä — Referat — Abstract

Large deviations theory is a branch of probability theory which studies the exponential decay of probabilities for extremely rare events in the context of sequences of probability distributions. The theory originates from actuaries studying risk and insurance from a mathematical perspective, but today it has become its own field of study, and is no longer as tightly linked to insurance mathematics. Large deviations theory is nowadays frequently applied in various fields, such as information theory, queuing theory, statistical mechanics and finance. The connection to insurance mathematics has not grown obsolete, however, and these new results can also be applied to develop new results in the context of insurance.

This paper is split into two main sections. The first presents some basic concepts from large deviations theory as well as the Gärtner-Ellis theorem, the first main topic of this thesis, and then provides a fairly detailed proof of this theorem. The Gärtner-Ellis theorem is an important result in large deviations theory, as it gives upper and lower bounds relating to asymptotic probabilities, while allowing for some dependence structure in the sequence of random variables.

The second main topic of this thesis is the presentation of two large deviations results developed by H. Nyrhinen, concerning the random time of ruin as a function of the given starting capital. This section begins with introducing the specifics of this insurance setting of Nyrhinen's work as well as the ruin problem, a central topic of risk theory. Following this are the main results, and the corresponding proofs, which rely to some part on convex analysis, and also on a continuous version of the Gärtner-Ellis theorem.

Recommended preliminary knowledge: Probability Theory, Risk Theory.

Avainsanat — Nyckelord — Keywords

Probability, Gärtner-Ellis, Large Deviations, Risk Theory, Ruin Probabilities, Insurance

Säilytyspaikka — Förvaringsställe — Where deposited

Kumpulan tiedekirjasto

Muita tietoja — Övriga uppgifter — Additional information

Contents

1 Introduction

Whereas Cramér [6] in 1938 already went beyond the traditional Law of Large Numbers and Central Limit Theorem, the large deviation principle, formalised in 1966 by Varadhan [5], goes even further and describes the behaviour of a sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ at very unlikely events. Informally, the idea is that far-off events become exponentially unlikely, and for very large n the sequence of probability measures behaves like $\mathrm{d}\mathbb{P}_n(x) \approx e^{-nI(x)}$, where I is called a rate function, and describes the rate of decay. This, and some more background, is laid out in more detail in [1, 2, 5].

Here one might stop to ask oneself why we even care about highly unlikely events. To give an example, sometimes extremely anomalous situations can have a massive impact, take for instance a supervolcano erupting, a large asteroid hitting earth, or perhaps something more mundane, such as a person winning the lottery. These events are significant enough not to be disregarded purely based on their rarity. Large deviations theory can be applied in many situations where we care about very uncommon events. It is frequently applied in information theory, queuing theory, statistical physics, finance, or, as in this paper, to the ruin problem, a fundamental topic of risk theory.

Continuing the previous train of thought, it is worth noting that Cramér's theorem only concerns empirical means of i.i.d. random variables, a rather sterile situation. The Gärtner-Ellis Theorem, developed by Gärtner and Ellis [3, 4], and presented as Theorem 2.29 in this thesis, considers a more general setting, one where the object of study is instead a sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$, distributed respectively according to laws $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$, which in some sense resembles the sequence of empirical means studied by Cramér, but which admits a weak dependence structure. Under certain limit conditions on the sequence $\{X_n\}_{n\in\mathbb{N}}$, the Gärtner-Ellis theorem implies the Large Deviation Principle for the sequence of probability measures $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$.

If one then, instead of examining sequences of empirical means and generalisations thereof, considers a stochastic process Y_n as a generalisation of a random walk, one can still obtain some large deviations results. In an insurance setting where the starting capital is M, and Y_n is the loss process, considering the time of ruin $T_M = \inf\{n : Y_n > M\}$, we present two results from Nyrhinen [9, 10], which together imply the Large Deviation Principle for the measures $\mathbb{P}_M(\cdot) = \mathbb{P}(T/M \in \cdot)$, when $M \to \infty$.

As follows is the structure of this thesis: Chapter 2 concerns generalisations of empirical means, and consists of an introduction of some relevant and necessary concepts and results, followed by a fairly detailed presentation and proof of the Gärtner-Ellis Theorem. The proof is split into several parts. Chapter 3 shifts our perspective to generalisations of random walks. Here we introduce an insurance setting in a more detailed manner, define a few necessary ideas, state a number of assumptions we will be working under, as well as Theorems 3.11 and 3.14, the two main results of this Chapter. We then recall or prove several auxiliary results from convex analysis and large deviations theory. What remains of Chapter 3 is dedicated to the proofs of the main theorems.

2 The Gärtner-Ellis Theorem

This chapter aims to introduce some concepts from large deviations theory, and present a quite detailed proof of the Gärtner-Ellis theorem, a central result in this field. Here it is presented as Theorem 2.29, and we will be proving the version of the theorem concerning random variables in $\mathbb{R}^d$.

Before introducing any new material, we recall a few concepts, in the forms they will be useful to us, as well as clarify some notation. First up is Hölder's inequality. Let X and Y be real-valued random variables, and suppose $p, q \geq 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$\mathbb{E}(XY) \leq \mathbb{E}\left(X^{\frac{1}{p}}\right)^p \mathbb{E}\left(Y^{\frac{1}{q}}\right)^q,$$

or, in a formulation that is more useful to us, if $\theta \in [0, 1]$, then

$$\mathbb{E}\left(X^\theta Y^{(1-\theta)}\right) \leq \mathbb{E}(X)^\theta \mathbb{E}(Y)^{(1-\theta)}. \tag{2.1}$$

Second, the exponential form of Chebychev's inequality. For each $a \in \mathbb{R}$, it holds that

$$\begin{aligned} \mathbb{P}(X \geq a) &\leq e^{-ta}\mathbb{E}\left(e^{tX}\right) \text{ for all } t \geq 0, \text{ and} \\ \mathbb{P}(X \leq a) &\leq e^{-ta}\mathbb{E}\left(e^{tX}\right) \text{ for all } t \leq 0. \end{aligned} \tag{2.2}$$

It is worth noting that the above bounds hold even if the expectation $\mathbb{E}\left(e^{tX}\right)$ is not finite, although they are not very interesting in that case.

Third, exponentially tilted probability measures. Let X be a random variable distributed according to law $\mathbb{P}$ and $\theta \in \mathbb{R}^d$ a point such that $M_X(\theta) \in (0, \infty)$, where M_X is the moment generating function of X. Then the θ-exponentially tilted random variable $\widehat{X}$ is distributed according to the density

$$\mathrm{d}\widehat{\mathbb{P}}(x) = \exp(\theta x - \log M_X(\theta))\,\mathrm{d}\mathbb{P}(x).$$

Note that $\widehat{\mathbb{P}}$ is guaranteed to be a probability measure because

$$\int_{\mathbb{R}^d} \mathrm{d}\widehat{\mathbb{P}} = \int_{\mathbb{R}^d} e^{\theta x - \log M_X(\theta)}\,\mathrm{d}\mathbb{P} = \frac{1}{M_X(\theta)} \int_{\mathbb{R}^d} e^{\theta x}\,\mathrm{d}\mathbb{P} = 1.$$

Throughout the following, $\langle \cdot, \cdot \rangle$ denotes the inner product and $\| \cdot \|$ the norm in $\mathbb{R}^d$. For any set $\mathcal{A}$, we denote by $\overline{\mathcal{A}}$ the closure, by $\operatorname{int}\mathcal{A}$ the interior, and by $\partial\mathcal{A}$ the boundary of $\mathcal{A}$. The complement is simply written as $\mathbb{R}^d \setminus \mathcal{A}$. For any function $f : \mathbb{R}^d \to \mathbb{R}$, the gradient is denoted ∇f. We will also be using the convention that $\inf \emptyset = \infty$. Any other notation will be defined in context.

2.1 The Large Deviation Principle

This section aims to present the large deviation principle and a few other important notions, as well as tie together properties of these new concepts into Lemma 2.16, a powerful result we will use later as part of the main proof in Section 2.4.

First we define the notion of lower semi-continuity, and also the concept of a rate function: a lower semi-continuous mapping to the non-negative extended real numbers.

Definition 2.3. A function $f : \mathbb{R}^d \to [-\infty, \infty]$ is called *lower semi-continuous* if any of the following equivalent conditions hold:

(I) For any sequence $x_n \to x \in \mathbb{R}^d$, it holds that $\liminf_{n\to\infty} f(x_n) \geq f(x)$.

(II) For any $x \in \mathbb{R}^d$, with $B_\varepsilon(x) := \{y \in \mathbb{R}^d : \|x-y\| < \varepsilon\}$, it holds that $\lim_{\varepsilon\to 0} \inf_{y\in B_\varepsilon(x)} f(y) = f(x)$.

(III) For any $c \in \mathbb{R}$, the set $\mathcal{L}_c(f) := \{x \in \mathbb{R}^d : f(x) \leq c\} = f^{-1}((-\infty, c])$ is closed. In other words, f has closed level sets.

Proofs of the equivalence of the above definitions are simple and can be found in Chapter IV.6.2 of [7].

Example 2.4. *It is worth noting that any continuous function is also lower semi-continuous. Other examples of lower semi-continuous functions are the function*

$$f(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ 1, & \text{if } x > 0, \end{cases}$$

and the indicator function $\mathbb{I}_{x\in A}(x)$, where A is an open set.

In the one-dimensional case, a function being lower semi-continuous intuitively means that in the case of a jump discontinuity, the value of the function at that point is less than or equal to the lowest of the one-sided limits.

An important property of lower semi-continuous functions is that they obtain their minimum over non-empty, compact sets.

Lemma 2.5. *If $f : \mathbb{R}^d \to [-\infty, \infty]$ is a lower semi-continuous function, and $\mathcal{K} \subset \mathbb{R}^d$ is a compact set, then there exists a point $x_0 \in \mathcal{K}$ such that $f(x_0) = \inf_{x\in K} f(x)$.*

Proof. Denote $m := \inf_{x\in\mathcal{K}} f(x)$. Then for any $n \geq 1$, by property 2.3 (III), the set $\mathcal{C}_n := \mathcal{K} \cap f^{-1}((-\infty, m + \frac{1}{n}])$ is closed, as an intersection of two closed sets. From here we notice that the collection $\{\mathcal{C}_n\}$ is nested, i.e. $\cdots \subset \mathcal{C}_2 \subset \mathcal{C}_1$. Because of property 2.3 (II), each $\mathcal{C}_n$ is non-empty, and as a closed subset of a compact set, $\mathcal{C}_n \subset \mathcal{K}$ is compact.

As nested sequences of non-empty, compact sets have non-empty intersection in $\mathbb{R}^d$, the set $\mathcal{C} := \cap_{n=1}^{\infty} \mathcal{C}_n$ must be non-empty. Then for all $x \in \mathcal{C}$, we have $f(x) \leq m$ and $x \in \mathcal{K}$, by the definition of $\mathcal{C}_n$. Since $\mathcal{C} \subset \mathcal{K}$ is non-empty, f achieves its minimum on $\mathcal{K}$. □

We can now define the concept of a rate function, and state the Large Deviation Principle, a fundamental concept of this thesis. The LDP attempts to characterise the behaviour of very unlikely events using a rate function, which describes the exponential rate at which the probabilities of these very unlikely events decrease.

Definition 2.6. A function $I : \mathbb{R}^d \to [0, \infty]$ is called a *rate function* if it is lower semi-continuous. If, in addition, I har compact level sets, it is called a *good rate function.*

Example 2.7. *To give two non-continuous examples of rate functions, we will use the ceiling function, denoted $\lceil \cdot \rceil$. The function $f(x) = \lceil x \rceil$ is lower semi-continuous, thus a rate function, but does not have bounded level sets, so it is not a good rate function. On the other hand, the function $g(x) = \lceil x^2 \rceil$ is lower semi-continuous and has bounded level sets, and is therefore a good rate function.*

Definition 2.8 (The Large Deviation Principle). A family of probability measures $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ satisfies the *large deviation principle* with rate function I if

(I) for any closed set $\mathcal{C} \subset \mathbb{R}^d$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}_n(\mathcal{C}) \leq - \inf_{x \in \mathcal{C}} I(x),$$

(II) for any open set $\mathcal{O} \subset \mathbb{R}^d$,

$$\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}_n(\mathcal{O}) \geq - \inf_{x \in \mathcal{O}} I(x).$$

Now, one might wonder why a rate function is required to be lower semi-continuous. For one, it is a looser condition than true continuity, but it is also sufficient to ensure the uniqueness of the rate function.

Theorem 2.9. *If the family $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function I, then I is unique.*

Proof. Fix any $x \in \mathbb{R}^d$, and for $\varepsilon > 0$ denote $B_\varepsilon(x) := \{y \in \mathbb{R}^d : \|x - y\| < \varepsilon\}$. Then

$$\begin{aligned}
-I(x) &\leq - \inf_{y \in B_\varepsilon(x)} I(y) \\
&\overset{\text{LDP}}{\leq} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(B_\varepsilon(x)) \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(B_\varepsilon(x)) \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\overline{B}_\varepsilon(x)) \\
&\overset{\text{LDP}}{\leq} - \inf_{y \in \overline{B}_\varepsilon(x)} J(y) \\
&\leq - \inf_{y \in B_{2\varepsilon}(x)} J(y).
\end{aligned}$$

Then letting $\varepsilon \to 0$ we get $\inf_{y \in B_{2\varepsilon}(x)} J(y) \to J(x)$, since J is lower semi-continuous. This yields $-I(x) \leq -J(x)$ for all $x \in \mathbb{R}^d$, since x was arbitrary. Also, since I and J were arbitrarily ordered in the above proof, swapping them yields the opposite inequality $-J(x) \leq -I(x)$ for all $x \in \mathbb{R}^d$, meaning $I \equiv J$. □

In addition to the regular concept of tightness of measures, the intuition behind which is that we want to prevent any of the probability mass from 'escaping to infinity', there exists a stronger version of this concept, often used in large deviations theory, known as exponential tightness. It will be very useful later, in Section 2.4.x

Definition 2.10. A family of probability measures $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ on $\mathbb{R}^d$ is *exponentially tight* if for any $M < \infty$ there exists some corresponding compact set $\mathcal{K}_M \subset \mathbb{R}^d$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathbb{R}^d \setminus \mathcal{K}_M) < -M. \tag{2.11}$$

Example 2.12. *For a simple case of exponential tightness on $\mathbb{R}$, take the mean of n independent normally distributed random variables ξ_j with zero mean and unit variance. Denote this mean by S_n and note that*

$$S_n = \frac{1}{n} \sum_{j=1}^{n} \xi_j \sim \mathcal{N}\left(0, \frac{1}{n}\right).$$

Let $\mathbb{P}_n$ be the law of the variable S_n. Then for any $t \geq 0$ and for any $a \in \mathbb{R}$, Chebychev's inequality (2.2) gives us

$$\mathbb{P}(S_n \geq a) \leq e^{-at} \mathbb{E}\left(e^{tS_n}\right) \quad \textit{and} \quad \mathbb{P}(S_n \leq -a) \leq e^{-at} \mathbb{E}\left(e^{tS_n}\right). \tag{2.13}$$

These inequalities hold for any n and any $t \geq 0$. Recall that the moment generating function of a normally distributed random variable Y with mean μ and variance σ^2 is $\mathbb{E}\left(e^{tY}\right) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$. Now let $\varepsilon > 0$ be small and note that, when choosing $t = n$ in (2.13),

$$\mathbb{P}(|S_n| > a) \leq \mathbb{P}(S_n \geq a - \varepsilon) + \mathbb{P}(S_n \leq -a + \varepsilon) \leq 2e^{-n(a-\varepsilon)+\frac{n^2}{2n}}.$$

Then

$$\begin{aligned}\frac{1}{n}\log \mathbb{P}(|S_n| > a) &\leq \frac{1}{n}\log\left(2e^{-n(a-\varepsilon)+\frac{n}{2}}\right)\\ &= \frac{1}{n}\left(\log 2 - na + n\varepsilon + \frac{n}{2}\right)\\ &= \frac{\log 2}{n} - a + \varepsilon + \frac{1}{2}.\end{aligned}$$

Letting $n \to \infty$ and $\varepsilon \to 0$ yields

$$\frac{1}{n}\log \mathbb{P}(|S_n| > a) \leq \frac{1}{2} - a.$$

Finally, for any $M < \infty$, choose $a > M + \frac{1}{2}$, and then we have the compact set $\mathcal{K}_M := [-a, a]$ which satisfies Definition 2.10.

Here we present a very short and simple lemma, which will be used in the proof of Lemma 2.16, as well as in the proof of the main theorem of this chapter, specifically in Section 2.4. It will also be used in the subsequent chapter, which is why we prove the result in the continuous case.

Lemma 2.14. *If $K < \infty$ is a fixed integer, and for each $i = 1, \ldots, K$ and each $\rho > 0$, the value of a_ρ^i is non-negative, then*

$$\limsup_{\rho\to\infty} \frac{1}{\rho}\log\left(\sum_{i=1}^{K} a_\rho^i\right) = \max_{i=1,\ldots,K}\limsup_{\rho\to\infty}\frac{1}{\rho}\log a_\rho^i.$$

Proof. For any $\rho > 0$, we have

$$\max_{i=1,\ldots,K} a_\rho^i \leq \sum_{i=1}^{K} a_\rho^i \leq K \max_{i=1,\ldots,K} a_\rho^i,$$

thus since the logarithm is strictly increasing, we have

$$0 \leq \log\left(\sum_{i=1}^{K} a_\rho^i\right) - \log\left(\max_{i=1,\ldots,K} a_\rho^i\right) \quad \text{and} \quad \log\left(\sum_{i=1}^{K} a_\rho^i\right) - \log\left(\max_{i=1,\ldots,K} a_\rho^i\right) \leq \log K,$$

thus

$$0 \le \frac{1}{\rho} \log \left(\sum_{i=1}^{K} a_\rho^i \right) - \max_{i=1,\dots,K} \frac{1}{\rho} \log a_\rho^i \le \frac{1}{\rho} \log K.$$

Then since K is finite we can switch the order of maximum and limes supremum, so that

$$\begin{aligned}
&\limsup_{\rho\to\infty} \left(\frac{1}{\rho} \log \left(\sum_{i=1}^{K} a_\rho^i \right) - \max_{i=1,\dots,K} \frac{1}{\rho} \log a_\rho^i \right) \\
&= \limsup_{\rho\to\infty} \frac{1}{\rho} \log \left(\sum_{i=1}^{K} a_\rho^i \right) - \max_{i=1,\dots,K} \limsup_{\rho\to\infty} \frac{1}{\rho} \log a_\rho^i.
\end{aligned}$$

Finally, since this value is sandwiched between 0 and $\limsup_{\rho\to\infty} \frac{1}{\rho} \log K = 0$, its value, by the sandwich theorem, must be zero, thus we have our result,

$$\limsup_{\rho\to\infty} \frac{1}{\rho} \log \left(\sum_{i=1}^{K} a_\rho^i \right) = \max_{i=1,\dots,K} \limsup_{\rho\to\infty} \frac{1}{\rho} \log a_\rho^i.$$

□

In addition to the LDP, stated in Definition 2.8, there also exists a Weak Large Deviation Principle, in which the upper bound holds only for compact sets. We will later use this, in combination with the subsequent Lemma 2.16, as a convenient stepping stone in our proof of Theorem 2.29.

