Università di Pisa
Dipartimento di Matematica
Corso di Laurea Triennale in Matematica

$\Gamma-$convergence and large deviations

Tesi di Laurea Triennale

Relatore
Dott. Dario Trevisan

Candidato
Vincenzo Mariani

Anno Accademico 2017/2018

Università di Pisa
Dipartimento di Matematica

Corso di Laurea Triennale in Matematica

$\Gamma-$convergence and large deviations

Tesi di
Vincenzo Mariani

Relatore: Dott. Dario Trevisan
Candidato: Vincenzo Mariani

Sessione di Laurea: 15 Febbraio 2019
Anno Accademico 2017/2018

Contents

Introduction

Γ-convergence is a notion of convergence for functionals introduced by Ennio de Giorgi. The theory of Γ-convergence is commonly recognized as a flexible tool for the description of the asymptotic behaviour of minimum problems in Calculus of Variations. Its strength is its adaptability, due to its being linked to no a priori assumption about the form of minimizers: the latter being in a sense automatically described by a process of optimization.

On the other hand the theory of Large Deviations concerns the asymptotic behaviour of remote tails of sequences of probability measures. Large Deviations theory formalizes the heuristic ideas of concentration of measures and makes more general statements about the notion of convergence of probability measures. Principles of Large Deviations may be applied to collect information of a probabilistic model, consequently they find their applications in information theory and mathematical physics. In physics, the best known application of large deviations theory emerges in thermodynamics and statistical mechanics (in connection with relating entropy with rate function).

In this work we will show how it is possible to find a rigorous connection between the two theories of Γ-convergence and Large Deviations. The main results will be presented, showing an application at the end. The thesis is widely based on the paper of Mauro Mariani (see [6]) *A Γ-convergence approach to large deviations.*

In the first chapter we explain the general setting and prove some basic statements that link the property of *tightness* of a sequence of probability measures with the equicoercivity of an associated family of Relative Entropy functionals. Moreover we find a connection between weak convergence in the space of probability measures and the Γ-lim of the entropy functionals.

In the second chapter we introduce the concept of Large Deviations, by proving a Large Deviations version of the previous statements: the *tightness* is replaced by *exponential tightness*, the functionals are weighted with an appropriate sequence of real numbers (the inverse of speed) and finally *weak convergence* is replaced by *Large Deviations bounds.*

The third chapter is dedicated to present a simple application of the conclusions showed in chapter 2, proving classical results of Large Deviations theory.

Notation

In propositions and theorems of the following chapters the statements concerning Large Deviations bounds, that express the probabilistic side of the medal, will be usually labeled by (P) or (P1), (P2); whereas the statements concerning Γ-

convergence and in particular concerning convergence of an entropy functionals sequence $(H_n)_n$, will be labeled with (H) or (H1), (H2). This kind of notation will be useful to quickly understand from an heuristic point of view the *direction* of the implications that we will prove.

Topological and set theory-notation

- Let X be a topological space and let A be a subset of X. $\mathring{A}$ is the interior of A. $\overline{A}$ denotes the closure of A with respect to the topology of X. $\partial A := \overline{A} \backslash \mathring{A}$.
- Let (X, d) be a metric space. Let r be a real positive number. Fixed $x \in X$ the symbol $B_r(x)$ denotes the ball centered in x with radius r i.e. the set of all $y \in X$ s.t. $d(x, y) < r$.
- Let (X, d) be a metric space and let A be a subset of X:

$$diam(A) = \sup\{d(x, y) \quad | \quad (x, y) \in A \times A\}$$

- $\mathbb{R}_0^+$ are non-negative real numbers.
- $\mathbb{R}^+$ are positive real numbers.
- $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$.
- $\mathbb{N}^+$ are positive natural numbers.