Definition 2.15. A family of probability measures $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ satisfies the *weak large deviation principle* with rate function I if

(I) for any compact set $\mathcal{C} \subset \mathbb{R}^d$, $\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathcal{C}) \le -\inf_{x\in\mathcal{C}} I(x)$,

(II) for any open set $\mathcal{O} \subset \mathbb{R}^d$, $\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathcal{O}) \ge -\inf_{x\in\mathcal{O}} I(x)$.

The following lemma is the main result of this section, and will be paramount in proving Theorem 2.29 later, as it means we can separately prove exponential tightness of the probability measures and the upper bound in Definition 2.15 for compact sets, thus implying the upper bound in Definition 2.8 for closed sets. Additionally, proving the lower bound then immediately implies the goodness of the rate function.

Lemma 2.16. *Let $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ be an exponentially tight family of probability measures.*

(I) *If the upper bound 2.15* (I) *holds for all compact sets for $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ with rate function I, then it also holds for all closed sets.*

(II) *If the lower bound 2.8* (II) *holds for* $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ *with rate function* I*, then* I *is a good rate function.*

Proof. (I) Fix a closed set $\mathcal{C}$, and choose $M < \infty$ such that $\mathcal{C} \subset (\mathbb{R}^d \setminus \mathcal{L}_M)$, where $\mathcal{L}_M$ denotes the level set

$$\mathcal{L}_M = \{x \in \mathbb{R}^d : I(x) \leq M\}.$$

This means all $x \in \mathcal{C}$ satisfy $I(x) > M$. As the family $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ is exponentially tight, this M has a corresponding compact set $\mathcal{K}_M$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathbb{R}^d \setminus \mathcal{K}_M) < -M.$$

Now clearly $\mathcal{C} \cap \mathcal{K}_M$ is compact, so the upper bound 2.15 (I) holds for this set. Also, since $\mathcal{C} \cap \mathcal{K}_M \subset \mathcal{C} \subset (\mathbb{R}^d \setminus \mathcal{L}_M)$, we have $\inf_{x\in\mathcal{C}\cap\mathcal{K}_M} I(x) \geq M$. The final thing to notice is that since $\mathcal{C} \subset (\mathcal{C} \cap \mathcal{K}_M) \cup (\mathbb{R}^d \setminus \mathcal{K}_M)$, we have, for all $n \geq 1$, the estimation

$$\mathbb{P}_n(\mathcal{C}) \leq \mathbb{P}_n(\mathcal{C} \cap \mathcal{K}_M) + \mathbb{P}_n(\mathbb{R}^d \setminus \mathcal{K}_M).$$

Now we apply Lemma 2.14, getting

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathcal{C}) &\leq \limsup_{n\to\infty} \frac{1}{n} \log(\mathbb{P}_n(\mathcal{C} \cap \mathcal{K}_M) + \mathbb{P}_n(\mathbb{R}^d \setminus \mathcal{K}_M)) \\
&= \max\Big\{ \underbrace{\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathcal{C} \cap \mathcal{K}_M)}_{\leq -\inf_{x\in\mathcal{C}\cap\mathcal{K}_M} I(x) \leq -M}, \underbrace{\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathbb{R}^d \setminus \mathcal{K}_M)}_{<-M} \Big\} \\
&\leq -M.
\end{aligned}$$

Then because

$$\inf\{-M : \mathcal{C} \subset (\mathbb{R}^d \setminus \mathcal{L}_M)\} = -\sup\{M < \infty : \forall x \in \mathcal{C}, I(x) > M\} = -\inf_{x\in\mathcal{C}} I(x),$$

We have

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathcal{C}) \leq -\inf_{x\in\mathcal{C}} I(x),$$

so the upper bound holds for all closed sets as well.

(II) Let $M < \infty$. Then as the family $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ is exponentially tight, there exists a compact set $\mathcal{K}_M$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathbb{R}^d \setminus \mathcal{K}_M) < -M.$$

As $\mathcal{K}_M$ is compact, it is closed, thus $\mathbb{R}^d \setminus \mathcal{K}_M$ is open, so the lower bound 2.8 (II) gives us

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathbb{R}^d \setminus \mathcal{K}_M) \geq - \inf_{x\in\mathbb{R}^d\setminus\mathcal{K}_M} I(x),$$

thus $M < \inf_{x\in\mathbb{R}^d\setminus\mathcal{K}_M} I(x)$. This means $I(x) > M$ for all $x \in \mathbb{R}^d\setminus\mathcal{K}_M$, so $(\mathbb{R}^d\setminus\mathcal{K}_M) \subset (\mathbb{R}^d \setminus \mathcal{L}_M)$. This can be rewritten as $\mathcal{L}_M \subset \mathcal{K}_M$, thus since $\mathcal{L}_M$ is closed, it is also compact. Since $M < \infty$ was arbitrary, I is a good rate function.

□

2.2 The Gärtner-Ellis theorem

A central concept in the proof of Theorem 2.29 is that of the limiting scaled cumulant generating function of a sequence of random variables, and the convex conjugate of this function. Consider random variables $X_n \in \mathbb{R}^d$, distributed according to the law $\mathbb{P}_n$, respectively, where $n \in \mathbb{N}$. The cumulant generating function $\Lambda_n : \mathbb{R}^d \to \mathbb{R}$ of the scaled variable nX_n is defined as

$$\Lambda_n(\lambda) := \log \mathbb{E}\left(e^{\langle\lambda, nX_n\rangle}\right).$$

By then scaling this function appropriately, and then taking the limit, we end up with what we call the *limiting scaled cumulant generating function* of the sequence $\{X_n\}_{n\in\mathbb{N}}$

$$\Lambda(\lambda) := \lim_{n\to\infty} \frac{1}{n}\Lambda_n(\lambda). \tag{2.17}$$

Example 2.18. *In the case of Example 2.11, we have*

$$X_n = S_n = \frac{1}{n}\sum_{j=1}^{n} \xi_j \sim \mathcal{N}\left(0, \frac{1}{n}\right),$$

and the moment generating function of the normal distribution again gives us

$$\mathbb{E}\left(e^{n\lambda X_n}\right) = e^{\frac{n^2\lambda^2}{2n}} = e^{\frac{n\lambda^2}{2}},$$

so

$$\Lambda(\lambda) = \lim_{n\to\infty} \frac{1}{n} \log \Lambda_n(\lambda) = \lim_{n\to\infty} \frac{1}{n} \log e^{\frac{n\lambda^2}{2}} = \frac{\lambda^2}{2}.$$

Where the Gärtner-Ellis theorem differs from Cramér's theorem is in its assumptions of dependence. Whereas Cramér's theorem works with a sum of i.i.d. random variables, the Gärtner-Ellis theorem imposes a looser condition. There can be some form of weak dependence between the variables, and the setting can be extended to not only include the distributions of sums of random variables, but in fact any sequence of probability distributions, as long as they abide by the following assumption.

Assumption 2.19. For every $\lambda \in \mathbb{R}^d$, the limit (2.17) exists as an extended real number. In addition, $0 \in \operatorname{int} \mathcal{D}_\Lambda$, where

$$\mathcal{D}_\Lambda := \{\lambda \in \mathbb{R}^d : \Lambda(\lambda) < \infty\}$$

denotes the *effective domain* of Λ.

Remark 2.20. A little bit later Lemma 2.31 tells us in case the limit (2.17) exists, Λ is a convex function. This implies $\mathcal{D}_\Lambda$ must be an interval, because if this was not the case, there must exist real numbers $a < b < c$ such that $\Lambda(a) < \infty$ and $\Lambda(c) < \infty$ but $\Lambda(b) = \infty$, which would violate convexity.

It is worth noting that although this assumption does not require independence, it is still a stringent condition. It only allows for very weak dependence in the sequence of random variables. In what follows of Chapter 2 we will be working under Assumption 2.19. With this in mind, we also need to introduce the following notion.

Definition 2.21. Denote by $\Lambda^* : \mathbb{R}^d \to [-\infty, \infty]$ the convex conjugate of Λ, defined as

$$\Lambda^*(x) := \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, x \rangle - \Lambda(\lambda)\}.$$

And similarly as above, $\mathcal{D}_{\Lambda^*} := \{\lambda \in \mathbb{R}^d : \Lambda^*(\lambda) < \infty\}$.

The reasoning behind introducing the function Λ and its convex conjugate is that Λ^* will act as the rate function with which the family of probability measures may satisfy the LDP.

Example 2.22. *In the case of Example 2.11, we have $\Lambda(\lambda) = \frac{\lambda^2}{2} \in (-\infty, \infty)$ everywhere, so $\operatorname{int} \mathcal{D}_\Lambda = \mathbb{R}$, and also, since $\lambda x - \frac{\lambda^2}{2}$ is a parabola with its maximum at $\lambda = x$, we have*

$$\Lambda^*(x) := \sup_{\lambda \in \mathbb{R}} \left\{ \lambda x - \frac{\lambda^2}{2} \right\} = \frac{x^2}{2}.$$

We still need to define a handful of concepts before we are ready to state our main theorem.

Definition 2.23. The point $y \in \mathbb{R}^d$ is called an *exposed point* of the function Λ^* if there exists some $\lambda \in \mathbb{R}^d$ such that for all $x \neq y$,

$$\langle \lambda, x \rangle - \Lambda^*(x) < \langle \lambda, y \rangle - \Lambda^*(y).$$

In the above, the point λ is called the *exposing hyperplane* of the point y.

Remark 2.24. Intuitively, a function f has an exposed point y if f is convex, but also strictly convex in some neighbourhood around y.

Example 2.25. *In the case of Example 2.12, we have* $\Lambda^*(x) = \frac{x^2}{2}$*, so every point* y *in* $\mathbb{R}$ *is exposed, with exposing hyperplane* $\lambda = \frac{\mathrm{d}}{\mathrm{d}y}\Lambda^*(y) = 2y$.

Definition 2.26. Let $\Lambda : \mathbb{R}^d \to (-\infty, \infty]$ be a convex function. Then Λ is called *essentially smooth* if the following hold:

(I) $\operatorname{int} \mathcal{D}_\Lambda \neq \emptyset$, i.e. Λ is finite in a neighbourhood of the origin.

(II) Λ is differentiable on $\operatorname{int} \mathcal{D}_\Lambda$.

(III) If $\{\lambda_i\} \in \operatorname{int} \mathcal{D}_\Lambda$ is a sequence converging to a point $\lambda \in \partial\mathcal{D}_\Lambda$, then $\lim_{n\to\infty} |\nabla\Lambda(\lambda_n)| = \infty$. When Λ has this property, we say Λ is *steep.*

Remark 2.27. Essential smoothness eliminates the possible case of $\Lambda = \pm\infty$ outside the origin, and, in addition to differentiability throughout the effective domain, imposes a condition on the behaviour of a function at its boundary, known as steepness. When approaching the boundary, the absolute value of the derivative (in general, the norm of the gradient) must approach $\pm\infty$, so that the function does not 'abruptly change direction towards infinity' at the edge of its effective domain.

Example 2.28. *In the case of example* 2.12*, the effective domain is the whole* $\mathbb{R}$*, so the boundary points are* $\{-\infty, \infty\}$*. Since* $\lim_{\lambda\to\infty} \frac{\lambda^2}{2} = \lim_{\lambda\to-\infty} \frac{\lambda^2}{2} = \infty$*,* Λ *is essentially smooth. Another function that is essentially smooth is*

$$f(x) = \begin{cases} \cos^{-1}(x), & \textit{if } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \\ \infty, & \textit{otherwise.} \end{cases}$$

The interior of the effective domain is $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$*, and* f *clearly is differentiable on this interval. Since*

$$\lim_{x\to\pi^-} \frac{\mathrm{d}}{\mathrm{d}x}\left|\frac{1}{\cos(x)}\right| = \lim_{x\to\pi^-}\left|\frac{\sin(x)}{\cos^2(x)}\right| = \infty, \quad \textit{and} \quad \lim_{x\to-\pi^+} \frac{\mathrm{d}}{\mathrm{d}x}\left|\frac{1}{\cos(x)}\right| = \lim_{x\to-\pi^+}\left|\frac{\sin(x)}{\cos^2(x)}\right| = \infty,$$

we also have the steepness of f.

The convex function

$$g(x) = \begin{cases} x^2, & \textit{if } x \in (-1,1), \\ \infty, & \textit{otherwise.} \end{cases}$$

is not essentially smooth, because it violates the steepness condition.

We are now ready to state our main result, a proof of which will be presented in Section 2.4.

Theorem 2.29 (Gärtner-Ellis). *Under Assumption 2.19, the following holds:*

(I) *For any closed set $\mathcal{C}$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathcal{C}) \leq -\inf_{x\in\mathcal{C}} \Lambda^*(x).$$

(II) *For any open set $\mathcal{O}$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathcal{O}) \geq -\inf_{x\in\mathcal{O}\cap\mathcal{E}} \Lambda^*(x),$$

where $\mathcal{E}$ denotes the set of exposed points of Λ^*, whose exposing hyperplane belong to $\operatorname{int} \mathcal{D}_\Lambda$

(III) *If Λ is essentially smooth and lower semi-continuous, then Λ^* is a good rate function, with which the Large Deviation Principle holds for the family $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$.*

The power of this theorem lies in that it can tell us when the LDP applies in more general settings than the one of Cramér's theorem. It also gives us an even more general lower bound, which can be turned into the lower bound in the LDP by removing the intersection with the set $\mathcal{E}$ from the infimum in the lower bound, which is allowed if Λ is essentially smooth.

2.3 A few results from convex analysis

This section presents a few important properties of the previously defined functions Λ and Λ^*. These will be of great use in our proof of Theorem 2.29, which constitutes the entirety of the subsection following this one.

Lemma 2.30. *The function Λ is convex.*

Proof. Let $n \geq 1$ and $\theta \in [0,1]$. Then Hölder's inequality (2.1) gives

$$\mathbb{E}\left((e^{\langle \lambda_1, nX_n\rangle})^{\theta}(e^{\langle \lambda_2, nX_n\rangle})^{(1-\theta)}\right) \leq \mathbb{E}\left(e^{\langle \lambda_1, nX_n\rangle}\right)^{\theta}\mathbb{E}\left(e^{\langle \lambda_2, nX_n\rangle}\right)^{(1-\theta)}.$$

so

$$\begin{aligned}
\Lambda_n(\theta\lambda_1 + (1-\theta)\lambda_2) &= \log \mathbb{E}\left(e^{\langle(\theta\lambda_1+(1-\theta)\lambda_2), nX_n\rangle}\right) \\
&= \log \mathbb{E}\left((e^{\langle \lambda_1, nX_n\rangle})^{\theta}(e^{\langle \lambda_2, nX_n\rangle})^{(1-\theta)}\right) \\
&\leq \log\left(\mathbb{E}\left(e^{\langle \lambda_1, nX_n\rangle}\right)^{\theta}\mathbb{E}\left(e^{\langle \lambda_2, nX_n\rangle}\right)^{(1-\theta)}\right) \\
&= \theta \log \mathbb{E}\left(e^{\langle \lambda_1, nX_n\rangle}\right) + (1-\theta)\log \mathbb{E}\left(e^{\langle \lambda_2, nX_n\rangle}\right) \\
&= \theta\Lambda_n(\lambda_1) + (1-\theta)\Lambda_n(\lambda_2),
\end{aligned}$$

thus Λ_n is convex for every $n \geq 1$. Since λ_1 and λ_2 i the above were arbitrary, it also holds that

$$\underbrace{\frac{1}{n}\Lambda_n(\theta\lambda_1 + (1-\theta)\lambda_2)}_{\to\Lambda(\theta\lambda_1+(1-\theta)\lambda_2)} \leq \underbrace{\theta\frac{1}{n}\Lambda_n(\lambda_1) + (1-\theta)\frac{1}{n}\Lambda_n(\lambda_2)}_{\to\theta\Lambda(\lambda_1)+(1-\theta)\Lambda(\lambda_2)},$$

so the limit Λ is also convex. □

Lemma 2.31. *Under Assumption 2.19, the following holds:*

(I) $\Lambda(\lambda) > -\infty$ *for all* $\lambda \in \mathbb{R}^d$.

(II) Λ^* *is a convex, good rate function.*

Proof. We prove the statements in order.

(I) First note that $\Lambda(0) = 0$. We proceed with a proof by contradiction. Suppose instead there is some $\lambda \in \mathbb{R}^d$ with $\Lambda(\lambda) = -\infty$. Since Λ is convex, as we just proved, we must, for all $0 < \theta \leq 1$, have

$$\Lambda(\theta\lambda) = \Lambda(\theta\lambda + (1-\theta)\cdot 0) \leq \theta\underbrace{\Lambda(\lambda)}_{=-\infty} + (1-\theta)\underbrace{\Lambda(0)}_{=0} = -\infty.$$

Convexity also implies, using $\theta = \frac{1}{2}$,

$$0 = \Lambda(0) = \Lambda\left(\frac{1}{2}(-\lambda+\lambda)\right) \leq \frac{1}{2}\Lambda(-\lambda) + \frac{1}{2}\Lambda(\lambda).$$

This means

$$-\Lambda(-\lambda) \leq \Lambda(\lambda) = -\infty$$

so we must have $\Lambda(-\lambda) = \infty$, which in turn implies, for all $0 < \theta \leq 1$,

$$\Lambda(-\theta\lambda) = \Lambda(\theta(-\lambda) + (1-\theta)\cdot 0) \leq \theta \underbrace{\Lambda(-\lambda)}_{=\infty} + (1-\theta)\underbrace{\Lambda(0)}_{=0} = \infty.$$

But Assumption 2.19 states that the interior of $\mathcal{D}_\Lambda$ is non-empty, which contradicts what we have just arrived at. Therefore there cannot exist any λ for which $\Lambda(\lambda) = -\infty$.

(II) From the definition of Λ^* follows that for any $x_1, x_2 \in \mathbb{R}^d$ and $\theta \in [0,1]$,

$$\begin{aligned}
\Lambda^*(\theta x_1 + (1-\theta)x_2) &= \sup_{\lambda\in\mathbb{R}^d}\{(\theta x_1 + (1-\theta)x_2)\lambda - \Lambda(\lambda)\} \\
&= \sup_{\lambda\in\mathbb{R}^d}\{\theta(\lambda x_1 - \Lambda(\lambda)) + (1-\theta)(\lambda x_2 - \Lambda(\lambda))\} \\
&\leq \theta \sup_{\lambda\in\mathbb{R}^d}\{\lambda x_1 - \Lambda(\lambda)\} + (1-\theta)\sup_{\lambda\in\mathbb{R}^d}\{\lambda x_2 - \Lambda(\lambda)\} \\
&= \theta\Lambda^*(x_1) + (1-\theta)\Lambda^*(x_2),
\end{aligned}$$

thus Λ^* is convex.

Note that by definition, $\Lambda^*(x) \geq 0\cdot x - \Lambda(0) = 0$, so Λ^* is non-negative. Now to prove that it is a good rate function, it suffices to show that it is lower semi-continuous and has bounded level sets. In order to show that it satisfies Definition 2.3 (I), let $x_n \to x$ be a sequence in $\mathbb{R}^d$. By the definition of Λ^* we have, for all $\lambda \in \mathbb{R}^d$,

$$\liminf_{n\to\infty} \Lambda^*(x_n) \geq \liminf_{n\to\infty}(\lambda x_n - \Lambda(\lambda)) = \lambda x - \Lambda(\lambda),$$

thus, taking the supremum over all $\lambda \in \mathbb{R}^d$,

$$\liminf_{n\to\infty} \Lambda^*(x_n) \geq \sup_{\lambda\in\mathbb{R}^d}\{\lambda x - \Lambda(\lambda)\} = \Lambda^*(x),$$

so Λ^* is lower semi-continuous. In order to show that it has bounded level sets, first note that the interior is always an open set, thus Λ is continuous in $\operatorname{int}\mathcal{D}_\Lambda$. Then recall that Assumption 2.19 ensures we can find small enough $\delta > 0$ such that the closed ball $\overline{B}_\delta(0) \subset \operatorname{int}\mathcal{D}_\Lambda$. By the continuity of Λ inside this ball we have

$\sup_{\lambda \in \overline{B}_\delta(0)} \Lambda(\lambda) < \infty$. Now we can estimate

$$\begin{aligned}
\Lambda^*(x) &= \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, x\rangle - \Lambda(\lambda)\} \\
&\geq \sup_{\lambda \in \overline{B}_\delta(0)} \{\langle \lambda, x\rangle - \Lambda(\lambda)\} \\
&\geq \sup_{\lambda \in \overline{B}_\delta(0)} \langle \lambda, x\rangle - \sup_{\lambda \in \overline{B}_\delta(0)} \Lambda(\lambda), \\
&= \delta|x| - \underbrace{\sup_{\lambda \in \overline{B}_\delta(0)} \Lambda(\lambda)}_{<\infty} \\
&\to \infty, \text{ as } x \to \pm\infty.
\end{aligned}$$

therefore the level sets $\mathcal{L}_c = \{x \in \mathbb{R}^d : \Lambda^*(x) \leq c\}$ are bounded for all $c \in \mathbb{R}$, so Λ^* has compact level sets because of the equivalence of Definitions 2.3 (I) and 2.3 (III). All together this means Λ^* is a convex, good rate function, as we wanted.

□

Some times we may only care about a function's values on some lower-dimensional subspace of the entire space, and want to remove the boundary points from some set with respect to this subspace without losing the entire set. In these kinds of situations a useful concept is that of the relative interior.

Definition 2.32. Let $\mathcal{A} \subset \mathbb{R}^d$ be a convex set. The *relative interior* of $\mathcal{A}$ is the subset defined as

$$\operatorname{ri} \mathcal{A} := \{x \in \mathcal{A} : \text{or all } a \in \mathcal{A} \text{ there is some } \varepsilon > 0 \text{ such that } x - \varepsilon(x - a) \in \mathcal{A}\}.$$

Example 2.33. *Take as an example the unit line segment in two dimensions* $A := \{(x, 0) \in \mathbb{R}^2 : 0 \leq x \leq 1\}$*. The interior in the traditional sense is* $\operatorname{int} A = \emptyset$*, but the relative interior instead gives us* $\operatorname{ri} A = \{(x, 0) \in \mathbb{R}^2 : 0 < x < 1\}$*.*

The following is a result from R. Tyrrell Rockafellar's book on convex analysis, and will be useful in proving 2.29 (III).

Lemma 2.34 (Rockafellar). *Let* $\Lambda : \mathbb{R}^d \to (-\infty, \infty]$ *be a convex function. If* Λ *is essentially smooth and lower semi-continuous, then* $\operatorname{ri} \mathcal{D}_{\Lambda^*} \subset \mathcal{E}$*.*

Proof. See Corollary 26.4.1 in [8]. □

2.4 A proof of Gärtner-Ellis

We now arrive at the proof of our main theorem. We will begin this proof by proving the upper bound in two steps, first showing that it holds for compact sets, and then showing that the family $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ is exponentially tight, two facts which, together with Lemma 2.16 (I) give us 2.29 (I).

Following this, we will define an exponentially tilted family of probability measures and, after verifying that the limiting scaled cumulant generating function that arises from this family also satisfies Assumption 2.19, use the already proven upper bound applied to this new family and its corresponding rate function, to show the lower bound 2.29 (II).

Finally, we use Lemma 2.34 to show that, under the conditions imposed in 2.29 (III), $\mathcal{O}\cap\mathcal{E}$ can be replaced by $\mathcal{O}$ in the infimum in the lower bound, which by 2.16 (II) implies Λ^* is a good rate function. This, in conjunction with 2.29 (I) and 2.29 (II), means the LDP holds for the family $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ with good rate function Λ^*.

The upper bound: compact sets

First and foremost we will define an auxiliary function. For any $\delta > 0$, the *δ-rate function* associated with Λ^* is

$$I^\delta(x) := \min\{\Lambda^*(x) - \delta, \frac{1}{\delta}\},$$

and note especially that $I^\delta(x) < \Lambda^*(x)$ for all $\delta > 0$, $x \in \mathbb{R}^d$, and that

$$\lim_{\delta\to 0}\inf_{x\in F} I^\delta(x) = \inf_{x\in F}\Lambda^*(x), \tag{2.35}$$

for any set $F \subset \mathbb{R}^d$.

The reason for introducing this function is that in proving the upper bound it is instrumental to work with a function whose range is bounded, while still having, in the limit, properties which agree with those of the rate function we care about.

Now let $\mathcal{K} \subset \mathbb{R}^d$ be a compact subset and choose $\delta > 0$ arbitrary. The definition of Λ^* ensures that, for any $x \in \mathcal{K}$, there must exist some corresponding $\lambda_x \in \mathbb{R}^d$ such that

$$I^\delta(x) \le \langle\lambda_x, x\rangle - \Lambda(\lambda_x),$$

since we can pick λ_x to make the expression $\langle\lambda_x, x\rangle - \Lambda(\lambda_x)$ arbitrarily close to its supremum $\Lambda^*(x)$. Since the scalar product is continuous, there exists some $\gamma > 0$ such that for any $y \in B_\gamma(x)$, we have

$$|\langle\lambda_x, y\rangle - \langle\lambda_x, x\rangle| < \delta,$$

and more specifically,

$$\inf_{y\in B_\gamma(x)}\{\langle\lambda_x, y\rangle - \langle\lambda_x, x\rangle\} \ge -\delta. \tag{2.36}$$

Chebychev's inequality (2.2) and (2.36) now give us, for all $t \geq 0$,

$$\begin{aligned}\mathbb{P}_n(B_\gamma(x)) &\leq \mathbb{P}(\langle \lambda_x, X_n\rangle - \langle \lambda_x, x\rangle \geq -\delta) \\ &\leq e^{\delta t}\mathbb{E}\left(e^{t(\langle \lambda_x, X_n\rangle - \langle \lambda_x, x\rangle)}\right).\end{aligned}$$

Choosing $t = n$ and noting that $n\langle \lambda_x, X_n\rangle = \langle \lambda_x, nX_n\rangle$, we have, by rearranging the above,

$$\mathbb{P}_n(B_\gamma(x) \leq e^{n(\delta - \langle \lambda_x, x\rangle)}\mathbb{E}\left(e^{\langle \lambda_x, nX_n\rangle}\right).$$

Thus

$$\frac{1}{n}\log \mathbb{P}_n(B_\gamma(x)) \leq \delta - \langle \lambda_x, x\rangle + \frac{1}{n}\Lambda_n(\lambda_x).$$

Now, as the set $\mathcal{K}$ is compact with a trivial cover $\bigcup_{x\in\mathcal{K}} B_\gamma(x)$, for any $\gamma > 0$, we can find some finite subcover, which we will denote $\bigcup_{i=1}^N B_{\gamma_i}(x_i)$. Now we can estimate

$$\mathbb{P}_n(\mathcal{K}) \leq \mathbb{P}_n\left(\bigcup_{i=1}^N B_{\gamma_i}(x_i)\right) \leq \sum_{i=1}^N \mathbb{P}_n\left(B_{\gamma_i}(x_i)\right) \leq N \max_{i=1,\ldots,N}\{\mathbb{P}_n\left(B_{\gamma_i}(x_i)\right)\},$$

thus

$$\begin{aligned}\frac{1}{n}\log \mathbb{P}_n(\mathcal{K}) &\leq \frac{1}{n}\log\left(N \max_{i=1,\ldots,N}\{\mathbb{P}_n\left(B_{\gamma_i}(x_i)\right)\}\right) \\ &= \frac{1}{n}\log N + \max_{i=1,\ldots,N}\left\{\frac{1}{n}\log \mathbb{P}_n\left(B_{\gamma_i}(x_i)\right)\right\} \\ &\leq \frac{1}{n}\log N + \delta + \max_{i=1,\ldots,N}\left\{\frac{1}{n}\Lambda_n(\lambda_{x_i}) - \langle \lambda_{x_i}, x_i\rangle\right\}.\end{aligned}$$

Now by taking the limes supremum we get

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\log \mathbb{P}_n(\mathcal{K}) &\leq \delta + \limsup_{n\to\infty}\left(\frac{1}{n}\log N + \max_{i=1,\ldots,N}\left\{\frac{1}{n}\Lambda_n(\lambda_{x_i}) - \langle \lambda_{x_i}, x_i\rangle\right\}\right) \\ &= \delta + \max_{i=1,\ldots,N}\left\{\limsup_{n\to\infty}\frac{1}{n}\Lambda_n(\lambda_{x_i}) - \langle \lambda_{x_i}, x_i\rangle\right\} \\ &= \delta - \min_{i=1,\ldots,N}\left\{\langle \lambda_{x_i}, x_i\rangle - \Lambda(\lambda_{x_i})\right\} \\ &\leq \delta - \min_{i=1,\ldots,N} I^\delta(x_i) \\ &\leq \delta - \inf_{x\in\mathcal{K}} I^\delta(x),\end{aligned}$$

where $\limsup_{n\to\infty} \frac{1}{n}\Lambda_n(\lambda_{x_i}) = \Lambda(\lambda_{x_i})$ because of Assumption 2.19 and the penultimate inequality is due to how we chose each λ_{x_i}. Letting $\delta \to 0$ then yields, by (2.35),

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(\mathcal{K}) \leq \lim_{\delta\to 0} \left(\delta - \inf_{x\in\mathcal{K}} I^\delta(x) \right) = -\inf_{x\in\mathcal{K}} \Lambda^*(x),$$

so we have the upper bound for compact sets.

The upper bound: exponential tightness

For the random variable $X_n = (X_n^1, \dots, X_n^d) \in \mathbb{R}^d$ following law $\mathbb{P}_n$, define for any $i = 1, \dots, d$ the law $\mathbb{P}_n^j$ such that for any set measurable set $A \subset \mathbb{R}$,

$$\mathbb{P}_n^i(A) = \mathbb{P}_n(X_n^i \in A) = \mathbb{P}_n(\underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{i-1} \times A \times \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{d-j}).$$

Now recall that Assumption 2.19 states $0 \in \operatorname{int} \mathcal{D}_\Lambda$, thus by 2.31 (I) we know Λ is of finite value around the origin. For $i = 1, \dots, d$, let $\mathbf{e}_i$ denote the i:th unit vector of the space $\mathbb{R}^d$, and let $\phi_i > 0, \varphi_i > 0$ be such that

$$\Lambda(-\varphi_i \mathbf{e}_i) < \infty, \text{ and } \Lambda(\phi_i \mathbf{e}_i) < \infty.$$

Let $\alpha > 0$. For all $i = 1, \dots, d$ and $n \geq 1$, the Chebychev's inequality (2.2) gives us two estimations, the first being

$$\begin{aligned}
\mathbb{P}_n^i((-\infty, -\alpha]) &\leq e^{-\alpha n \varphi_i} \mathbb{E}\left(e^{-n\varphi_i X_n^i}\right) \\
&= e^{-\alpha n \varphi_i} \mathbb{E}\left(e^{\langle -\varphi_i \mathbf{e}_i, nX_n \rangle}\right) \\
&= e^{-\alpha n \varphi_i} \exp\left(\log\left(\mathbb{E}\left(e^{\langle -\varphi_i \mathbf{e}_i, nX_n \rangle}\right)\right)\right) \\
&= \exp(-\alpha n \varphi_i + \Lambda_n(-\varphi_i \mathbf{e}_i)),
\end{aligned}$$

which turns into

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n^i((-\infty, -\alpha]) &\leq \limsup_{n\to\infty} \frac{1}{n}(-\alpha n \varphi_i + \Lambda_n(-\varphi_i \mathbf{e}_i)) \\
&= \Lambda(-\varphi_i \mathbf{e}_i) - \alpha \varphi_i \\
&\to -\infty, \quad \text{as } \alpha \to \infty.
\end{aligned}$$

as $\varphi_i > 0$ and $\Lambda(-\varphi_i \mathbf{e}_i) < \infty$. The equality above is due to Assumption 2.19.

The second estimation achieved from the Chebychev's inequality (2.2) is

$$\begin{aligned}\mathbb{P}_n^i([\alpha,\infty)) &\le e^{-\alpha n\phi_i}\mathbb{E}\left(e^{n\phi_i X_n^i}\right)\\ &= e^{-\alpha n\phi_i}\mathbb{E}\left(e^{\langle \phi_i \mathbf{e}_i, nX_n\rangle}\right)\\ &= e^{-\alpha n\phi_i}\exp\left(\log\left(\mathbb{E}\left(e^{\langle \phi_i \mathbf{e}_i, nX_n\rangle}\right)\right)\right)\\ &= \exp(-\alpha n\phi_i + \Lambda_n(\phi_i\mathbf{e}_i)),\end{aligned}$$

giving us the limit

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\log \mathbb{P}_n^i([\alpha,\infty)) &\le \limsup_{n\to\infty}\frac{1}{n}(-\alpha n\phi_i + \Lambda_n(\phi_i\mathbf{e}_i))\\ &= \Lambda(\phi_i\mathbf{e}_i) - \alpha\phi_i\\ &\to -\infty, \quad \text{as } \alpha\to\infty,\end{aligned}$$

as $\phi_i > 0$ and $\Lambda(\phi_i\mathbf{e}_i) < \infty$.

We now examine the event where X belongs to the complement of the compact set $[-\alpha,\alpha]^d$. This can also be written

$$X_n^i \in (-\infty,-\alpha]\cup[\alpha,\infty), \text{ for some } i = 1,\dots,d.$$

Then we have the union bound

$$\mathbb{P}_n(\mathbb{R}^d\setminus[-\alpha,\alpha]^d) \le \sum_{i=1}^d \left(\mathbb{P}_n^i((-\infty,-\alpha]) + \mathbb{P}_n^i([\alpha,\infty))\right).$$

Now, using Lemma 2.14 together with what we shower earlier, we find the limit

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}_n(\mathbb{R}^d\setminus[-\alpha,\alpha]^d)\\ &\le \limsup_{n\to\infty}\frac{1}{n}\log\sum_{i=1}^d\left(\mathbb{P}_n^i((-\infty,-\alpha]) + \mathbb{P}_n^i([\alpha,\infty))\right)\\ &\overset{(2.14)}{=} \max_{i=1,\dots,d}\left\{\max\left\{\underbrace{\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}_n^i((-\infty,-\alpha])}_{\to-\infty \text{ as } \alpha\to\infty,\ \forall i=1,\dots,d}, \underbrace{\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}_n^i([\alpha,\infty))}_{\to-\infty \text{ as } \alpha\to\infty,\ \forall i=1,\dots,d}\right\}\right\}\\ &\to -\infty, \quad \text{as } \alpha\to\infty.\end{aligned}$$

Thus, for any $M > 0$ there is some compact set $[-\alpha_M,\alpha_M]^d =: \mathcal{K}_M$ that satisfies the inequality 2.11 in Definition 2.10, and we have proved exponential tightness. Then 2.16 (I) implies the upper bound 2.29 (I).

The lower bound: exponential tilt of measure

Denote by $\mathcal{E}$ the set of all exposed points of Λ^* whose exposing hyperplane belongs to $\operatorname{int}\mathcal{D}_\Lambda$. Then for any open set $\mathcal{O}, x \in \mathcal{O}\cap\mathcal{E}$ and $\delta > 0$ small enough, we have

$$\mathbb{P}_n(\mathcal{O}) \geq \mathbb{P}_n(B_\delta(x)),$$

for all $n \geq 1$, which means to prove the lower bound 2.29 (II) it suffices to show that

$$\lim_{\delta\to 0}\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}_n(B_\delta(x)) \geq -\Lambda^*(x), \quad \text{for all } x \in \mathcal{E}. \tag{2.37}$$

First fix $x \in \mathcal{E}$, and let $\kappa \in \operatorname{int}\mathcal{D}_\Lambda$ be an exposing hyperplane of the point x. Then by Assumption 2.19, we know $\Lambda(\kappa)$ exists, so for this κ and all n large enough, $\Lambda_n(\kappa) < \infty$. We can define an exponentially tilted probability measure $\widehat{\mathbb{P}}_n$ by the measure densities

$$\mathrm{d}\widehat{\mathbb{P}}_n(y) = \exp\left(n\langle\kappa, y\rangle - \Lambda_n(\kappa)\right)\mathrm{d}\mathbb{P}_n(y), \quad \text{for all } y \in \mathbb{R}^d.$$

Now for n large enough that $\Lambda_n(\kappa) < \infty$, we can calculate

$$\begin{aligned}
\frac{1}{n}\log\mathbb{P}_n(B_\delta(x)) &= \frac{1}{n}\log\int_{B_\delta(x)}\mathrm{d}\mathbb{P}_n(y)\\
&= \frac{1}{n}\log\int_{B_\delta(x)}\exp\left(\Lambda_n(\kappa) - n\langle\kappa, y\rangle\right)\mathrm{d}\widehat{\mathbb{P}}_n(y)\\
&= \frac{1}{n}\log\left(\exp\left(\Lambda_n(\kappa) - n\langle\kappa, x\rangle\right)\int_{B_\delta(x)} e^{-n\langle\kappa, y-x\rangle}\,\mathrm{d}\widehat{\mathbb{P}}_n(y)\right)\\
&= \frac{1}{n}\Lambda_n(\kappa) - \langle\kappa, x\rangle + \frac{1}{n}\log\int_{B_\delta(x)} e^{-n\langle\kappa, y-x\rangle}\,\mathrm{d}\widehat{\mathbb{P}}_n(y)\\
&\geq \frac{1}{n}\Lambda_n(\kappa) - \langle\kappa, x\rangle + \frac{1}{n}\log\left(e^{-n\delta|\kappa|}\int_{B_\delta(x)}\mathrm{d}\widehat{\mathbb{P}}_n(y)\right)\\
&= \frac{1}{n}\Lambda_n(\kappa) - \langle\kappa, x\rangle - \delta|\kappa| + \frac{1}{n}\log\int_{B_\delta(x)}\mathrm{d}\widehat{\mathbb{P}}_n(y)\\
&= \frac{1}{n}\Lambda_n(\kappa) - \langle\kappa, x\rangle - \delta|\kappa| + \frac{1}{n}\log\widehat{\mathbb{P}}_n(B_\delta(x)).
\end{aligned} \tag{2.38}$$

Now notice that

$$\lim_{\delta\to 0}\liminf_{n\to\infty}\frac{1}{n}\Lambda_n(\kappa) - \langle\kappa, x\rangle - \delta|\kappa| \to \Lambda(\kappa) - \langle\kappa, x\rangle \geq -\Lambda^*(x),$$

by the definition of Λ^*. Thus, when taking the limes infimum as $n \to \infty$ and then the limit as $\delta \to 0$, we can further estimate (2.38) as

$$\lim_{\delta\to 0}\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}_n(B_\delta(x)) \geq -\Lambda^*(x) + \lim_{\delta\to 0}\liminf_{n\to\infty}\frac{1}{n}\log\widehat{\mathbb{P}}_n(B_\delta(x)). \tag{2.39}$$

Now if we can manage to show that

$$\lim_{\delta\to 0}\liminf_{n\to\infty}\frac{1}{n}\log\widehat{\mathbb{P}}_n(B_\delta(x))=0, \tag{2.40}$$

then 2.39 implies 2.37, and we have the lower bound 2.29 (II).