Probabilistic and analytical notation

- The notation $(a_n)_n$ denotes a sequence, indexed on positive natural numbers.
- Given a real valued sequence $(a_n)_n$ we denote

$$\underline{\lim}_n a_n := \liminf_{n \to +\infty} a_n \text{ and } \overline{\lim}_n a_n := \limsup_{n \to +\infty} a_n.$$

  Moreover, given a space X with a notion of convergence, and given a sequence $(x_n)_n$ with elements in X, we denote

$$\lim_n x_n = \lim_{n \to +\infty} x_n.$$

- Let (X, d) be a metric space. $\mathcal{P}(X)$ denotes the space of (posive) Borel probability measures.
- Let (X, d) be a metric space, $\nu \in \mathcal{P}(X)$ and let $f : X \to \mathbb{R}$ be a ν-Borel-measurable function. We denote

$$\nu(f) := \int_X f(x) d\nu(x) = \int_X f d\nu$$

- Let (X, d) be a metric space and let $(E_n)_n$ be a sequence of subsets of X. We denote $\sigma((E_n)_n)$ the σ-algebra generated by $(E_n)_n$.

- Let X be a topological space. We denote $\mathcal{B}(X)$ the Borel σ-algebra on X.
- Let X be a set and A a subset of X. We denote $\chi_A : X \to \mathbb{R}$ the indicator function
$$\chi_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ otherwise} \end{cases}$$
- We write $f \in C_b(X)$ if $f : X \to \mathbb{R}$ is continuous and uniformly bounded.
- Let X be a metric space. $\forall \mu \in \mathcal{P}(X) \quad \forall A \in \mathcal{B}(X) : \mu(A) > 0$ we denote $\mu^A(\cdot) = \mu(\cdot|A)$. For every $x \in X$, we define the Dirac measure as $\delta_x \in \mathcal{P}(X)$ where for any A Borel set
$$\delta_x(A) = \begin{cases} 0, & x \notin A; \\ 1, & x \in A. \end{cases}$$

Chapter 1

Preliminaries

In this chapter we are going to introduce the general setting used in the work, main definitions and the first results. In all the thesis (X, d) will be a separable and complete metric space. In particular we will show a first connection between tightness of a sequence of measures in $\mathcal{P}(X)$ and equicoercivity of entropy functionals, whereupon we are going to prove a relation between weak convergence in $\mathcal{P}(X)$ and Γ-liminf of the associated entropy functionals.

1.1 $\Gamma-$ convergence definitions

Let (E, d) be a metric space. We are going to use the following definitions.

Definition 1 (Lower semicontinuity)**.** A functional $I : E \longrightarrow [0, +\infty]$ is lower semicontinuos if and only if $\forall l \geqslant 0$ the set $\{x \in E | I(x) \leqslant l\}$ is closed. Equivalently given $(x_n)_n$ in E such that x_n converges to $x \in E$ then

$$\varliminf_n I(x_n) \geqslant I(x) \tag{1.1}$$

Definition 2 (Precompact)**.** Let E be a metric space. Let $Y \subseteq E$. A precompact subset Y of a topological space E is a set where any sequence in Y has a subsequence convergent in E.

Definition 3 (Coercivity)**.** A functional I is coercive if and only if $\forall l \geqslant 0$ the set $\{x \in E | I(x) \leqslant l\}$ is precompact.

Definition 4 (Equicoercivity)**.** Let (I_n) be a sequence of functionals $I_n : E \to [0, +\infty]$. The sequence is equicoercive on E if and only if $\forall l > 0$ the set

$$\bigcup_n \{x \in E | I_n(x) \leqslant l\}$$

is precompact.

Definition 5 ($\Gamma-$convergence)**.** Let E be a metric space and $F_n : E \to [0, +\infty)$ a sequence of functionals on E. Then F_n are said to Γ-converge to the Γ-limit $F : E \to [0, +\infty)$ if the following two conditions hold:

- lower bound inequality: for every sequence $(x_n)_n \subseteq E$ such that $x_n \to x$ as $n \to +\infty$,
$$F(x) \leqslant \varliminf_n F_n(x_n)$$
- upper bound inequality: $\forall x \in E \quad \exists (x_n)_n \subseteq E$ such that $x_n \to x$ as $n \to +\infty$
$$F(x) \geqslant \varlimsup_n F_n(x_n).$$

Definition 6 ($\Gamma-\varliminf$ and $\Gamma-\varlimsup$)**.** Let E be a metric space and $F_n : E \to [0, +\infty)$ a sequence of functionals on E. For $x \in E$ we define:

$$(\Gamma - \varliminf_n F_n)(x) := \inf\{\varliminf_n F_n(x_n)|(x_n)_n \text{ sequence in } E \text{ s. t. } x_n \to x\}$$

and

$$(\Gamma - \varlimsup_n F_n)(x) := \inf\{\varlimsup_n F_n(x_n)|(x_n)_n \text{ sequence in } E \text{ s. t. } x_n \to x\}$$