The lower bound: using the upper bound

Let $\widehat{X}_n$ be the random variable whose distribution law is $\widehat{\mathbb{P}}_n$ and denote by $\widehat{\Lambda}_n$ the cumulant generating function of the scaled variable $n\widehat{X}_n$. In order to show 2.40, we will show that $\widehat{\Lambda}$ also satisfies Assumption 2.19, and use the recently proven upper bound. To this end, notice that

$$\begin{aligned}\widehat{\Lambda}_n(\lambda)&=\log\int_\Omega e^{\langle\lambda,ny\rangle}\,\mathrm{d}\widehat{\mathbb{P}}_n(y)\\&=\log\int_\Omega e^{\langle\lambda,ny\rangle+n\langle\kappa,y\rangle-\Lambda_n(\kappa)}\,\mathrm{d}\mathbb{P}_n(y)\\&=\log\int_\Omega e^{\langle\lambda+\kappa,ny\rangle}\,\mathrm{d}\mathbb{P}_n(y)-\log e^{-\Lambda_n(\kappa)}\\&=\Lambda_n(\lambda+\kappa)-\Lambda_n(\kappa),\end{aligned}$$

which gives us the existence of the limit

$$\widehat{\Lambda}(\lambda):=\lim_{n\to\infty}\frac{1}{n}\widehat{\Lambda}_n(\lambda)=\Lambda(\lambda+\kappa)-\Lambda(\kappa)\in[-\infty,\infty]. \tag{2.41}$$

Now Λ has a defined finite value at the point κ and in some neighbourhood around it, since $\kappa\in\operatorname{int}\mathcal{D}_\Lambda$. Then $\widehat{\Lambda}$ also is finite in a neighbourhood around 0, meaning $0\in\operatorname{int}\mathcal{D}_{\widehat{\Lambda}}$. Thus $\widehat{\Lambda}$ satisfies Assumption 2.19.

Now in 2.40, the set $B_\delta(x)$ is open, so we apply the upper bound 2.29 (I) to its complement. This yields

$$\limsup_{n\to\infty}\frac{1}{n}\log\widehat{\mathbb{P}}_n(\mathbb{R}^d\setminus B_\delta(x))\le-\inf_{y\in\mathbb{R}^d\setminus B_\delta(x)}\widehat{\Lambda}^*(y).$$

Now Lemma 2.31 (II) says $\widehat{\Lambda}^*$ is a good rate function, meaning it is lower semi-continuous and has compact level sets. By Lemma 2.5, this implies that $\widehat{\Lambda}^*$ achieves its infimum over closed sets. Then there must exist some $y_0\in\mathbb{R}^d\setminus B_\delta(x)$ for which

$$\inf_{y\in\mathbb{R}^d\setminus B_\delta(x)}\widehat{\Lambda}^*(y)=\widehat{\Lambda}^*(y_0).$$

From the definition of Λ^* we have $\Lambda(\kappa) \geq \langle \kappa, x\rangle - \Lambda^*(x)$ for all $\kappa \in \mathbb{R}^d$, and since x is an exposed point of Λ^* with exposing hyperplane κ, by Definition 2.23, we also have the inequality

$$\langle \kappa, x\rangle - \langle \kappa, y_0\rangle - \Lambda^*(x) + \Lambda^*(y_0) > 0.$$

Putting all this together and recalling (2.41) gives

$$\begin{aligned}
\widehat{\Lambda}^*(y_0) &= \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, y_0\rangle - (\Lambda(\lambda+\kappa) - \Lambda(\kappa))\} \\
&= \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda+\kappa, y_0\rangle - \Lambda(\lambda+\kappa)\} - \langle \kappa, y_0\rangle + \Lambda(\kappa) \\
&= \Lambda^*(y_0) - \langle \kappa, y_0\rangle + \Lambda(\kappa) \\
&\geq \Lambda^*(y_0) - \langle \kappa, y_0\rangle + \langle \kappa, x\rangle - \Lambda^*(x) > 0.
\end{aligned}$$

Then for all $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \widehat{\mathbb{P}}_n(\mathbb{R}^d \setminus B_\delta(x)) \leq - \inf_{y \in \mathbb{R}^d \setminus B_\delta(x)} \widehat{\Lambda}^*(y) = -\widehat{\Lambda}^*(y_0) < 0.$$

This means that for any fixed $\delta > 0$ we must have $\log \widehat{\mathbb{P}}_n(\mathbb{R}^d \setminus B_\delta(x)) \to -\infty$, and therefore also $\widehat{\mathbb{P}}_n(B_\delta(x)) \to 1$ as $n \to \infty$, and so we get

$$\lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \widehat{\mathbb{P}}_n(B_\delta(x)) = 0,$$

which is 2.40, as we wanted, meaning 2.39 implies 2.37, giving us the lower bound 2.29 (II).

The case in which the LDP holds

Employing Lemma 2.34, we now see that under the conditions imposed on Λ in 2.29 (III), we have, for any open set $\mathcal{O}$,

$$\inf_{x \in \mathcal{O}\cap\mathcal{E}} \Lambda^*(x) \leq \inf_{x \in \mathcal{O}\cap \operatorname{ri} \mathcal{D}_{\Lambda^*}} \Lambda^*(x),$$

meaning that, in light of the already proven upper and lower bounds, it is enough to show that

$$\inf_{x \in \mathcal{O}\cap \operatorname{ri} \mathcal{D}_{\Lambda^*}} \Lambda^*(x) \leq \inf_{x\in\mathcal{O}} \Lambda^*(x), \tag{2.42}$$

as this implies the lower bound 2.8 (II) of the Large Deviation Principle. After this the goodness of the rate function Λ^* follows from Lemma 2.16 (II), as we have already proven the exponential tightness of the family $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$.

If $\mathcal{O} \cap \mathcal{D}_{\Lambda^*} = \emptyset$, then both sides of the inequality (2.42) evaluate to ∞, and there is nothing more to prove. We assume instead that $\mathcal{D}_{\Lambda^*}$ is non-empty. This means we can choose some arbitrary $a \in \operatorname{ri} \mathcal{D}_{\Lambda^*}$ and $b \in \mathcal{O} \cap \mathcal{D}_{\Lambda^*}$. Now for small enough $\theta > 0$ it holds for the intermediary point that

$$\theta a + (1-\theta)b \in \mathcal{O} \cap \operatorname{ri} \mathcal{D}_{\Lambda^*}.$$

Since Λ^* is convex on its effective domain $\mathcal{D}_{\Lambda^*}$, it is continuous on $\operatorname{ri} \mathcal{D}_{\Lambda^*}$, in the sense that if $x \in \operatorname{ri} \mathcal{D}_{\Lambda^*}$ and $x_n \to x$ is a sequence with and $\{x_n\}_{n \in \mathbb{N}} \in \operatorname{ri} \mathcal{D}_{\Lambda^*}$, then $f(x_n) \to f(x)$. With this realisation we arrive at the estimation

$$\begin{aligned}
\inf_{x \in \mathcal{O} \cap \operatorname{ri} \mathcal{D}_{\Lambda^*}} \Lambda^*(x) &\leq \lim_{\theta \to 0} \Lambda^*(\theta a + (1-\theta)b) \\
&\leq \lim_{\theta \to 0} \theta \Lambda^*(a) + \lim_{\theta \to 0} (1-\theta) \Lambda^*(b) \\
&= \Lambda^*(b),
\end{aligned}$$

where the second inequality is due to the convexity of Λ^*. This, since a and especially b were arbitrarily chosen, implies

$$\inf_{x \in \mathcal{O} \cap \operatorname{ri} \mathcal{D}_{\Lambda^*}} \Lambda^*(x) \leq \inf_{x \in \mathcal{O} \cap \mathcal{D}_{\Lambda^*}} \Lambda^*(x) = \inf_{x \in \mathcal{O}} \Lambda^*(x),$$

where the equality is due to the definition of $\mathcal{D}_{\Lambda^*}$. This is what we wanted to prove. □

3 Large deviations results for the time of ruin

Whereas Cramér's theorem concerns empirical means of i.i.d. random variables, the Gärtner-Ellis theorem, presented in the previous chapter, concerns a sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ as a generalisation of such a mean. In this chapter we will instead be shifting our perspective to a sequence $\{Y_n\}_{n\in\mathbb{N}}$ which is a generalisation not of an empirical mean, but of a random walk (one might liken Y_n to nX_n). Also, instead of a discrete family of probability measures $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$, we now consider the continuous case $\{\mathbb{P}_M\}_{M>0}$.

We will be presenting some large deviations results developed by Nyrhinen in [9, 10], pertaining to the probabilities $\mathbb{P}_M(A) = \mathbb{P}(T/M \in A)$, where $T = T(M) = \inf\{n \in \mathbb{N} : Y_n > M\}$ is the time of ruin when the starting capital is $M > 0$. These results are stated here as Theorems 3.11 and 3.14.

A main stepping stone in the proofs of these results is Theorem 3.24, which can be thought of as a continuous version of the Gärtner-Ellis Theorem, with some modifications.

3.1 Necessary preliminaries and the two main theorems

Let $\{Y_n\}_{n\in\mathbb{N}}$ be a sequence of random variables distributed according to laws $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$, respectively, and let $M > 0$. Denote the time of ruin by $T = T_M := \inf\{n \in \mathbb{N} : Y_n > M\}$. Here we say $T = \infty$ if $Y_n \le M$ for all $n \in \mathbb{N}$. Denote the cumulant generating function of Y_n by

$$c_n(t) = \log \mathbb{E}\left(e^{tY_n}\right) \in (-\infty, \infty].$$

Also define the *upper limiting scaled cumulant generating function* of Y_n by

$$c(t) := \limsup_{n\to\infty} \frac{1}{n} c_n(t) \in [-\infty, \infty], \tag{3.1}$$

and denote

$$w = \sup\{t \in \mathbb{R} : c(t) \le 0\} \in [0, \infty]. \tag{3.2}$$

Remark 3.3. The functions c_n and c are very similar to Λ_n and Λ in the previous chapter, the main difference being the absence of scaling of the variable Y_n (which is really just a shift in the behaviour of the random variables we are considering) and the relaxed assumption of (3.1) holding as a true limit. We use this notation in this chapter to mirror Nyrhinen's own, and to avoid confusing definitions with those of the previous chapter.

Similarly to Lemma 2.30, it can be shown that c is convex, meaning there is some largest open interval $J = (a_1, a_2)$ for which c is finite and strictly increasing on J. Suppose J is non-empty. Then $w \in [a_1, a_2]$. Now denote the right and left limits

$$b_1 = \lim_{t \to a_1^+} c(t),$$

and

$$b_2 = \lim_{t \to a_2^-} c(t).$$

Now the restriction of c to the open interval J has an inverse, denoted c^{-1}. Using this, define the function $\Gamma : \mathbb{R} \to (-\infty, \infty]$ as

$$\Gamma(t) = \begin{cases} -a_2, & \text{if } t \in (-\infty, -b_2] \\ -c^{-1}(-t), & \text{if } t \in (-b_2, -b_1) \\ -a_1, & \text{if } t = -b_1 \\ \infty, & \text{if } t \in (-b_1, \infty). \end{cases} \tag{3.4}$$

Next, we will present some assumptions we will be working under in what follows. Our proof of Theorem 3.11 requires only the first three, while all of them are used in our proof of Theorem 3.14. The reasoning behind these conditions is explained in Section 4 of [9], and in Section 3 of [10].

Assumption 3.5. The following hold:

(I) The open interval J is non-empty.

(II) For all $n \in \mathbb{N}$ and $t \in J$, we have $c_n(t) < \infty$.

(III) If $w > 0$, then $a_1 < w$.

(IV) For all $t \in J$, (3.1) holds as a limit.

(V) The derivative $c'(t)$ exists for all $t \in J$.

(VI) If $a_2 \in (-\infty, \infty)$, we have $\lim_{t \to a_2^-} c'(t) = \infty$.

(VII) If $a_1 \in (-\infty, \infty)$, we have $\lim_{t \to a_1^+} c'(t) = 0$.

We still need to define a few further concepts, and some characteristics relating to them. For any $t, u \in \mathbb{R}$, and $\vartheta > 0$, denote the function

$$C_M(t, u; \vartheta) = \frac{1}{M} \log \mathbb{E}\left(e^{tY_{\lceil \vartheta M \rceil} + uY_{\lceil M \rceil}}\right). \tag{3.6}$$

Then for each $\tau \geq 0$ and $\delta > 0$ denote

$$H_\delta(t,u;\tau) = \limsup_{M\to\infty} \sup_{\vartheta\in B_\delta(\tau)\cap(0,\infty)} C_M(t,u;\vartheta), \tag{3.7}$$

and let

$$H(t,u;\tau) = \lim_{\delta\to 0} H_\delta(t,u;\tau). \tag{3.8}$$

Now denote the right partial derivative of H with respect to t at the point 0 by the function

$$p(u;\tau) = \lim_{h\to 0^+} \frac{H(h,u;\tau) - H(0,u;\tau)}{h}. \tag{3.9}$$

Finally, write

$$\tilde{J} = \{u \in J : c'(u) \in \operatorname{int} c'(J)\}.$$

Remark 3.10. Let $u \in J$ and fix $\tau \geq 0$, $\vartheta > 0$. Because $H_\delta(t,u;\tau)$ is monotone in δ for any $t \in \mathbb{R}$, the limit (3.8) must exist as an extended real number. In addition, $C_M(t,u;\vartheta)$ is convex in t, because for any two points $t_1, t_2 \in \mathbb{R}$ and $\theta \in [0,1]$, we have

$$\begin{aligned}
C_M(\theta t_1 + (1-\theta)t_2, u, \vartheta) &= \frac{1}{M}\log\mathbb{E}\left(e^{uY_{\lceil M\rceil}+(\theta t_1+(1-\theta)t_2)Y_{\lceil\vartheta M\rceil}}\right)\\
&= \frac{1}{M}\log\mathbb{E}\left(e^{\theta\left(uY_{\lceil M\rceil}+t_1Y_{\lceil\vartheta M\rceil}\right)+(1-\theta)\left(uY_{\lceil M\rceil}+t_2Y_{\lceil\vartheta M\rceil}\right)}\right)\\
&= \frac{1}{M}\log\mathbb{E}\left(\left(e^{uY_{\lceil M\rceil}+t_1Y_{\lceil\vartheta M\rceil}}\right)^{\theta}\left(e^{uY_{\lceil M\rceil}+t_2Y_{\lceil\vartheta M\rceil}}\right)^{(1-\theta)}\right)\\
&\overset{(2.1)}{\leq} \frac{1}{M}\log\left(\mathbb{E}\left(e^{uY_{\lceil M\rceil}+t_1Y_{\lceil\vartheta M\rceil}}\right)^{\theta}\mathbb{E}\left(e^{uY_{\lceil M\rceil}+t_2Y_{\lceil\vartheta M\rceil}}\right)^{(1-\theta)}\right)\\
&= \theta\frac{1}{M}\log\mathbb{E}\left(e^{uY_{\lceil M\rceil}+t_1Y_{\lceil\vartheta M\rceil}}\right) + (1-\theta)\frac{1}{M}\log\mathbb{E}\left(e^{uY_{\lceil M\rceil}+t_2Y_{\lceil\vartheta M\rceil}}\right)\\
&= \theta C_M(t_1,u;\vartheta) + (1-\theta)C_M(t_2,u;\vartheta),
\end{aligned}$$

by use of Hölder's inequality. Therefore also $H_\delta(t,u;\tau)$ and $H(t,u;\tau)$ are convex in t. Also, notice that

$$\begin{aligned}
H_\delta(0,u;\tau) &= \limsup_{M\to\infty} \sup_{\vartheta\in B_\delta(\tau)\cap(0,\infty)} \frac{1}{M}\log\mathbb{E}\left(e^{0\cdot Y_{\lceil\vartheta M\rceil}+uY_{\lceil M\rceil}}\right)\\
&= \limsup_{M\to\infty}\frac{1}{M}\log\mathbb{E}\left(e^{uY_{\lceil M\rceil}}\right) = c(u) \in (-\infty,\infty).
\end{aligned}$$

We can now state the two main results in this chapter.

Theorem 3.11. *Suppose assumptions 3.5* (I)*-3.5* (III) *hold. Then for every* $t \in \mathbb{R} \setminus \{-b_1\}$ *we have*

$$\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{E}\left(e^{tT}\mathbb{I}\{T<\infty\}\right) \le \Gamma(t). \tag{3.12}$$

Also, for every closed set $\mathcal{C} \subset \mathbb{R}$ *we have*

$$\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{C}\right) \le -\inf_{x\in\mathcal{C}} \Gamma^*(x). \tag{3.13}$$

Theorem 3.14. *Suppose assumptions 3.5* (I)*-3.5* (VII) *hold and suppose* $p(u;\tau) < c'(u)$ *for every* $u \in \tilde{J}$ *and* $0 \le \tau < 1$*. Then for every* $t \in \mathbb{R}$ *we have*

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{E}\left(e^{tT}\mathbb{I}\{T<\infty\}\right) \ge \Gamma(t). \tag{3.15}$$

Also, for every open set $\mathcal{O} \subset \mathbb{R}$ *we have*

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{O}\right) \ge -\inf_{x\in\mathcal{O}} \Gamma^*(x). \tag{3.16}$$

Theorem 3.11 is one of two main results in [9], and implies the upper bound of the LDP for the family of probabilities $\{\mathbb{P}_M\}_{M>0}$ under relatively weak conditions on the behaviour of the sequence $\{Y_n\}_{n\in\mathbb{N}}$. Theorem 3.14 on the other hand, is the main result in [10], and is a strengthening of another result in [9]. It implies the lower bound of the LDP under some additional conditions on the behaviour of the sequence $\{Y_n\}_{n\in\mathbb{N}}$. Theorems 3.11 and 3.14 together imply the LDP for the family $\{\mathbb{P}_M\}_{M>0}$, and are proved hereunder in Sections 3.3 and 3.4. For some justification of the assumptions regarding the derivative $p(u,\tau)$, the reader is referred to Section 3 of [10].

3.2 Auxiliary concepts and results

Many of the functions we will be working with will turn out to be convex, and it can therefore be of great help to turn to the field of convex analysis for some useful results. We will be using several concepts and results from R. T. Rockafellar's book on the subject [8], including the notion of proper convexity, and that of the lower semi-continuous hull. More background surrounding these, and on convex analysis in general, can be found in the aforementioned book.

Definition 3.17. A convex function f is called *proper convex*, if $f(x) > -\infty$ for all x, and $f(x) < \infty$ for some x.

Definition 3.18. Let f be a proper convex function. The *lower semi-continuous hull* of f is defined as

$$f_L(x) = \liminf_{y \to x} f(y).$$

Remark 3.19. If the function f in the above definition is lower semi-continuous, then $f_L \equiv f$.

The following three results are stated and proved in [8]. They will be very useful in the proofs of our main results, and also of the soon to be presented Lemma 3.25.

Theorem 3.20. *Let f be a proper convex function. Then the convex conjugate f^* is lower semi-continuous and proper convex. Additionally, $(f_L)^* \equiv f^*$, and $(f^*)^* \equiv f_L$.*

Proof. See Theorem 12.2 in [8]. □

Corollary 3.21. *If f is a convex function on $\mathbb{R}$, then $f^*(x) = \sup_{t \in \operatorname{int} \mathcal{D}_f} \{xt - f(t)\}$.*

Proof. See Corollary 12.2.2 in [8]. □

Theorem 3.22. *If f is lower semi-continuous and proper convex, then* $\inf f = -f^*(0)$.

Proof. See Theorem 27.1 in [8]. □

We have now introduced everything necessary from convex analysis. The following lemma is simple but very important, as it enables the final steps of the proof of Theorem 3.14 in Section 3.4.

Lemma 3.23. *For each $\rho > 0$, suppose that $0 \leq b_\rho \leq a_\rho$. Suppose also that there is some $c \in \mathbb{R}$ such that*

$$\limsup_{\rho \to \infty} \frac{1}{\rho} \log b_\rho < c \leq \liminf_{\rho \to \infty} \frac{1}{\rho} \log a_\rho.$$

Then

$$\liminf_{\rho \to \infty} \frac{1}{\rho} \log (a_\rho - b_\rho) \geq c.$$

Proof. Let $\varepsilon > 0$ be small enough that $\limsup_{\rho \to \infty} \frac{1}{\rho} \log b_\rho < c - 2\varepsilon$. Then we can find a suitably large $P > 0$ so that when $\rho > P$, it holds that

$$a_\rho > e^{\rho(c-\varepsilon)} \quad \text{and} \quad b_\rho < e^{\rho(c-2\varepsilon)}.$$

Then by our choice of ε, we have $e^{\rho(c-\varepsilon)} > e^{\rho(c-2\varepsilon)}$, thus we can write

$$\begin{aligned}
\frac{1}{\rho}\log\left(a_\rho - b_\rho\right) &> \frac{1}{\rho}\log\left(e^{\rho(c-\varepsilon)} - e^{\rho(c-2\varepsilon)}\right) \\
&= \frac{1}{\rho}\log e^{\rho(c-\varepsilon)}\left(1 - e^{-\rho\varepsilon}\right) \\
&= c - \varepsilon + \underbrace{\frac{1}{\rho}\log\left(1 - e^{-\rho\varepsilon}\right)}_{\to 0,\text{ as } \rho\to\infty}.
\end{aligned}$$

Therefore

$$\liminf_{\rho\to\infty}\frac{1}{\rho}\log\left(a_\rho - b_\rho\right) \geq c - \varepsilon,$$

and the proof is completed by letting $\varepsilon \to 0$. □

The following theorem can be seen as a continuous version of Theorem 2.29, and although it is stated a bit differently, its proof is in many ways extremely similar. As our current case concerns the random variables $\frac{T}{M}$ and the corresponding probability measures $\mathbb{P}_M$, which are indexed by the real number $M > 0$, we cannot use Theorem 2.29, so we will be needing this one instead. It will be used in several places in the proofs in Sections 3.3 and 3.4.

Theorem 3.24. *For each $\rho > 0$, let Z_ρ be a random variable on the probability space $(\Omega, \Sigma, \mathbb{Q}_\rho)$. For each $t \in \mathbb{R}$, denote*

$$\bar{\Lambda}(t) = \limsup_{\rho\to\infty}\frac{1}{\rho}\log\mathbb{E}\left(e^{tZ_\rho}\right).$$

Then the following hold:

(I) *For each compact set $\mathcal{K} \subset \mathbb{R}$, we have*

$$\limsup_{\rho\to\infty}\frac{1}{\rho}\log\mathbb{Q}_\rho\left(\frac{Z_\rho}{\rho} \in \mathcal{K}\right) \leq -\inf_{x\in\mathcal{K}}\bar{\Lambda}^*(x),$$

and if $0 \in \operatorname{int}\mathcal{D}_{\bar{\Lambda}}$, the above also holds for all closed sets.

(II) *If $t_0 \in \mathbb{R}$ is such that*

$$\bar{\Lambda}(t) = \lim_{\rho\to\infty}\frac{1}{\rho}\log\mathbb{E}\left(e^{tZ_\rho}\right) \in (-\infty, \infty),$$

for every t in some neighbourhood of t_0, and if $\bar{\Lambda}$ is differentiable at t_0, then for any open set $\mathcal{O} \ni \bar{\Lambda}'(t_0)$,

$$\liminf_{\rho\to\infty} \frac{1}{\rho} \log \mathbb{Q}_\rho \left(\frac{Z_\rho}{\rho} \in \mathcal{O} \right) \geq \bar{\Lambda}(t_0) - t_0\bar{\Lambda}'(t_0).$$

Proof. See Theorem 4.5.3 in [2] for the proof of 3.24 (I) for compact sets. For the extension to closed sets, see Theorem 4.5.20 in [2] noting that the assumption $0 \in \operatorname{int} \mathcal{D}_{\bar{\Lambda}}$ implies exponential tightness of the probability measures $\mathbb{Q}_\rho$. The proof of this fact in the continuous case is largely identical to Part 2 of Section 2.4 in this thesis.

By Lemma 2.3.9 in [2], the existence of $\bar{\Lambda}'(t_0)$ implies that it is an exposed point of $\bar{\Lambda}^*$, with exposing hyperplane t_0. Noting that

$$\bar{\Lambda}^*(\bar{\Lambda}'(t_0)) = \sup_{t\in\mathbb{R}} \left\{ t\bar{\Lambda}'(t_0) - \bar{\Lambda}(t) \right\} = t_0\bar{\Lambda}'(t_0) - \bar{\Lambda}(t_0),$$

since $\bar{\Lambda}'(t_0)t - \bar{\Lambda}(t)$ is maximised when $t = t_0$, the proof of 3.24 (II) then follows from Theorem 4.5.20 in [2]. $\square$

Finally, before we begin with our main proofs, it will be to our benefit to clarify some of the properties of the function Γ and its convex conjugate Γ^*.

Lemma 3.25. *Suppose assumption 3.5* (I) *holds. Then Γ is lower semi-continuous and proper convex, and its convex conjugate is*

$$\Gamma^*(x) = \begin{cases} \infty, & \text{if } x \in (-\infty, 0) \\ \lim_{y\to 0^+}, yc^*\left(\frac{1}{y}\right) & \text{if } x = 0 \\ xc^*\left(\frac{1}{x}\right), & \text{if } x \in (0, \infty). \end{cases} \tag{3.26}$$

Proof. As c is finite, convex and strictly increasing on (a_1, a_2), its inverse c^{-1} is finite, concave and strictly increasing on (b_1, b_2), meaning Γ is finite, convex and strictly increasing on $(-b_2, -b_1)$. Then by its definition (3.4), Γ is convex on $\mathbb{R}$. Noticing that a_2 is infinite if and only if b_2 is infinite yields the proper convexity of Γ. Also, by finiteness and convexity, Γ is also continuous on $(-b_2, -b_1)$. The definitions of a_1, a_2, b_1 and b_2 imply that Γ is lower semi-continuous. As for the convex conjugate Γ^*, we have three cases.

(I) First let $x \in (0, \infty)$. Notice that Γ being constant on $(-\infty, -b_2]$ implies that $xt - \Gamma(t)$ is a increasing function of t on this interval. Then as $\Gamma(t) = \infty$ on $(-b_1, \infty)$ we can write

$$\Gamma^*(x) = \sup_{t\in\mathbb{R}} \{xt - \Gamma(t)\} = \sup_{t\in[-b_2,-b_1]} \{xt - \Gamma(t)\}.$$

Now define the function

$$\bar{\Gamma}(t) = \begin{cases} \infty, & \text{if } t \in (-\infty, -b_2) \\ \Gamma(t), & \text{if } t \in [-b_2, \infty), \end{cases}$$

and notice that

$$(\bar{\Gamma})^*(x) = \sup_{t\in\mathbb{R}}\{xt - \bar{\Gamma}(t)\} = \sup_{t\in[-b_2,-b_1]}\{xt - \bar{\Gamma}(t)\} = \sup_{t\in[-b_2,-b_1]}\{xt - \Gamma(t)\} = \Gamma^*(x).$$

In fact, by Corollary 3.21, the supremum above can be taken over the open interval, so

$$\Gamma^*(x) = \bar{\Gamma}^*(x) = \sup_{t\in(-b_2,-b_1)}\{xt - \bar{\Gamma}(x)\} = \sup_{t\in(-b_2,-b_1)}\{xt - \Gamma(x)\}.$$

Thus

$$\begin{aligned} \Gamma^*(x) &= \sup_{t\in(-b_2,-b_1)}\{xt - \Gamma(t)\} \\ &= \sup_{t\in(-b_2,-b_1)}\{xt + c^{-1}(-t)\} \\ &= \sup_{c(t)\in(b_1,b_2)}\{-xc(t) + t\} \\ &= \sup_{t\in J} x\left\{\frac{t}{x} - c(t)\right\}. \end{aligned}$$

Also, as J is the largest open interval on which c is finite and strictly increasing, we must have $c(t) = \infty$ for $t \in [a_2, \infty)$, if a_2 is finite. Also, c is convex but not strictly increasing on $(-\infty, a_1]$, meaning $xt - c(t)$ is an increasing function on this interval. Then

$$c^*(x) = \sup_{t\in\mathbb{R}}\{xt - c(t)\} = \sup_{t\in[a_1,a_2]}\{xt - c(t)\}.$$

Now, as above with Γ, we can construct a new function whose convex conjugate agrees with that of c, and whose effective domain is $[a_1, a_2]$, and, using Corollary 3.21 again, arrive at

$$c^*(x) = \sup_{t\in(a_1,a_2)}\{xt - c(t)\}.$$

Therefore

$$xc^*\left(\frac{1}{x}\right) = \sup_{t\in J} x\left\{\frac{t}{x} - c(t)\right\} = \Gamma^*(x),$$

which is the first case.

(II) Now let $x \in (-\infty, 0)$. Recall first the definition of w from (3.2). When J is non-empty, we have $w \in [a_1, a_2]$, meaning $0 = c(w) \in [b_1, b_2]$, and thus

$$\Gamma(0) = -c^{-1}(-c(w)) = -c^{-1}(c(w)) = -w \leq 0.$$

Then since Γ is increasing by (3.4), we have $\Gamma(t) \leq 0$ for every $t < 0$. Thus

$$\Gamma^*(x) = \sup_{t \in \mathbb{R}} \{xt - \Gamma(t)\} \geq \lim_{t \to -\infty} (xt - \Gamma(t)) = \infty,$$

and we have shown the second case.

(III) For the final case where $x = 0$, we turn to Theorem 3.20, which, by the previously noted proper convexity of Γ, implies Γ^* is proper convex and lower semi-continuous. Then as Γ^* is infinite to the left of the origin, the only possibility is

$$\Gamma^*(0) = \lim_{x \to 0^+} x c^* \left(\frac{1}{x} \right),$$

as any other choice would violate either convexity or lower semi-continuity. This is the final case.

□

With these results stated, we are now ready to move on to the proofs of our main results. We begin with the upper bound, i.e. Theorem 3.11, the proof of which constitutes the entirety of the following section.

3.3 Proof of Theorem 3.11

Throughout this entire section, we assume 3.5 (I)-3.5 (III). We begin with an observation. First denote $P := \mathbb{P}(T < \infty)$. We will show in Part 1 that $P > 0$. For each $M > 0$, choose $W_M = T\mathbb{I}\{T < \infty\}$, define the probability measures $\mathbb{Q}_M$ such that for all $\mathcal{A} \subset \mathbb{R}$,

$$\mathbb{Q}_M(\mathcal{A}) = \frac{1}{P} \mathbb{P}\left(\mathcal{A} \cap \{T < \infty\}\right).$$

Denote the functions

$$\bar{\Lambda}(t) = \limsup_{M \to \infty} \frac{1}{M} \log \int_\Omega e^{tW_M(\omega)} \, \mathrm{d}\mathbb{Q}_M(\omega),$$

and

$$\gamma(t) = \limsup_{M \to \infty} \frac{1}{M} \log \mathbb{E}\left(e^{tT} \mathbb{I}\{T < \infty\}\right).$$

Then by noticing that

$$P \int_\Omega e^{tW_M(\omega)}\, \mathrm{d}\mathbb{Q}_M(\omega) = \int_\Omega e^{tT(\omega)\mathbb{I}\{T(\omega)<\infty\}} \mathbb{I}\{T(\omega) < \infty\}\, \mathrm{d}\mathbb{P}(\omega) = \mathbb{E}\left(e^{tT}\mathbb{I}\{T < \infty\}\right),$$

we have

$$\begin{aligned}\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{E}\left(e^{tT}\mathbb{I}\{T<\infty\}\right) &= \limsup_{M\to\infty} \frac{1}{M} \log P \int_\Omega e^{tW_M(\omega)}\, \mathrm{d}\mathbb{Q}_M(\omega) \\ &= \underbrace{\limsup_{M\to\infty} \frac{\log P}{M}}_{=0} + \limsup_{M\to\infty} \frac{1}{M} \log \int_\Omega e^{tW_M(\omega)}\, \mathrm{d}\mathbb{Q}_M(\omega)\end{aligned}$$

and thus the identity

$$\bar{\Lambda} \equiv \gamma, \tag{3.27}$$

which will be very useful later.

The structure of this proof is as follows. In Part 1 we show that $P > 0$, which makes the above identity (3.27) meaningful. In Part 2 we show that $\gamma(t) \leq \Gamma(t)$ for every $t \in \mathbb{R} \setminus \{-b_1\}$, which is (3.12). In Part 3 we show that $\gamma^*(x) \geq \Gamma^*(x)$ for every $x \in \mathbb{R}$, and how this, in light of (3.27), together with Theorem 3.24 (I) implies (3.13) for compact sets. In Part 4 we extend (3.13) to closed sets.

Part 1

Let $\varepsilon > 0$. In order to prove that $P = \mathbb{P}(T < \infty) > 0$, let $t > 0$ be such that $\sup_{n\geq 1} \mathbb{E}\left(e^{tY_n}\right) = \infty$. With the intention of arriving at a contradiction, suppose it holds for some $a > 0$ that for all $n \geq 1$ and $x \geq a$ that

$$\mathbb{P}\left(e^{tY_n} > x\right) \leq x^{-(1+\varepsilon)}.$$

Let $N \geq 1$ and notice that

$$\begin{aligned}
\mathbb{E}\left(e^{tY_n}\mathbb{I}\{e^{tY_n} \leq N\}\right) &\leq \sum_{m=1}^{N} m\mathbb{P}\left(m-1 < e^{tY_n} \leq m\right) \\
&= \sum_{m=1}^{N} \left(\mathbb{P}\left(m-1 < e^{tY_n} \leq m\right) + \cdots + \mathbb{P}\left(N-1 < e^{tY_n} \leq N\right)\right) \\
&= \sum_{m=1}^{N} \mathbb{P}\left(m-1 < e^{tY_n} \leq N\right) \\
&= \sum_{m=0}^{N-1} \mathbb{P}\left(e^{tY_n} > m\right) - \underbrace{N\mathbb{P}\left(e^{tY_n} > N\right)}_{\geq 0} \\
&\leq \sum_{m=0}^{N-1} \mathbb{P}\left(e^{tY_n} > m\right) \\
&\leq \mathbb{P}\left(e^{tY_n} > 0\right) + \cdots + \mathbb{P}\left(e^{tY_n} > \lfloor a \rfloor\right) + \sum_{m=\lceil a \rceil + 1}^{\infty} \mathbb{P}\left(e^{tY_n} > m\right) \\
&\leq 1 + \lfloor a \rfloor + \underbrace{\sum_{m=\lceil a \rceil + 1}^{\infty} m^{-(1+\varepsilon)}}_{\text{converges}} < \infty,
\end{aligned}$$

even as $N \to \infty$, which is in contradiction to the assumption that $\sup_{n\geq 1} \mathbb{E}\left(e^{tY_n}\right) = \infty$. This means that for all $i \geq 1$ we can choose a sequence $a_i > 0$, with $a_i \to \infty$ as $i \to \infty$, for which we can find sequences $n_i \geq 1$ and $x_i \geq a$ such that $x_i \to \infty$ as $i \to \infty$, and for which it holds that

$$\mathbb{P}\left(e^{tY_n} > x_i\right) > x_i^{-(1+\varepsilon)}. \tag{3.28}$$

Now by writing $M_i := \frac{1}{t}\log x_i$ and noticing that $\mathbb{P}(T_{M_i} < \infty) \geq \mathbb{P}(Y_{n_i} > M_i)$ and that $M_i \to \infty$, as $i \to \infty$, we can use (3.28) to get

$$\begin{aligned}
\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}(T < \infty) &\geq \limsup_{i\to\infty} \frac{1}{M_i} \log \mathbb{P}(Y_{n_i} > M_i) \\
&= \limsup_{i\to\infty} \frac{1}{M_i} \log \mathbb{P}(e^{tY_{n_i}} > x_i) \\
&\geq \limsup_{i\to\infty} \frac{1}{M_i} \log(e^{-(1+\varepsilon)tM_i}) = -(1+\varepsilon)t.
\end{aligned}$$

Now letting $\varepsilon \to 0$ yields

$$\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}(T < \infty) \geq -t.$$

By assumption the interval J is non-empty, and c is convex and strictly increasing on J. Thus w must be finite. Also, noticing that $\sup_{n\geq 1} \mathbb{E}\left(e^{tY_n}\right) < \infty$ implies $c(t) \leq 0$ gives us

$$\sup\left\{t > 0 : \sup_{n\geq 1} \mathbb{E}\left(e^{tY_n}\right) < \infty\right\} \leq \sup\{t \in \mathbb{R} : c(t) \leq 0\} = w < \infty.$$

Then

$$\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}(T < \infty) > -\infty,$$

and since $\mathbb{P}(T_M < \infty)$ is decreasing in M, this means that for all $M > 0$,

$$\mathbb{P}(T_M < \infty) > 0.$$

Part 2

To prove (3.12) we show that

$$\gamma(t) \leq \Gamma(t) \tag{3.29}$$

for every $t \in \mathbb{R} \setminus \{-b_1\}$. First suppose $w > 0$. We proceed with three cases:

(I) For $t \in (-b_1, \infty)$, $\Gamma(t) = \infty$, so (3.29) holds without further investigation.