1.2 Weak topology on probability measures

The space $\mathcal{P}(X)$ of Borel probability measures on X will be equipped with the *weak* topology. All probability measures from now on will be elements of $\mathcal{P}(X)$. The weak topology on $\mathcal{P}(X)$ is the coarsest among all topologies that make continuous all the real functions
$$\mu \longmapsto \int_X f d\mu$$
where $f \in C_b(X)$. The convergence of measures with respect to this topology is said weak convergence. In order that a sequence $(\mu_n)_n$ of measure weakly converges towards a measure μ it is necessary and sufficient that

$$\int_X f d\mu = \lim_n \int_X f d\mu_n$$

for all real valued, continuous and bounded function f on X.

In the following lines we are going to state some theorems about weak convergence, without proof[1]. The first theorem is about a characterization of weak convergence.

Theorem 1. *Let μ be a measure and $(\mu_n)_n$ a sequence of measures on $(X, \mathcal{B}(X))$. The following facts are equivalent.*

(a) The sequence $(\mu_n)_n$ weakly converges to μ.

(b) It holds
$$\int_X f d\mu = \lim_n \int_X f d\mu_n$$
for every bounded and Lipschitz function $f : X \to \mathbb{R}$.

[1]A proof can be found in [2].

(c) It holds

$$\int_X f d\mu \leqslant \varliminf_n \int_X f d\mu_n$$

for every lower bounded and lower semicontinuous function $f : X \to \mathbb{R}$.

(d) It holds

$$\int_X f d\mu \geqslant \varlimsup_n \int_X f d\mu_n$$

for every upper bounded and upper semicontinuous function $f : X \to \mathbb{R}$.

(e) It holds

$$\int_X f d\mu = \lim_n \int_X f d\mu_n$$

for every bounded and Borel function $f : X \to \mathbb{R}$, μ-almost evereywhere continuous.

The second theorem is a version of the previous one, where sets play the role of functions.

Theorem 2. *Let μ be a measure and $(\mu_n)_n$ a sequence of measures on $(X, \mathcal{B}(X))$. The following facts are equivalent.*

(a) The sequence $(\mu_n)_n$ weakly converges to μ.

(b) We have

$$\mu(X) = \lim_n \mu_n(X) \text{ and } \mu(O) \leqslant \varliminf_n \mu_n(O)$$

for every $O \subseteq X$ open.

(c) We have

$$\mu(X) = \lim_n \mu_n(X) \text{ and } \mu(C) \geqslant \varlimsup_n \mu_n(C)$$

for every $C \subseteq X$ closed.

(d) We have

$$\mu(B) = \lim_n \mu_n(B)$$

for every $B \in \mathcal{B}(X)$ such that $\mu(\partial B) = 0$.

Definition 7 (Tightness)**.** A probability measure μ is said tight if $\forall \varepsilon > 0$ exists a compact K of X such that $\mu(K^c) < \varepsilon$.

This condition of tightness equals to impose that the measure μ is concentrated on an appropriate countable union of compact sets.

Theorem 3 (Ulam)**.** *If X is a separable and complete metric space, then every probability measure is tight.*

So we could always assume the tightness for every measure, because in all the thesis X will be a separable and complete metric space.

The following definition and the next theorem will be extensively used in the proofs of the main results, so it is essential to focus on them.

Definition 8 (Tightness for a family of measure). Given $\mathcal{H} \subseteq \mathcal{P}(X)$ we say that $\mathcal{H}$ is tight if $\forall \varepsilon > 0$ exists a compact $K \subseteq X$ such that

$$\sup_{\mu \in \mathcal{H}} \mu(K^c) \leqslant \varepsilon$$

Theorem 4 (Prokhorov). *Given $\mathcal{H} \subseteq \mathcal{P}(X)$. $\mathcal{H}$ is tight if and only if it is sequentially compact in the space $\mathcal{P}(X)$ equipped with the topology of weak convergence.*

For a proof of Prokhorov's theorem and Ulam's theorem see [2].