(II) For $t \in (-b_2, -b_1)$, let $\varepsilon > 0$ be small and $t_\varepsilon \in J \subset (0, \infty)$ for which $c(t_\varepsilon) = -t - \varepsilon$. Now c_n is finite throughout J for all n by assumption, and since for any $M > 0$,

$$\begin{aligned} e^{c_n(t_\varepsilon)} &= \exp\left(\log \mathbb{E}\left(e^{t_\varepsilon Y_n}\right)\right) \\ &\geq \mathbb{E}\left(e^{t_\varepsilon Y_n} \mathbb{I}\{Y_n > M\}\right) \\ &\geq e^{t_\varepsilon M} \mathbb{E}\left(\mathbb{I}\{Y_n > M\}\right) \\ &= e^{t_\varepsilon M} \mathbb{P}\left(Y_n > M\right), \end{aligned}$$

we have, by rearranging,

$$\mathbb{P}\left(Y_n > M\right) \leq e^{c_n(t_\varepsilon) - t_\varepsilon M}.$$

Then we can estimate

$$\begin{aligned}
\mathbb{E}\left(e^{tT}\mathbb{I}\{T<\infty\}\right) &= \sum_{n=1}^{\infty} e^{tn}\mathbb{P}\left(T=n\right) \\
&\le \sum_{n=1}^{\infty} e^{tn}\mathbb{P}\left(Y_n>M\right) \\
&\le \sum_{n=1}^{\infty} e^{tn+c_n(t_\varepsilon)-t_\varepsilon M} \\
&= e^{-t_\varepsilon M}\underbrace{\sum_{n=1}^{\infty} e^{n\left(t+\frac{1}{n}c_n(t_\varepsilon)\right)}}_{\text{converges, }=:\alpha<\infty} \\
&= \alpha e^{-t_\varepsilon M}.
\end{aligned}$$

Then

$$\gamma(t) \le \limsup_{M\to\infty} \frac{1}{M}\log \alpha e^{-t_\varepsilon M} = \limsup_{M\to\infty} \frac{\log\alpha}{M} - t_\varepsilon = -t_\varepsilon = -c^{-1}(-t-\varepsilon),$$

and letting $\varepsilon \to 0$ yields (3.29).

(III) Finally, if $b_2 < \infty$ and $t \le -b_2$, then as we just showed $\gamma(t) \le \Gamma(t)$ for $t \in (-b_2, -b_1)$ to get

$$\gamma(t) \le \limsup_{r\to -b_2^+} \gamma(t) \le \limsup_{r\to -b_2^+} \Gamma(t) = \Gamma(-b_2) = -a_2 = \Gamma(t)$$

because γ is increasing.

We have now shown (3.12) when $w > 0$. Next suppose $w = 0$. Then $c(t) > 0$ for any $t > 0$ by the definition of w. Also, since we assumed J is non-empty, either b_1 is negative, or b_2 is positive. This means we have five possible cases:

(I) For $t \in (-b_1, \infty)$, the case is trivial, just as when $w > 0$.

(II) For $t = 0$, the case is also clear, as as $\gamma(0) = \Gamma(0) = 0$.

(III) If $b_2 > 0$ and $t \in (-b_2, 0)$, then $c(c^{-1}(-t)) = -t \in (0, b_2)$. As $c(0) = 0$, we have $c^{-1}(-t) \in (0, a_2)$. Proceeding as in case (II) when $w > 0$, but choosing $\varepsilon = 0$ and $t_\varepsilon = t_0 = c^{-1}(-t)$, we get

$$\gamma(t) \le -t_0 = -c^{-1}(-t) = \Gamma(t).$$

(IV) If $b_1 < 0$ and $t \in (0, -b_1)$, then some part of J overlaps the interval $(-\infty, 0)$. Let $\varepsilon > 0$ and find some $t_\varepsilon \in J \cap (-\infty, 0)$ such that $c(t_\varepsilon) = -t - \varepsilon$. Now since t_ε is negative we get, for all $M > 0$,

$$\begin{aligned} e^{c_n(t_\varepsilon)} &= \exp\left(\log \mathbb{E}\left(e^{t_\varepsilon Y_n}\right)\right) \\ &\geq \mathbb{E}\left(e^{t_\varepsilon Y_n} \mathbb{I}\{Y_n \leq M\}\right) \\ &\geq e^{t_\varepsilon M} \mathbb{E}\left(\mathbb{I}\{Y_n \leq M\}\right) \\ &= e^{t_\varepsilon M} \mathbb{P}\left(Y_n \leq M\right), \end{aligned}$$

and by rearranging,

$$\mathbb{P}(Y_n \leq M) \leq e^{c_n(t_\varepsilon) - t_\varepsilon M}.$$

In addition, we notice that for each $n \geq 1$, it holds that

$$\mathbb{P}(T = n) \leq \mathbb{P}(Y_{n-1} \leq M).$$

We can now estimate

$$\begin{aligned} \mathbb{E}\left(e^{tT} \mathbb{I}\{T < \infty\}\right) &= \sum_{n=1}^{\infty} e^{tn} \mathbb{P}\left(T = n\right) \\ &\leq \sum_{n=1}^{\infty} e^{tn} \mathbb{P}\left(Y_{n-1} \leq M\right) \\ &\leq \sum_{n=1}^{\infty} e^{tn + c_{n-1}(t_\varepsilon) - t_\varepsilon M} \\ &= e^{-t_\varepsilon M} \underbrace{\sum_{n=1}^{\infty} e^t e^{(n-1)\left(t + \frac{1}{n-1} c_{n-1}(t_\varepsilon)\right)}}_{\text{converges, } =: \beta < \infty} \\ &= \beta e^{-t_\varepsilon M}. \end{aligned}$$

Then similarly as before

$$\gamma(t) \leq \limsup_{M \to \infty} \frac{1}{M} \log \beta e^{-t_\varepsilon M} = \limsup_{M \to \infty} \frac{\log \beta}{M} - t_\varepsilon = -t_\varepsilon = -c^{-1}(-t - \varepsilon),$$

which yields (3.29) when $\varepsilon \to 0$.

(V) If $b_2 < \infty$ and $t \leq -b_2$, we use the fact that we have just proven (3.29) for $t \in (-b_2, -b_1)$, and see that once again,

$$\gamma(t) \leq \limsup_{r \to -b_2^+} \gamma(t) \leq \limsup_{r \to -b_2^+} \Gamma(t) = -a_2 = \Gamma(t).$$

Part 3

Next we want to show that

$$\gamma^*(x) \geq \Gamma^*(x) \tag{3.30}$$

for each $x \in \mathbb{R}$. Let $\mathcal{K} \subset \mathbb{R}$ be a compact set. Recall from the beginning of this section the definitions of P and $\mathbb{Q}_M$, and notice that

$$\mathbb{P}\left(\frac{T}{M} \in \mathcal{K}\right) = \mathbb{P}\left(\frac{W_M}{M} \in \mathcal{K} \cap \{T < \infty\}\right) = P\mathbb{Q}_M\left(\frac{W_M}{M} \in \mathcal{K}\right).$$

meaning

$$\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{K}\right) = \underbrace{\limsup_{M\to\infty} \frac{\log P}{M}}_{=0} + \limsup_{M\to\infty} \frac{1}{M} \log \mathbb{Q}_M\left(\frac{W_M}{M} \in \mathcal{K}\right).$$

By Theorem 3.24 (I) as well as (3.27) and (3.30), this then implies

$$\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{K}\right) \leq -\inf_{x\in\mathcal{K}} \bar{\Lambda}^*(x) = -\inf_{x\in\mathcal{K}} \gamma^*(x) \leq -\inf_{x\in\mathcal{K}} \Gamma^*(x),$$

for any compact set $\mathcal{K} \subset \mathbb{R}$, meaning we have (3.13) for compact sets. In showing (3.30), we have two cases:

(I) If $\gamma(t) = -\infty$ for any $t \in \mathbb{R}$, then

$$\gamma^*(x) = \sup_{t\in\mathbb{R}}\{xt - \gamma(t)\} = xt + \infty = \infty,$$

for all $x \in \mathbb{R}$, so the case is trivial.

(II) If instead $\gamma(t) > -\infty$ for all $t \in \mathbb{R}$, then by the recently proven (3.29), γ is proper convex. Now denote by γ_L the lower semi-continuous hull of γ. Then $\gamma_L(t) \leq \gamma(t) \leq \Gamma(t)$ for all $t \in \mathbb{R} \setminus \{-b_1\}$, and also

$$\gamma_L(-b_1) = \liminf_{t\to -b_1} \gamma(t) \leq \liminf_{t\to -b_1} \Gamma(t) = -a_1 = \Gamma(-b_1),$$

so $\gamma_L(t) \leq \Gamma(t)$ for all $t \in \mathbb{R}$. Then for all $x \in \mathbb{R}$,

$$(\gamma_L)^*(x) = \sup_{t\in\mathbb{R}}\{xt - \gamma_L(t)\} \geq \sup_{t\in\mathbb{R}}\{xt - \Gamma(t)\} = \Gamma^*(x).$$

Now Theorem 3.20 tells us that $(\gamma_L)^* \equiv \gamma^*$, thus (3.30) holds.

We have now shown (3.13) for compact sets. In order to extend it to closed sets, it again suffices to check two cases.

(I) If $a_1 < w$, then by the definition of w, because $w \in [a_1, a_2]$, and because c is strictly increasing on J, there must exist some $t_0 \in (a_1, w) \subset J$, such that $c(t_0) < 0$. Now the definition of Γ and the fact that $-c(t_0) \in (-b_2, -b_1)$ implies that

$$\Gamma(t) \leq \Gamma(-c(t_0)) = -t_0 < \infty,$$

for every $t \in (-\infty, -c(t_0))$. Thus by (3.27) and (3.29), for every such t we also have

$$\bar{\Lambda}(t) = \gamma(t) \leq \Gamma(t) < \infty,$$

for every $t \in (-\infty, -c(t_0))$. As the origin belongs to this interval, we have $0 \in \operatorname{int} \mathcal{D}_{\bar{\Lambda}}$, so Theorem 3.24 (I) implies that (3.13) also holds for closed sets.

(II) If instead $a_1 \geq w$, then since $w \in [a_1, a_2]$ we must have $w = a_1$, but since $w \in [0, \infty]$ and we assumed that $a_1 < w$ for positive w, our case must be $w = a_1 = 0$. As c is convex and finite, and strictly increasing on (a_1, a_2), we must have

$$b_1 = \lim_{t \to a_1^+} c(t) = \lim_{t \to 0^+} c(t) = c(0) = 0.$$

Then the definition of Γ says that $\Gamma(t) = \infty$ for all positive t. This means that

$$\Gamma^*(x) = \sup_{t \in \mathbb{R}} \{tx - \Gamma(t)\} = \sup_{t \leq 0} \{tx - \Gamma(t)\},$$

which lets us conclude that Γ^* is decreasing, as $x_1 < x_2$ implies

$$\Gamma^*(x_1) = \sup_{t \leq 0} \{tx_1 - \Gamma(t)\} \geq \sup_{t \leq 0} \{tx_2 - \Gamma(t)\} = \Gamma^*(x_2).$$

By Lemma 3.25, Γ is proper convex and lower semi-continuous, so Theorem 3.20 and Remark 3.19 yield $(\Gamma^*)^* \equiv \Gamma_L \equiv \Gamma$, and Γ^* is proper convex and lower semi-continuous. Theorem 3.22 then lets us conclude that

$$\inf_{x \in \mathbb{R}} \Gamma^*(x) = -(\Gamma^*)^*(0) = -\Gamma(0) = -\Gamma(-b_1) = a_1 = 0.$$

As we recently showed Γ^* is decreasing, we also have

$$\lim_{x \to \infty} \Gamma^*(x) = \inf_{x \in \mathbb{R}} \Gamma^*(x) = 0. \tag{3.31}$$

Now that we have shown this, we can get to proving (3.13) for closed sets. To this end, let $\mathcal{C}$ be a closed set. If $\mathcal{C}$ is compact, we have already shown (3.13) above. Suppose then $\mathcal{C}$ is closed and non-compact. Then $\mathcal{C}$ is unbounded, and we have two cases:

(a) If $\mathcal{C}$ is unbounded to the right, then

$$\inf_{x\in\mathcal{C}} \Gamma^*(x) = \lim_{x\to\infty} \Gamma^*(x) = 0,$$

Now since

$$\mathbb{P}\left(\frac{T}{M} \in \mathcal{C}\right) \leq 1,$$

we have

$$\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{C}\right) \leq 0 = -\inf_{x\in\mathcal{C}} \Gamma^*(x),$$

which is (3.13).

(b) If $\mathcal{C}$ is instead bounded to the right, but unbounded to the left, then $\mathcal{C} \cap [0,\infty)$ is compact. Notice also that since we always have $T \geq 0$ and $M > 0$, it holds for any closed set $\mathcal{C}$ that

$$\mathbb{P}\left(\frac{T}{M} \in \mathcal{C}\right) = \mathbb{P}\left(\frac{T}{M} \in \mathcal{C} \cap [0,\infty)\right).$$

Then as we have already proven (3.13) for compact sets we have

$$\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{C} \cap [0,\infty)\right) = -\inf_{x\in\mathcal{C}\cap[0,\infty)} \Gamma^*(x) = \inf_{x\in\mathcal{C}} \Gamma^*(x),$$

where the second equality is due to Lemma 3.25. Putting this together yields

$$\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{C}\right) = -\inf_{x\in\mathcal{C}} \Gamma^*(x),$$

which is (3.13).

□

We have now proven Theorem 3.11. The following section constitutes the proof of the lower bound, Theorem 3.14.

3.4 Proof of Theorem 3.14

Throughout this entire section, we assume 3.5 (I)-3.5 (VII). Before we begin, denote

$$\mathcal{F} := \left\{\frac{1}{c'(u)} : u \in \tilde{J}\right\} \quad \text{and} \quad \mu := \begin{cases} \infty, & \text{if } w = a_1, \\ \frac{1}{c'(w)}, & \text{if } w \in (a_1, a_2), \\ 0, & \text{if } w = a_2. \end{cases}$$

The structure of this proof is as follows. Part 1 proves a set of important properties of the set $\mathcal{F}$. We then assume in all that follows that $p(u,\tau) < c'(u)$ for all $u \in \tilde{J}$ and $0 \le \tau < 1$. Part 2 introduces an auxiliary function and proves an inequality involving said function. In Parts 3 and 4 we prove that if $\mathcal{F}$ is non-empty, then for all $x \in \mathcal{F}$ and $\varepsilon \in (0,x)$,

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in (x-\varepsilon, x+\varepsilon)\right) \ge -xc^*\left(\frac{1}{x}\right). \tag{3.32}$$

Part 3 proves the above when $w > 0$, while Part 4 concerns the case $w = 0$. In Part 5 we show the desired (3.15) and (3.16), using the result (3.32) in the case of non-empty $\mathcal{F}$.

Part 1

The following lemma serves as a collection of properties of the set $\mathcal{F}$.

Lemma 3.33. *Suppose $\mathcal{F}$ is non-empty. Then the following hold:*

(I) *$\mathcal{F} \subset (0,\infty)$, and $\mathcal{F}$ is an open interval.*

(II) *The function c^* is finite and continuous at $\frac{1}{x}$, for each $x \in \mathcal{F}$. Also, $c^*\left(\frac{1}{x}\right) = \frac{u}{x} - c(u)$, with $u \in J$ chosen such that $c'(u) = \frac{1}{x}$.*

(III) *The function Γ^* is finite and continuous throughout $\mathcal{F}$, and $\mathcal{D}_{\Gamma^*} \subset \overline{\mathcal{F}}$.*

(IV) *If $\mu < \infty$, then $\Gamma^*(\mu) = \inf_{x\in\mathbb{R}} \Gamma^*(x) = w$, and μ is the unique point where this minimum is attained. Also, Γ^* is strictly decreasing on $(-\infty,\mu) \cap \mathcal{F}$ and strictly increasing on $(\mu,\infty) \cap \mathcal{F}$.*

Proof. (I) Suppose there exists some point $t_0 \in J$ such that c' is discontinuous at t_0. Then as c is convex, finite and strictly increasing on J, this must be a jump discontinuity, meaning $c'(t_0)$ is undefined. This is in contradiction with assumption 3.5 (V), stating the derivative c' exists at every point $t \in J$, therefore c' must be continuous throughout J. Then J being an open interval implies $\tilde{J}$ being an open interval, by which in turn $\mathcal{F}$ must be an open interval. As c is strictly increasing on J, c' is positive on $\tilde{J}$, thus $\frac{1}{c'(u)} > 0$ for each $u \in \tilde{J}$.

(II) Fix $x \in \mathcal{F}$. Since $\frac{t}{x} - c(t)$ is maximised when $t = u \in J$, i.e. when $\frac{1}{x} = c'(t)$, we have

$$c^*\left(\frac{1}{x}\right) = \sup_{t\in\mathbb{R}}\left\{\frac{t}{x} - c(t)\right\} = \frac{u}{x} - c(u) \in (-\infty,\infty).$$

We just showed $\mathcal{F}$ is an open interval, thus also $c^*(y) \in (-\infty, \infty)$ when $y \in B_\varepsilon\left(\frac{1}{x}\right)$ for some small enough $\varepsilon > 0$. Now Theorem 3.20 says c^* is convex, which together with finiteness implies c^* is continuous on $B_\varepsilon(x)$.

(III) Finite- and continuousness follows from (II) and from Lemma 3.25. Denote the boundary points of $\mathcal{F}$ by $\kappa_1 = \lim_{t \to a_2^-} \frac{1}{c'(t)}$ and $\kappa_2 = \lim_{t \to a_1^+} \frac{1}{c'(t)}$. We want to show that if $x < \kappa_1$ or $x > \kappa_2$, then $\Gamma^*(x) = \infty$. We have two cases.

(a) Let $x < \kappa_1$. As $\mathcal{F} \subset (0, \infty)$, it must hold that $\kappa_1 \geq 0$. If $\kappa_1 = 0$, then by Lemma 3.25, $\Gamma^*(x) = \infty$ for $x < 0$, so we are done. Suppose instead $\kappa_1 > 0$ and $x \in (0, \kappa_1)$. Then assumption 3.5 (VI) implies $a_2 = \infty$. Letting $t \to \infty = a_2$, we see that $\frac{t}{x} - c(t) \to \infty$, because it holds for the derivative that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t}{x} - c(t)\right) = \frac{1}{x} - c'(t) \to \frac{1}{x} - \frac{1}{\kappa_1} > 0.$$

Then by Lemma 3.25,

$$\Gamma^*(x) = xc^*\left(\frac{1}{x}\right) = x \sup_{t \in \mathbb{R}} \left\{\frac{t}{x} - c(t)\right\} = x \lim_{t \to \infty} \left(\frac{t}{x} - c(t)\right) = \infty,$$

and case $x = 0$ also follows by Lemma 3.25.