1.3 Large Deviations definitions

Definition 9 (Large Deviations bounds). Let $I : X \to [0, +\infty]$ be a lower semi-continuous functional and let $(a_n)_n$ be a sequence of positive real numbers such that $\lim_n a_n = +\infty$. The sequence $(\mu_n)_n \subseteq \mathcal{P}(X)$ satisfies:

- a Large Deviations Lower Bound with speed $(a_n)_n$ and rate I, if and only if $\forall O \subseteq X$ open we have

$$\underline{\lim}_n \frac{1}{a_n} \log \mu_n(O) \geqslant - \inf_{x \in O} I(x) \tag{1.2}$$

- a Large Deviations Weak Upper Bound with speed $(a_n)_n$ and rate I, if and only if $\forall K \subseteq X$ compact we have

$$\overline{\lim}_n \frac{1}{a_n} \log \mu_n(K) \leqslant - \inf_{x \in K} I(x) \tag{1.3}$$

- a Large Deviations Upper Bound with speed $(a_n)_n$ and rate I, if and only if $\forall C \subseteq X$ closed we have

$$\overline{\lim}_n \frac{1}{a_n} \log \mu_n(C) \leqslant - \inf_{x \in C} I(x) \tag{1.4}$$

Definition 10 (Relative Entropy). [2] Given $\mu, \nu \in \mathcal{P}(X)$ and $\mathcal{F} \subset \mathcal{B}(X)$ a σ-algebra, the relative entropy of ν with respect to μ on $\mathcal{B}$ is defined as

$$H_{\mathcal{F}}(\nu|\mu) := \sup_{\varphi} \{\nu(\varphi) - \log \mu(e^{\varphi})\} \in [0, +\infty] \tag{1.5}$$

where the supremum runs over the bounded $\mathcal{F}$-measurable functions φ on X. For a fixed μ, $H_{\mathcal{F}}(\cdot|\mu)$ is a convex functional on $\mathcal{P}(X)$. If $\mathcal{F} = \mathcal{B}(X)$ the subindex will be dropped hereafter. In such case $H(\cdot|\mu)$ is also lower semicontinuous and coercive on $\mathcal{P}(X)$. Given $(\mu_n)_n$ sequence in $\mathcal{P}(X)$ we will define

$$H_n(\nu) := H(\nu|\mu_n)$$

[2] The relative entropy is also known as *Kullback–Leibler divergence* in information theory and mathematical statistics. It is a measure of how one probability distribution is different from a second reference probability distribution. The Kullback–Leibler divergence was introduced by Solomon Kullback and Richard Leibler in 1951 (see [5]) as the directed divergence between two distributions.

Moreover, we can write the entropy functional in a more explicit way, very helpful in some steps later.

$$H(\nu|\mu) = \begin{cases} \int_X \frac{d\nu}{d\mu} \log\left(\frac{d\nu}{d\mu}\right) d\mu & \text{if } \nu \ll \mu \\ +\infty & \text{otherwise} \end{cases} \tag{1.6}$$

Remark 1. The two formulas of Relative Entropy presented above are equivalent. We can prove the equivalence showing that the two inequalities in 1.5 hold with respect to the integral formulation. Without loss of generality we assume $\nu \ll \mu$. In the following lines we use $D(\nu|\mu)$ to denote the 1.6 form.

- Simply
$$\varphi(x) = \log\left(\frac{d\nu}{d\mu}(x)\right) \Rightarrow H(\nu|\mu) \geqslant D(\nu|\mu)$$

- Fix φ, we define
$$\mu^{\varphi}(dx) = \frac{e^{\varphi}\mu(dx)}{\mu(e^{\varphi})}$$
and of course $\mu^{\varphi} \ll \mu$.
$$\nu(\varphi) - \log(\mu(e^{\varphi})) = \mu\left(\log\left(\frac{d\mu^{\varphi}}{d\mu}\right)\right) = D(\nu|\mu) - D(\nu|\mu^{\varphi}) \leqslant D(\nu|\mu).$$
Hence $H(\nu|\mu) \leqslant D(\nu|\mu)$.

Moreover because of the strict convexity of the function $g : x \longmapsto x \log x$ defined on $\mathbb{R}^+$, we can use Jensen's inequality to show that $H(\nu|\mu) \geqslant 0$. Indeed

$$H(\nu|\mu) = \mu\left(g\left(\frac{d\nu}{d\mu}\right)\right) \geqslant g\left(\mu\left(\frac{d\nu}{d\mu}\right)\right) = g(1) = 0$$

and $H(\nu|\mu) = 0$ if and only if $\nu = \mu$ almost everywhere.

Now let us prove convexity. Let $\nu_1, \nu_2, \mu \in \mathcal{P}(X)$ and $\lambda \in \mathbb{R}$: for g convexity we have:

$$\begin{aligned} H(\lambda\nu_1 + (1-\lambda)\nu_2|\mu) &= \mu\left(g\left(\lambda\frac{d\nu_1}{d\mu} + (1-\lambda)\frac{d\nu_2}{d\mu}\right)\right) \leqslant \\ &\leqslant \mu\left(\lambda g\left(\frac{d\nu_1}{d\mu}\right) + (1-\lambda)g\left(\frac{d\nu_2}{d\mu}\right)\right) = \\ &= \lambda\mu\left(g\left(\frac{d\nu_1}{d\mu}\right)\right) + (1-\lambda)\mu\left(g\left(\frac{d\nu_2}{d\mu}\right)\right) \\ &= \lambda H(\nu_1|\mu) + (1-\lambda)H(\nu_2|\mu) \end{aligned}$$

With the following proposition we are going to prove lower semicontinuity of the relative entropy functional, when μ is fixed.

Proposition 5. *Let $\mu, \nu \in \mathcal{P}(X)$ and $(\mu_n)_n$, $(\nu_n)_n$ two sequence of probability measures such that $\mu_n \to \mu$ and $\nu_n \to \nu$ weakly in $\mathcal{P}(X)$ thus*

$$\underline{\lim}_n H(\nu_n|\mu_n) \geqslant H(\nu|\mu)$$

Proof. From 1.5 we have

$$H(\nu_n|\mu_n) := \sup_{\varphi}\{\nu_n(\varphi) - \log \mu_n(e^{\varphi})\}$$

Moreover $\mu_n \to \mu$ if and only if $\mu(\varphi) = \lim_n \mu_n(\varphi)$ for any bounded φ thanks to point (b) of Theorem 1. Taking the liminf concludes the proof. □

Lemma 6. *Given A Borel set we get*

$$\nu(A) \leqslant \frac{\log 2 + H(\nu|\mu)}{\log\left(1 + \frac{1}{\mu(A)}\right)} \tag{1.7}$$

if $H(\nu|\mu) < +\infty$ *and* $\mu(A) > 0$.

Proof. Choosing $\varphi(x) = \log(1 + \mu(A))\chi_A(x) + \log(\mu(A))\chi_{A^c}(x)$ in the definition of entropy, we get the thesis. Fixing

$$\varphi(x) = c_1\chi_A(x) + c_2\chi_{A^c}(x)$$

with $c_1 \neq c_2$ real numbers, from 1.5 we get easily that

$$\begin{aligned} c_1\nu(A) + c_2(1 - \nu(A)) &\leqslant H(\nu|\mu) + \log\left(e^{c_2}(1 - \mu(A)) + \mu(A)e^{c_1}\right) \\ \nu(A)(c_1 - c_2) &\leqslant H(\nu|\mu) + \log\left(e^{c_2}(1 - \mu(A)) + \mu(A)e^{c_1}\right) - c_2 \\ \nu(A) &\leqslant \frac{H(\nu|\mu) + \log\left(e^{c_2}(1 - \mu(A)) + \mu(A)e^{c_1}\right) - c_2}{(c_1 - c_2)} \end{aligned}$$

Choosing $c_1 = \log(1 + \mu(A))$ and $c_2 = \log(\mu(A))$ we have

$$c_1 - c_2 = \log\left(1 + \frac{1}{\mu(A)}\right)$$

whereas $e^{c_2} = \mu(A)$ and $e^{c_1} = 1 + \mu(A)$ thus

$$\log\left(\mu(A)(1 - \mu(A)) + \mu(A)e^{\log(1+\mu(A))}\right) - \log(\mu(A)) = \log 2$$

so the inequality.