(b) Suppose $\kappa_2 < \infty$ and let $x > \kappa_2$. Assumption 3.5 (VII) then implies $a_1 = -\infty$. Also, as c is convex and strictly increasing, c' is positive and increasing, meaning $\kappa_2 > 0$. Now if we let $t \to -\infty$ we see that $\frac{t}{x} - c(t) \to \infty$, because

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t}{x} - c(t)\right) = \frac{1}{x} - c'(t) \to \frac{1}{x} - \frac{1}{\kappa_2} < 0,$$

Then by Lemma 3.25,

$$\Gamma^*(x) = xc^*\left(\frac{1}{x}\right) = x \sup_{t \in \mathbb{R}} \left\{\frac{1}{x} - c(t)\right\} = x \lim_{t \to -\infty} \left(\frac{t}{x} - c(t)\right) = \infty.$$

(IV) Suppose $\mu < \infty$. Then $w \in (a_1, a_2)$, so $0 \in (-b_2, -b_1)$, meaning

$$\Gamma'(0) = \frac{1}{c'(c^{-1}(0))} = \frac{1}{c'(w)} = \mu.$$

Now by Lemma 3.25, Γ is proper convex, so Theorem 3.20 implies Γ^* is proper convex and lower semi-continuous, thus by Theorem 3.22,

$$\inf_{x \in \mathbb{R}} \Gamma^*(x) = -\Gamma(0) = c^{-1}(0) = w.$$

Also, $\frac{\mathrm{d}}{\mathrm{d}t} t\mu - \Gamma(t) = \mu - \Gamma'(t)$, which is equal to zero if and only if $t = 0$. Then

$$\Gamma^*(\mu) = \sup_{t \in \mathbb{R}} \{t\mu - \Gamma(t)\} = -\Gamma(0) = w,$$

so the minimum $\inf_{x\in\mathbb{R}} \Gamma^*(x) = w$ is attained only at the point $x = \mu$. Now Γ^* must be strictly decreasing on $(-\infty, \mu) \cap \mathcal{F}$ and strictly increasing on $(\mu, \infty) \cap \mathcal{F}$, because Γ^* is convex.

$\square$

Part 2

This part proves a bound used in Part 3. Suppose $\mathcal{F}$ is non-empty. Fix $x \in \mathcal{F}$, $u \in \tilde{J}$ and $0 \le \tau < 1$, and define for each $\delta > 0$ the function $g_\delta : \mathbb{R} \to [-\infty, \infty]$ as

$$g_\delta(t) = -t - u + xH_\delta(t, u; \tau),$$

and notice that g_δ is convex by the convexity of H in the variable t, shown in Remark 3.10. In the interest of expediting the proof in Part 3, we now aim to show that

$$\limsup_{\delta \to 0^+} g_\delta'(0+) \le -1 + xp(u; \tau), \tag{3.34}$$

where $g_\delta'(0+)$ denotes the right-hand derivative of g_δ at the origin. More precisely,

$$g_\delta'(0+) = \lim_{h \to 0^+} \frac{g_\delta(h) - g_\delta(0)}{h}.$$

Also remember that $H(0, u; \tau) = H_\delta(0, u; \tau) = c(u)$. In proving (3.34), we have two cases:

(I) If $p(u; \tau) > -\infty$, let $\varepsilon > 0$, and choose a small enough $h_\varepsilon > 0$ such that

$$\frac{H(h_\varepsilon, u; \tau) - H(0, u; \tau)}{h_\varepsilon} = \frac{H(h_\varepsilon, u; \tau) - c(u)}{h_\varepsilon} < p(u; \tau) + \varepsilon.$$

After this, choose $\delta_\varepsilon > 0$ small enough that for all $\delta \in (0, \delta_\varepsilon)$,

$$H_\delta(h_\varepsilon, u; \tau) \le H(h_\varepsilon, u; \tau) + \varepsilon h_\varepsilon.$$

Now the convexity of g_δ means its derivative is increasing, so for every $\delta \in (0, \delta_\varepsilon)$,

$$\begin{aligned}
g_\delta'(0+) &\le \frac{g_\delta(h_\varepsilon) - g_\delta(0)}{h_\varepsilon} \\
&= \frac{-h_\varepsilon - u + xH_\delta(h_\varepsilon, u; \tau) + u - xH_\delta(0, u; \tau)}{h_\varepsilon} \\
&= -1 + x\left(\frac{H_\delta(h_\varepsilon, u; \tau) - c(u)}{h_\varepsilon}\right) \\
&\le -1 + x\left(\frac{H(h_\varepsilon, u; \tau) - c(u)}{h_\varepsilon} + \varepsilon\right) \\
&< -1 + xp(u; \tau) + 2x\varepsilon.
\end{aligned}$$

Now letting $\varepsilon \to 0$ yields

$$\limsup_{\delta \to 0^+} g_\delta'(0+) \le -1 + xp(u; \tau),$$

which is (3.34), as we wanted.

(II) If instead $p(u; \tau) = -\infty$, let $\varepsilon > 0$. Then by (3.8), for every $h > 0$ we can choose $\delta_h > 0$ such that when $\delta \in (0, \delta_h)$,

$$H_\delta(h, u; \tau) \le H(h, u; \tau) + h\varepsilon.$$

Then by the convexity of g_δ,

$$\begin{aligned}
g_\delta'(0+) &\le \frac{g_\delta(h) - g_\delta(0)}{h} \\
&= -1 + x\left(\frac{H_\delta(h, u; \tau) - c(u)}{h}\right) \\
&\le -1 + x\left(\frac{H(h, u; \tau) - c(u)}{h} + \varepsilon\right).
\end{aligned}$$

Now by first taking the limes supremum as $\delta \to 0$, then the limit as $h \to 0^+$, and finally letting $\varepsilon \to 0$, we get

$$\lim_{\varepsilon \to 0} \lim_{h \to 0^+} \limsup_{\delta \to 0} g_\delta'(0+) \le \lim_{\varepsilon \to 0} \lim_{h \to 0^+} \limsup_{\delta \to 0} \left(-1 + x\left(\frac{H(h, u; \tau) - c(u)}{h} + \varepsilon\right)\right),$$

which reduces down to

$$\begin{aligned}
\limsup_{\delta \to 0} g_\delta'(0+) &\le \lim_{\varepsilon \to 0} \left(-1 + xp(u; \tau) + x\varepsilon\right) \\
&= -1 + xp(u; \tau) \\
&= -\infty,
\end{aligned}$$

as $x > 0$. Thus we have shown (3.34).

Part 3

In this part we want to show (3.32) when $w > 0$, whereas Part 4 concerns the case $w = 0$. The idea here is to first prove (3.38) in steps, using the auxiliary bounds (3.35), (3.36) and (3.37). After this we show (3.39), and finally use Lemma 3.23 to arrive at (3.32).

Suppose that $p(u;\tau) > c'(u)$ for every $u \in \tilde{J}$ and $0 \le \tau < 1$. Suppose that $\mathcal{F}$ is non-empty and fix $x \in \mathcal{F}$. Finally suppose that $w > 0$. Then J lies in its entirety to the right of the origin, and $0 \notin J$. Recall that $\mathcal{F} \subset (0,\infty)$. Let $u \in \tilde{J}$ be such that $x = \frac{1}{c'(u)}$. This is guaranteed to exist by the definition of $\mathcal{F}$, and in fact this means $u > 0$. Now let $t \ge 0$, $\varepsilon \in (0, x)$ and $\delta' \in (0, \varepsilon)$. Then by Chebychev's inequality (2.2),

$$\begin{aligned}
\mathbb{E}\left(e^{tY_{\lceil \vartheta M\rceil}+uY_{\lceil xM\rceil}}\right) &= \mathbb{E}\left(e^{(t+u)\left(\frac{t}{t+u}Y_{\lceil \vartheta' M\rceil}+\frac{u}{t+u}Y_{\lceil xM\rceil}\right)}\right)\\
&\overset{(2.2)}{\ge} e^{(t+u)M}\mathbb{P}\left(\frac{t}{t+u}Y_{\lceil \vartheta' M\rceil}+\frac{u}{t+u}Y_{\lceil xM\rceil} \ge M\right)\\
&= e^{(t+u)M}\mathbb{P}\left(tY_{\lceil \vartheta' M\rceil}+uY_{\lceil xM\rceil} \ge (t+u)M\right)\\
&\ge e^{(t+u)M}\mathbb{P}\left(Y_{\lceil \vartheta' M\rceil} > M \text{ and } Y_{\lceil xM\rceil} > M\right),
\end{aligned}$$

yielding

$$\frac{1}{M}\log\mathbb{E}\left(e^{tY_{\lceil \vartheta M\rceil}+uY_{\lceil xM\rceil}}\right) \ge t+u+\frac{1}{M}\log\mathbb{P}\left(Y_{\lceil \vartheta' M\rceil} > M \text{ and } Y_{\lceil xM\rceil} > M\right).$$

Let $y \in [0, x-\varepsilon]$, and denote $\tau = \frac{y}{x} \in [0,1)$. Also denote $\delta = \frac{\delta'}{x}$ and note that $\vartheta := \frac{\vartheta'}{x} \in B_\delta(\tau) \cap (0,\infty)$ for every $\vartheta' \in B_{\delta'}(y) \cap (0,\infty)$. Now we have

$$\begin{aligned}
&\limsup_{M\to\infty}\frac{1}{M}\log\sup_{\vartheta'\in B_{\delta'}(y)\cap(0,\infty)}\mathbb{P}\left(Y_{\lceil \vartheta' M\rceil} > M \text{ and } Y_{\lceil xM\rceil} > M\right)\\
&\le -t-u+\limsup_{M\to\infty}\sup_{\vartheta'\in B_{\delta'}(y)\cap(0,\infty)}\frac{1}{M}\log\mathbb{E}\left(e^{tY_{\lceil \vartheta' M\rceil}+uY_{\lceil xM\rceil}}\right)\\
&\le -t-u+x\left(\limsup_{M\to\infty}\sup_{\vartheta\in B_{\delta}(\tau)\cap(0,\infty)}\frac{1}{xM}\log\mathbb{E}\left(e^{tY_{\lceil \vartheta xM\rceil}+uY_{\lceil xM\rceil}}\right)\right)\\
&= -t-u+xH_\delta(t,u;\tau) = g_\delta(t).
\end{aligned} \tag{3.35}$$

By assumption $p(u;\tau) < c'(u)$, so by (3.34) and by the choice of u,

$$\limsup_{\delta\to 0} g_\delta'(0+) \le -1 + xp(u;\tau) < -1 + xc'(u) = -1+1 = 0,$$

Then, for small enough $\delta > 0$, the function $-t-u+xH_\delta(t,u;\tau)$ is decreasing in t on the right side at the origin. For these δ we can then find small enough $t > 0$ such that

$$-t-u+xH_\delta(t,u;\tau) < -u+xH_\delta(0,u;\tau) = -u+xc(u) = -xc^*\left(\frac{1}{x}\right), \tag{3.36}$$

where the last equality is because

$$
\begin{aligned}
xc^*\left(\frac{1}{x}\right) &= x\sup_{t\in\mathbb{R}}\left\{\frac{t}{x}-c(t)\right\}\\
&= \sup_{t\in\mathbb{R}}\left\{t-xc(t)\right\},
\end{aligned}
$$

and $t - xc(t)$ is maximised when $c'(t) = \frac{1}{x}$, i.e. when $t = u$.

Now, $\frac{T}{M} \in B_{\delta'}(y)$ means that $T = \vartheta' M$ for some $\vartheta' \in B_{\delta'}(y) \cap (0,\infty)$. Then it must also be true that $Y_{\lceil \vartheta' M\rceil} > M$. Thus

$$
\mathbb{P}\left(\frac{T}{M} \in B_{\delta'}(y)\right) \le \mathbb{P}\left(Y_{\lceil \vartheta' M\rceil} > M \text{ for some } \vartheta' \in B_{\delta'}(y)\cap(0,\infty)\right),
$$

and therefore

$$
\begin{aligned}
&\limsup_{M\to\infty}\frac{1}{M}\log\mathbb{P}\left(\frac{T}{M}\in B_{\delta'}(y) \text{ and } Y_{\lceil xM\rceil} > M\right)\\
\le\ &\limsup_{M\to\infty}\frac{1}{M}\log\mathbb{P}\left(Y_{\lceil \vartheta' M\rceil} > M \text{ for some } \vartheta' \in B_{\delta'}(y)\cap(0,\infty) \text{ and } Y_{\lceil xM\rceil} > M\right)\\
\le\ &\limsup_{M\to\infty}\frac{1}{M}\log\lceil (y+\delta')M\rceil \sup_{\vartheta'\in B_{\delta'}(y)\cap(0,\infty)}\mathbb{P}\left(Y_{\lceil \vartheta' M\rceil} > M \text{ and } Y_{\lceil xM\rceil} > M\right)\\
=\ &\limsup_{M\to\infty}\sup_{\vartheta'\in B_{\delta'}(y)\cap(0,\infty)}\frac{1}{M}\log\mathbb{P}\left(Y_{\lceil \vartheta' M\rceil} > M \text{ and } Y_{\lceil xM\rceil} > M\right).
\end{aligned}
\tag{3.37}
$$

As the balls $B'_\delta(y)$ together cover the compact set $[0, x-\varepsilon]$, we can extract a finite subcover from these. Call these $B_{\delta'_i}(y_i)$ with $i = 1,\dots,K$. Then by applying Lemma 2.14, we arrive at

$$
\begin{aligned}
&\limsup_{M\to\infty}\frac{1}{M}\log\mathbb{P}\left(\frac{T}{M}\in[0,x-\varepsilon] \text{ and } Y_{\lceil xM\rceil} > M\right)\\
\le\ &\limsup_{M\to\infty}\frac{1}{M}\log\sum_{i=1}^{K}\mathbb{P}\left(\frac{T}{M}\in B_{\delta'_i}(y_i) \text{ and } Y_{\lceil xM\rceil} > M\right)\\
=\ &\max_{i=1,\dots,K}\limsup_{M\to\infty}\frac{1}{M}\log\mathbb{P}\left(\frac{T}{M}\in B_{\delta'_i}(y_i) \text{ and } Y_{\lceil xM\rceil} > M\right)\\
<\ &-xc^*\left(\frac{1}{x}\right),
\end{aligned}
\tag{3.38}
$$

where the last inequality is due to (3.35), (3.36) and (3.37). The above is the first important bound in this part. We now want to show the second, i.e. (3.39).

For all $t \in \mathbb{R}$, we have

$$\lim_{M\to\infty} \frac{1}{xM} \log \mathbb{E}\left(e^{tY_{\lceil xM \rceil}}\right) = c(t).$$

By assumption 3.5 (IV), this limit exists, and is finite in a neighbourhood of $u \in J$. By assumption 3.5 (V), the derivative $c'(t)$ exists in some neighbourhood G_u around u. As $\mathcal{F}$ is non-empty, for any $\varepsilon' > 0$ small enough that $x - \varepsilon' \in \mathcal{F}$, we can find $t_{\varepsilon'} \in G_u$ such that $c'(t_{\varepsilon'}) = \frac{1}{x-\varepsilon'} > \frac{1}{x}$. Then by choosing $\varepsilon'' > 0$ small enough that $\frac{1}{x-\varepsilon'} - \varepsilon'' > \frac{1}{x}$ and applying Theorem 3.24 (II),

$$\begin{aligned}
\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(Y_{\lceil xM \rceil} > M\frac{x}{x-\varepsilon'} - xM\varepsilon''\right) &= x\left(\liminf_{M\to\infty} \frac{1}{xM} \log \mathbb{P}\left(\frac{Y_{\lceil xM \rceil}}{xM} > \frac{1}{x-\varepsilon'} - \varepsilon''\right)\right) \\
&\geq x\left(c(t_{\varepsilon'}) - t_{\varepsilon'}c'(t_{\varepsilon'})\right) \\
&= -x\left(\frac{t_{\varepsilon'}}{x-\varepsilon'} - c(t_{\varepsilon'})\right) \\
&\geq -xc^*\left(\frac{1}{x-\varepsilon'}\right).
\end{aligned}$$

Letting first $\varepsilon'' \to 0$ gives

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(Y_{\lceil xM \rceil} \geq M\frac{x}{x-\varepsilon'}\right) \geq -xc^*\left(\frac{1}{x-\varepsilon'}\right)$$

Then as c^* is continuous at $\frac{1}{x}$ by Lemma 3.33, letting $\varepsilon' \to 0$ yields

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(Y_{\lceil xM \rceil} > M\right) \geq -xc^*\left(\frac{1}{x}\right). \tag{3.39}$$

This is the second important bound in this part. We will now show (3.32) using (3.39) and (3.38).

Since $M(x+\varepsilon) > \lceil xM \rceil$ for large enough $M > 0$, it also holds that

$$\mathbb{P}\left(\frac{T}{M} \in [x+\varepsilon, \infty) \text{ and } Y_{\lceil xM \rceil} > M\right) = 0,$$

and thus

$$\begin{aligned}
\mathbb{P}\left(\frac{T}{M} \in (x-\varepsilon, x+\varepsilon)\right) &\geq \mathbb{P}\left(\frac{T}{M} \in (x-\varepsilon, x+\varepsilon) \text{ and } Y_{\lceil xM \rceil} > M\right) \\
&= \mathbb{P}\left(Y_{\lceil xM \rceil} > M\right) - \mathbb{P}\left(\frac{T}{M} \in [0, x-\varepsilon] \text{ and } Y_{\lceil xM \rceil} > M\right),
\end{aligned}$$

for large enough $M > 0$. Then recalling (3.38) and (3.39), and applying Lemma 3.23 gives us

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in (x-\varepsilon, x+\varepsilon)\right) \geq -xc^*\left(\frac{1}{x}\right),$$

which is (3.32).

Part 4

The previous part concerned the case $w > 0$. In this part we tackle the case $w = 0$, using a somewhat similar strategy.

Suppose, again, that $p(u;\tau) < c'(u)$ for all $u \in \tilde{J}$ and $0 \leq \tau < 1$, and that $\mathcal{F}$ is non-empty. Also suppose $w = 0$. If $x \in \mathcal{F} \cap (-\infty, \mu)$, we can choose $u \in \tilde{J}$ such that $c'(u) = \frac{1}{x}$. Then $c'(u) \geq c'(w)$, so since c is convex, and thus c' increasing, we must have $u \geq 0 = w$. The proof in Part 3 can then be repeated, so (3.32) holds when $x \in \mathcal{F} \cap (-\infty, \mu]$.