□

1.4 Weak convergence and Entropy

Proposition 7. *The following are equivalent.*

(*P*) $(\mu_n)_n$ *is tight in* $\mathcal{P}(X)$.

(*H*) $(H(\cdot|\mu_n))_n$ *is equicoercive on* $\mathcal{P}(X)$

Proof.

$(H) \Rightarrow (P)$ Since $H(\mu_n|\mu_n) \leqslant 0 \quad \forall n \in \mathbb{N}$, the sequence $(\mu_n)_n$ is contained in $\{\nu|H(\nu|\mu_n) \leqslant 0\}$ hence precompact and thus tight for Prokhorov's theorem.

$(P) \Rightarrow (H)$ Let $(\nu_n)_n$ be a subsequence in $\mathcal{P}(X)$ such that $\overline{\lim_n} H(\nu_n|\mu_n) < +\infty$.

To prove the thesis we need to show that

$$\forall l \geqslant 0 \quad \bigcup_{n \in \mathbb{N}} \{\nu \in \mathcal{P}(X) | H(\nu|\mu_n) \leqslant l\}$$

is precompact i.e. the set is tight. Namely we will prove precompactness of a sequence in

$$\{\nu \in \mathcal{P}(X) | H(\nu|\mu_n) \leqslant l\}.$$

Let us use the notation $H_n(\nu) = H(\nu|\mu_n)$. Thus

$$\begin{aligned}&H_n(\nu) = \sup_{\varphi}\{\nu(\varphi) - \log \mu_n(e^{\varphi})\} \leqslant l \Rightarrow \\ &\forall \varphi \quad \nu(\varphi) \leqslant l + \log \mu_n(e^{\varphi})\end{aligned}$$

A good choice of φ could be

$$\varphi(x) := \begin{cases} 0 & \text{if } x \in K \\ M & \text{if } x \in K^c \end{cases}$$

where $M \in \mathbb{R}^+$ will be specified later. This gives

$$\begin{aligned}M\nu(K^c) &\leqslant l + \log(\mu(K) + \mu(K^c)e^M) \\ \nu(K^c) &\leqslant \frac{l}{M} + \frac{\log(1 - \mu(K^c) + \mu(K^c)e^M)}{M} \\ \nu(K^c) &\leqslant \frac{l}{M} + \frac{\log(1 + \mu(K^c)(e^M - 1))}{M}\end{aligned}$$

Given $\varepsilon > 0$ we set $M = \frac{2l}{\varepsilon}$ so

$$\nu(K^c) \leqslant \frac{\varepsilon}{2} + \frac{\log(1 + \mu(K^c)(e^M - 1))}{M}$$

Now we use the hypothesis of tightness of $(\mu_n)_n$. Indeed

$$\forall \eta > 0 \quad \exists K := K_\eta : \mu_n(K^c) < \eta$$

thus choosing $\eta < \frac{Me^{\frac{\varepsilon}{2}} - 1}{e^M - 1}$ we will obtain $\frac{\log(1+\mu(K^c)(e^M-1))}{M} \leqslant \frac{\varepsilon}{2}$ and so the thesis. □

Before proving the next proposition it is appropriate to build a particular partitions of space X and to do some observations.

Let $(E^i)_{i=0}^N$ be a finite partition of X. Let us call $\mathcal{G} = \sigma((E^i)_{i=0}^N)$ the σ-algebra generated by the partition. So applying the integral definition of the entropy functional we can state:

$$H_{\mathcal{G}}(\nu|\mu) = \sum_{i=0}^{N} \nu(E^i) \log \frac{\nu(E^i)}{\mu(E^i)}$$

understanding $\nu(A) \log \frac{\nu(A)}{\mu(A)} = 0$ if $\nu(A) = 0$ and $\nu(A) \log \frac{\nu(A)}{\mu(A)} = +\infty$ if $\mu(A) = 0$ but $\nu(A) > 0$.

Lemma 8. *Let $\mu, \nu \in \mathcal{P}(X)$, and $(K_l)_{l\in\mathbb{N}}$ a sequence of compact subsets of X such that*

$$\lim_l \mu(K_l) = 1.$$

For every $l \in \mathbb{N}$ and $\delta > 0$, exists a finite collection of Borel subsets of X that we denote with $(E^i_{\delta,l})_{i=1}^{N_{\delta,l}}$ such that:

(i) if $i \neq i'$, then $E^i_{\delta,l} \cap E^{i'}_{\delta,l} = \emptyset$ and

$$K_l \subseteq \bigcup_{i=1}^{N_{\delta,l}} E^i_{\delta,l}$$

(ii) $\forall i \in \{1, ..., N_{\delta,l}\}$ we have diameter$(E^i_{\delta,l}) \leqslant \delta$

(iii) $\forall i \in \{1, ..., N_{\delta,l}\}$ we have $\mu(\partial E^i_{\delta,l}) = \nu(\partial E^i_{\delta,l}) = 0$

(iv) $\forall i \in \{1, ..., N_{\delta,l}\} \quad \mathring{E}^i_{\delta,l} \neq \emptyset$

(v) setting

$$E^0_{\delta,l} = X \setminus \bigcup_{i=1}^{N_{\delta,l}} E^i_{\delta,l}$$

we can assume without loss of generality that if $\delta \leqslant \delta'$ and $l \geqslant l'$ then $(E^i_{\delta,l})_{i=0}^{N_{\delta,l}}$ is finer[3] than $(E^i_{\delta,l})_{i=0}^{N_{\delta',l'}}$

(vi) let $\mathcal{G}_{\delta,l} = \sigma((E^i_{\delta,l})_{i=0}^{N_{\delta,l}})$ so

$$\lim_l \lim_\delta H_{\mathcal{G}_{\delta,l}}(\nu|\mu) = H(\nu|\mu)$$

Proof. Fixing $\delta > 0$ and $l \in \mathbb{N}$ we can take the open cover $\{B_{\frac{\delta}{2}}(x) | x \in K_l\}$ of K_l and because it is compact, we can extract a finite subcover. Thus it exists $N_{\delta,l} \in \mathbb{N}$ and exist $x_i \in K_l \quad \forall i \in \{1, \ldots, N_{\delta,l}\}$ for which

$$K_l \subseteq \bigcup_{i=1}^{N_{\delta,l}} B_{\frac{\delta}{2}}(x_i)$$

Setting $m = \min_{1 \leqslant i < j \leqslant N_{\delta,l}} d(x_i, x_j)$, let $r \in \mathbb{R}^+$ be such that $r \leqslant \min\{\frac{\delta}{2}, m\}$; so there exists $\delta' \in (\frac{\delta}{2}, \frac{\delta}{2} + r)$ such that $\mu(\partial B_{\delta'}(x_i)) = \nu(\partial B_{\delta'}(x_i)) = 0 \quad \forall i \in \{1, \ldots, N_{\delta,l}\}$. Indeed the set $\{\delta' | \mu(\partial B_{\delta'}(x_i)) > 0\}$ is countable.

Let us define the partition. We take

$$\begin{cases} E^1_{\delta,l} = B_{\delta'}(x_1) \\ E^i_{\delta,l} = B_{\delta'}(x_i) \setminus \bigcup_{j<i} E^j_{\delta,l} \text{ for } i > 1 \end{cases}$$

[3] A partition α of a set X is a refinement of a partition ρ of X -and we say that α is finer than ρ and that ρ is coarser than α- if every element of α is a subset of some element of ρ. Roughly speaking, this means that α is a further fragmentation of ρ.

Property (i) is trivial for construction. Property (ii) is really easy to check: $\forall i \in \{1, ..., N_{\delta,l}\}$

$$\mathrm{diam}(E^i_{\delta,l}) = \mathrm{diam}\left(B_{\delta'}(x_i)\setminus \bigcup_{j<i} E^j_{\delta,l}\right) \leqslant \mathrm{diam}(B_{\delta'}(x_i)) \leqslant 2\delta' \leqslant \delta$$

Property (iii) is due to $\mu(\partial B_{\delta'}(x_i)) = \nu(\partial B_{\delta'}(x_i)) = 0$. Property (iv) is due to the fact that balls have nonempty interior.