If $\mu = \infty$ there is nothing more to prove; suppose instead $\mu < \infty$. We now want to show that (3.32) holds when $x \in \mathcal{F} \cap (\mu, \infty)$. As $\mu < \infty$ implies $a_1 < w$, we can find $t_0 \in J$ such that $t_0 < 0$ and $c(t_0) < 0$. Then Chebychev's inequality (2.2) gives, for any $n \geq 1$ and $M > 0$,

$$\mathbb{E}\left(e^{t_0 Y_n}\right) \geq e^{t_0 M}\mathbb{P}(Y_n \leq M).$$

Thus

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Y_n \leq M) \leq \limsup_{n\to\infty} \frac{1}{n} \log e^{-t_0 M}\mathbb{E}\left(e^{t_0 Y_n}\right) \tag{3.40}$$

$$= \underbrace{\limsup_{n\to\infty} \frac{1}{n} \log e^{-t_0 M}}_{=0} + \underbrace{\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left(e^{t_0 Y_n}\right)}_{=c(t_0)} \tag{3.41}$$

$$= c(t_0) < 0. \tag{3.42}$$

Here we must have $\log \mathbb{P}(Y_n \leq M) \to -\infty$, hence also $\mathbb{P}(Y_n > M) \to 1$, as $n \to \infty$. This means $T < \infty$ with probability 1, for every $M > 0$. Suppose now $x \in \mathcal{F} \cap (\mu, \infty)$ and choose $u \in \tilde{J}$ such that $c'(u) = \frac{1}{x}$. Then $c'(u) < c'(w)$, so $u < 0 = w$. Let $\varepsilon > 0$ and choose $y \in [0, x-\varepsilon]$. For every $\delta' \in (0, \varepsilon)$ and small enough $t \geq 0$ that $t + u \leq 0$, Chebychev's inequality (2.2) gives

$$\begin{aligned}
\mathbb{E}\left(e^{tY_{\lceil \vartheta M\rceil}+uY_{\lceil xM\rceil}}\right) &= \mathbb{E}\left(e^{(t+u)\left(\frac{t}{t+u}Y_{\lceil \vartheta M\rceil}+\frac{u}{t+u}Y_{\lceil xM\rceil}\right)}\right)\\
&\overset{(2.2)}{\geq} e^{(t+u)M}\mathbb{P}\left(\frac{t}{t+u}Y_{\lceil \vartheta M\rceil}+\frac{u}{t+u}Y_{\lceil xM\rceil}\geq M\right)\\
&= e^{(t+u)M}\mathbb{P}\left(tY_{\lceil \vartheta M\rceil}+uY_{\lceil xM\rceil}\geq (t+u)M\right)\\
&\geq e^{(t+u)M}\mathbb{P}\left(tY_{\lceil \vartheta M\rceil}> tM \text{ and } uY_{\lceil xM\rceil}\geq uM\right)\\
&= e^{(t+u)M}\mathbb{P}\left(Y_{\lceil \vartheta M\rceil}> M \text{ and } Y_{\lceil xM\rceil}\leq M\right).
\end{aligned}$$

We now again proceed similarly to how we did in Part 3 to get the auxiliary bounds (3.35), (3.36) and (3.37), and arrive at

$$\limsup_{M\to\infty}\frac{1}{M}\log\mathbb{P}\left(\frac{T}{M}\in[0,x-\varepsilon] \text{ and } Y_{\lceil xM\rceil}\leq M\right) < -xc^*\left(\frac{1}{x}\right). \tag{3.43}$$

This is the first important bound in this part. Showing the second, (3.44), is slightly easier:

We let $\varepsilon'>0$ and apply Theorem 3.24 (II) as we did in Part 3, giving us

$$\begin{aligned}
\liminf_{M\to\infty}\frac{1}{M}\log\mathbb{P}\left(Y_{\lceil xM\rceil}< M+xM\varepsilon'\right) &= x\left(\liminf_{M\to\infty}\frac{1}{xM}\log\mathbb{P}\left(\frac{Y_{\lceil xM\rceil}}{xM}< c'(u)+\varepsilon'\right)\right)\\
&\geq x(c(u)-uc'(u))\\
&= -xc^*\left(\frac{1}{x}\right).
\end{aligned}$$

Letting $\varepsilon'\to 0$ yields the desired

$$\liminf_{M\to\infty}\frac{1}{M}\log\mathbb{P}(Y_{\lceil xM\rceil}\leq M)\geq -xc^*\left(\frac{1}{x}\right). \tag{3.44}$$

This is the second important bound in this part. The aim is now to use (3.44) and (3.43) to show (3.32).

To this end, notice that

$$\begin{aligned}
&\mathbb{P}\left(\frac{T}{M} \in (x-\varepsilon, \infty)\right) \\
&= 1 - \mathbb{P}\left(\frac{T}{M} \in [0, x-\varepsilon]\right) \\
&= 1 - \mathbb{P}\left(\frac{T}{M} \in [0, x-\varepsilon] \text{ and } Y_{\lceil xM \rceil} > M\right) - \mathbb{P}\left(\frac{T}{M} \in [0, x-\varepsilon] \text{ and } Y_{\lceil xM \rceil} \leq M\right) \\
&= \mathbb{P}\left(\frac{T}{M} \in (x-\varepsilon, \infty) \text{ or } Y_{\lceil xM \rceil} \leq M\right) - \mathbb{P}\left(\frac{T}{M} \in [0, x-\varepsilon] \text{ and } Y_{\lceil xM \rceil} \leq M\right) \\
&\geq \mathbb{P}\left(Y_{\lceil xM \rceil} \leq M\right) - \mathbb{P}\left(\frac{T}{M} \in [0, x-\varepsilon] \text{ and } Y_{\lceil xM \rceil} \leq M\right)
\end{aligned}$$

By (3.43) and (3.44) and Lemma 3.23 we then have

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in (x-\varepsilon, \infty)\right) \geq -xc^*\left(\frac{1}{x}\right).$$

Applying the previously proved Theorem 3.11 yields

$$\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in [x+\varepsilon, \infty)\right) \leq -\inf_{y \geq x+\varepsilon} \Gamma^*(y) = -\Gamma^*(x+\varepsilon) < -\Gamma^*(x) = -xc^*\left(\frac{1}{x}\right),$$

because Γ^* is strictly increasing on (μ, ∞), by Lemma 3.33. Now by writing as a difference

$$\mathbb{P}\left(\frac{T}{M} \in (x-\varepsilon, x+\varepsilon)\right) = \mathbb{P}\left(\frac{T}{M} \in (x-\varepsilon, \infty)\right) - \mathbb{P}\left(\frac{T}{M} \in [x+\varepsilon, \infty)\right),$$

and again applying Lemma 3.23, we get

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in (x-\varepsilon, x+\varepsilon)\right) \geq -xc^*\left(\frac{1}{x}\right),$$

which is (3.32).

Part 5

This is the final part, where we arrive at the main results (3.15) and (3.16). The proof of (3.16) is largely split into two cases: either the set $\mathcal{F}$ is empty, or it is not. In the non-empty case, we will use the bound (3.32), which we just spent some effort proving. The empty case is a little bit more straightforward.

We begin by proving (3.16). Let $\mathcal{O} \subset \mathbb{R}$ be an open set. We have two cases:

(I) Suppose $\mathcal{F}$ is non-empty. If $\mathcal{O} \cap \mathcal{F} = \emptyset$, then it holds trivially that

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{O}\right) \geq - \inf_{x\in\mathcal{O}\cap\mathcal{F}} \Gamma^*(x) = -\infty,$$

as we are using the convention that $\inf \emptyset = \infty$. If $\mathcal{O} \cap \mathcal{F} \neq \emptyset$ and $x \in \mathcal{O} \cap \mathcal{F}$, then choose $\varepsilon_x \in (0, x)$ such that $(x - \varepsilon_x, x + \varepsilon_x) \subset \mathcal{O} \cap \mathcal{F}$. By Lemma 3.25 and Theorem 3.20, Γ^* is proper convex and lower semi-continuous, thus also continuous throughout $\mathcal{D}_{\Gamma^*}$. Using (3.32) and Lemma 3.33(III) we get

$$\begin{aligned}
\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{O}\right) &\geq \liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{O} \cap \mathcal{F}\right) \\
&\geq \sup_{x\in\mathcal{O}\cap\mathcal{F}} \left\{\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in (x - \varepsilon_x, x + \varepsilon_x)\right)\right\} \\
(3.32) \quad &\geq \sup_{x\in\mathcal{O}\cap\mathcal{F}} \left\{-xc^*\left(\frac{1}{x}\right)\right\} \\
(\text{Lemma 3.25}) \quad &\geq - \inf_{x\in\mathcal{O}\cap\mathcal{F}} \Gamma^*(x) \\
(\text{proper convexity}) \quad &= - \inf_{x\in\mathcal{O}\cap\overline{\mathcal{F}}} \Gamma^*(x) \\
(\text{Lemma 3.33(III)}) \quad &= - \inf_{x\in\mathcal{O}\cap\mathcal{D}_{\Gamma^*}} \Gamma^*(x) \\
(\text{definition of } \mathcal{D}_{\Gamma^*}) \quad &= - \inf_{x\in\mathcal{O}} \Gamma^*(x).
\end{aligned}$$

This is (3.16) in the case of non-empty $\mathcal{F}$.

(II) Suppose now $\mathcal{F}$ is empty. Then the set $\{c'(t) : t \in J\}$ consists only of isolated point. By assumption 3.5 (V), the derivative function c' can have no jumps, thus $\{c'(t) : t \in J\} = \{\lambda\}$ is a singleton, with $\lambda > 0$. By assumptions 3.5 (VI) and 3.5 (VII), we must have $J = (-\infty, \infty)$, and thus $c(t) = \lambda t$, for all $t \in \mathbb{R}$. This means $0 = w \in J$, and thus $\frac{1}{\lambda} = \mu < \infty$. We then have

$$\Gamma(t) = -c^{-1}(-t) = \frac{t}{\lambda} = \mu t,$$

for $t \in \mathbb{R}$, so

$$\Gamma^*(x) = \sup_{t\in\mathbb{R}}\{tx - \Gamma(t)\} = \sup_{t\in\mathbb{R}}\{t(x - \mu)\} = \begin{cases} 0, & \text{if } x = \mu, \\ \infty, & \text{if } x \neq \mu. \end{cases} \tag{3.45}$$

We showed in Part 4 that in the case $w = 0$, we have $\mathbb{P}(T < \infty) = 1$ for all $M > 0$. Recall now from the proof of Theorem 3.11 the convex function

$$\gamma(t) = \limsup_{M\to\infty} \frac{1}{M} \log \mathbb{E}\left(e^{tT}\mathbb{I}\{T < \infty\}\right) = \limsup_{M\to\infty} \frac{1}{M} \log \mathbb{E}\left(e^{tT}\right).$$

We showed in Part 2 of the proof of Theorem 3.11 that $\gamma(t) \leq -c^{-1}(-t) = \mu t$ for all $t \in (-b_2, -b_1) = (-\infty, \infty)$. Because γ is convex and $\gamma(0) = -c^{-1}(-0) = \mu \cdot 0 = 0$, we must have that $\gamma(t) \equiv \mu t$, meaning $\gamma'(0) = \mu$.

We now aim to show that $\frac{T}{M} \to \mu$ in probability. To this end, let $\varepsilon > 0$. By applying Chebychev's inequality (2.2) for $t > 0$, it holds that

$$\mathbb{P}\left(\frac{T}{M} - \mu \geq \varepsilon\right) = \mathbb{P}(T \geq M(\mu+\varepsilon)) \leq e^{-tM(\mu+\varepsilon)}\mathbb{E}\left(e^{tT}\right).$$

Then

$$\begin{aligned}\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} - \mu \geq \varepsilon\right) &\leq -t\mu - t\varepsilon + \limsup_{M\to\infty} \frac{1}{M} \log \mathbb{E}\left(e^{tT}\right) \\ &= -t\mu - t\varepsilon + \gamma(t) \\ &= \underbrace{\gamma(t) - t\mu}_{=0} - t\varepsilon < 0.\end{aligned}$$

This means $\mathbb{P}\left(\frac{T}{M} - \mu \geq \varepsilon\right) \to 0$, as $M \to \infty$. Similarly, for $t < 0$, we get

$$\mathbb{P}\left(\frac{T}{M} - \mu \leq -\varepsilon\right) = \mathbb{P}(T \leq M(\mu-\varepsilon)) \leq e^{-tM(\mu-\varepsilon)}\mathbb{E}\left(e^{tT}\right),$$

and thus

$$\begin{aligned}\limsup_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} - \mu \leq -\varepsilon\right) &\leq -t\mu + t\varepsilon + \limsup_{M\to\infty} \frac{1}{M} \log \mathbb{E}\left(e^{tT}\right) \\ &= -t\mu + t\varepsilon + \gamma(t) \\ &= \underbrace{\gamma(t) - t\mu}_{=0} + t\varepsilon < 0,\end{aligned}$$

implying $\mathbb{P}\left(\frac{T}{M} - \mu \leq -\varepsilon\right) \to 0$, as $M \to \infty$. Thus

$$\mathbb{P}\left(\left|\frac{T}{M} - \mu\right| \geq \varepsilon\right) = \mathbb{P}\left(\frac{T}{M} - \mu \leq -\varepsilon\right) + \mathbb{P}\left(\frac{T}{M} - \mu \geq \varepsilon\right) \to 0,$$

as $M \to \infty$, for any $\varepsilon > 0$. This is convergence in probability.

Recall (3.45). If $\mu \in \mathcal{O}$, then $\mathbb{P}\left(\frac{T}{M} \in \mathcal{O}\right) \to 1$, as $M \to \infty$, and thus

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{O}\right) = 0 = -\Gamma^*(\mu) = -\inf_{x\in\mathcal{O}}\{\Gamma^*(x)\}.$$

On the other hand, if $\mu \notin \mathcal{O}$, then $\mathbb{P}\left(\frac{T}{M} \in \mathcal{O}\right) \to 0$, as $M \to \infty$, therefore

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in \mathcal{O}\right) = -\infty = -\inf_{x\in\mathcal{O}}\{\Gamma^*(x)\}.$$

In either case, we have (3.16).

Now that (3.16) is proven, what remains is to show (3.15). To this end, fix $x \in \mathbb{R}$ and $\varepsilon > 0$. We have, for all $t \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}\left(e^{tT}\mathbb{I}\{T < \infty\}\right) &\geq \mathbb{E}\left(e^{tT}\mathbb{I}\left\{\frac{T}{M} \in B_\varepsilon(x)\right\}\right) \\
&\geq \mathbb{E}\left(\exp\left(\inf_{y\in B_\varepsilon(x)} tyM\right)\mathbb{I}\left\{\frac{T}{M} \in B_\varepsilon(x)\right\}\right) \\
&= \exp\left(\inf_{y\in B_\varepsilon(x)} tyM\right)\mathbb{P}\left(\frac{T}{M} \in B_\varepsilon(x)\right).
\end{aligned}$$

Therefore, by the already proven (3.16),

$$\begin{aligned}
\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{E}\left(e^{tT}\mathbb{I}\{T < \infty\}\right) &\geq \liminf_{M\to\infty} \frac{1}{M} \inf_{y\in B_\varepsilon(x)} tyM + \liminf_{M\to\infty} \frac{1}{M} \log \mathbb{P}\left(\frac{T}{M} \in B_\varepsilon(y)\right) \\
&\geq \inf_{y\in B_\varepsilon(x)} ty - \inf_{y\in B_\varepsilon(x)} \Gamma^*(y) \\
&\geq tx - |t\varepsilon| - \Gamma^*(x) \\
&\to tx - \Gamma^*(x),
\end{aligned}$$

as $\varepsilon \to 0$. Now recall that by Lemma 3.25, Γ is proper convex and lower semi-continuous, so by Theorem 3.20, $(\Gamma^*)^* \equiv \Gamma$. Now as our choice of $x \in \mathbb{R}$ above was arbitrary, we have

$$\liminf_{M\to\infty} \frac{1}{M} \log \mathbb{E}\left(e^{tT}\mathbb{I}\{T < \infty\}\right) \geq \sup_{x\in\mathbb{R}}\{tx - \Gamma^*(x)\} = (\Gamma^*)^*(t) = \Gamma(t).$$

This is (3.15). □

4 Conclusions

The foundations of large deviations theory were developed by actuaries such as H. Cramér, studying risk and insurance from a mathematical perspective. Since then, it has grown into a field of its own, and many new important results have been developed. One of these is the Gärtner-Ellis theorem, presented here as Theorem 2.29. This result has even been expanded upon, to apply also to the continuous case under slightly different assumptions, see Theorem 4.5.20 in [2]. This result is stated in this paper in a slightly modified form as Theorem 3.24.

The connection between large deviations theory and actuarial science has not been undone by time, however, and more recent large deviations results are still frequently applied in the context of insurance. Two such results, both developed by H. Nyrhinen, were presented in this paper as Theorems 3.11 and 3.14. These two results were chosen partly because they depend on a result similar to the Gärtner-Ellis theorem, and partly because when put together, they yield a very similar result.

The Gärtner-Ellis theorem has later been expanded upon. Theorem 3.24 in this paper is a continuous version, and originates from Baldi [11]. Another extension of the Gärtner-Ellis theorem was developed by O'Brien and Vervaat [12], and later this extension was further strengthened by Comman [13].

The results by Nyrhinen which are presented in this paper have later been expanded upon by Nyrhinen himself, but their application can also be seen for example in Albrecher and Asmussen [14]. One can also see the results of [9] investigated via Monte Carlo-simulation using a Markov dependence structure in Albrecher and Kantor [15].

Finally, for a reference text on the current state-of-the-art of insurance mathematics in general see the book *Risk and Insurance* by S. Asmussen and M. Steffensen [16]. This book uses large deviations theory in several places as tools, but also uses methods from several other areas of mathematics.

Appendix of notations and definitions

$\mathbb{P}(\cdot)$	Probability	3
$\mathbb{E}(\cdot)$	Expectation	5
$\langle \cdot, \cdot \rangle$	Inner product	5
$\lVert \cdot \rVert$	Euclidian norm	5
$\bar{\mathcal{A}}$	Closure of set $\mathcal{A}$	5

References

[1] den Hollander, F. (2000). *Large Deviations.* American Mathematical Society.

[2] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications.* Second Edition, Springer.

[3] Gärtner, Jürgen. (1977). *On Large Deviations from an Invariant Measure.* Teor. Verojatnost. i Primenen. **22. No. 1**. 27-42.

[4] Ellis, Richard S. (1984). *Large Deviations for a General Class of Random Vectors.* Annals of Probability. **12. No. 1**. 1-12.

[5] Varadhan, S. R. S. (1966). *Asymptotic Probabilities and Differential Equations.* Communications on Pure and Applied Mathematics. **19**. 261-286.

[6] Cramér, H. (1938). *Sur un nouveau theorème-limite de la thèorie des probabilités.* Actualiés scientifiques et industrielles. **736**. Hermann & Cie. 5-23.

[7] Bourbaki, N. (1995). *Elements of Mathematics: General Topology.* Springer. Chapters 1-4.

[8] Rockafellar, R. T. (1972). *Convex Analysis.* Second Edition, Princeton University Press.

[9] Nyrhinen, H. (1998). *Rough descriptions of ruin for a general class of surplus processes.* Advances in Applied Probability. **30**. 1008-1026.

[10] Nyrhinen, H. (1999). *Large Deviations for the Time of Ruin.* Journal of Applied Probability. **36**. 733-746.

[11] Baldi, P. (1988). *Large deviations and stochastic homogenization.* Annali di Matematica Pura ed Applicata. **151**. 161-177.

[12] O'Brien, G. L. and Vervaat, W. (1995) *Compactness in the theory of large deviations.* Stochastic Processes and their Applications. **57**. 1-10.

[13] Comman, H. (2009). *Differentiability-free conditions on the free-energy function implying large deviations.* Confluentes Mathematici. **1. No. 2**. 1-15.

[14] Albrecher, H and Asmussen, S. (2006) *Ruin probabilities and aggregrate claims distributions for shot noise Cox processes.* Scandinavian Actuarial Journal. **2**. 86-110.

[15] Albrecher, H and Kantor, J. (2002) *Simulation of ruin probabilities for risk processes of Markovian type*. Monte Carlo Methods and Applications. **8. No. 2**. 111-127.

[16] Asmussen, S. and Steffensen, M. (2020). *Risk and Insurance*. Springer.