Properties (v): if $(E^i_{\delta,l})_{i=0}^{N_{\delta',l'}}$ is not finer than $(E^i_{\delta,l})_{i=0}^{N_{\delta,l}}$ we could intersect the sets $E^i_{\delta,l}$ of the two covers (of K_l) obtaining a new one with the properties from (i) to (v). Property (vi) is a corollary of Proposition 16 in the appendix.

□

Proposition 9. *The following are equivalent. Given a sequence $(\mu_n)_n$ and given μ in $\mathcal{P}(X)$,*

$(P1)$ $\mu_n \to \mu$ *in* $\mathcal{P}(X)$

$(P2)$ *For each sequence $(\varphi_n)_n$ of Borel measurable functions $\varphi_n : X \to \mathbb{R} \cup \{+\infty\}$ bounded from below we have*

$$\varliminf_n \mu_n(\varphi_n) \geqslant \mu(\Gamma - \varliminf_n(\varphi_n)) \tag{1.8}$$

(H) $H(\nu|\mu) = (\Gamma - \varliminf_n H_n)(\nu)$

Proof. The implications will be proved as follows: (P1)⇒(P2)⇒(H)⇒(P1).

(P1)⇒(P2) Let us assume $\mu_n \to \mu$ in $\mathcal{P}(X)$ and consider the construction of $(E^i_{\delta,l})_{i=0}^{N_{\delta,l}}$ as in Lemma 8 with $\nu = \mu$. Now consider a generic sequence $(\varphi_n)_n$, $\varphi_n : X \to \mathbb{R} \cup \{+\infty\}$ of Borel functions bounded from below. Fixed $x \in X$ we define

$$\varphi_{n,\delta,l}(x) := \inf_{y \in E^i_{\delta,l}} \varphi_n(y) \text{ if } x \in E^i_{\delta,l}$$

and

$$\varphi_{\delta,l}(x) := \varliminf_n \varphi_{n,\delta,l}(x) = \varliminf_n \inf_{y \in E^i_{\delta,l}} \varphi_n(y)$$

From bullets (iii), (iv) and (v) of Lemma 8 it follows that

$$\mu(\bigcup_{l\in\mathbb{N}^+,\delta>0} \bigcup_{i=1}^{N_{\delta,l}} \partial E^i_{\delta,l}) = 0 \text{ and } \lim_l \lim_\delta \sum_{i=0}^{N_{\delta,l}} \mu(E^i_{\delta,l}) = 0$$

where the first equality holds for (v) and the second one for (ii). Moreover if $x \notin \bigcup_{l\in\mathbb{N}^+,\delta>0} \bigcup_{i=1}^{N_{\delta,l}} \partial E^i_{\delta,l}$ it holds (see Proposition 15 in Appendix A)

$$\lim_{l\to+\infty} \lim_{\delta\to 0} \varphi_{\delta,l}(x) = (\Gamma - \varliminf_n \varphi_n)(x). \tag{1.9}$$

Thanks to point (v) of Lemma 8 the limit is monotone increasing and it holds μ-a.e., thus by monotone convergence

$$\mu(\Gamma - \varliminf_n \varphi_n) = \lim_l \lim_\delta \mu(\varphi_{\delta,l})$$

Inasmuch as $\varphi_{n,\delta,l}$ takes the same values $\forall x \in \overset{\circ}{E}{}^i_{\delta,l}$ it follows

$$\lim_l \lim_\delta \mu(\varphi_{\delta,l}) = \lim_l \lim_\delta \sum_{i=0}^{N_{\delta,l}} (\mu(E^i_{\delta,l}) \varliminf_n \inf_{y \in E^i_{\delta,l}} \varphi_n(y))$$

The right hand side of the last equality is smaller than

$$\lim_l \lim_\delta \varliminf_n \sum_{i=0}^{N_{\delta,l}} (\mu(E^i_{\delta,l}) \inf_{y \in E^i_{\delta,l}} \varphi_n(y)) \leqslant \varliminf_n \mu_n(\varphi_n)$$

Thus $\mu(\Gamma - \varliminf_n \varphi_n) \leqslant \varliminf_n \mu_n(\varphi_n)$ as wanted.

(P2)⇒(H) We want to show that $H(\nu|\mu) = (\Gamma - \lim_n H_n)(\nu)$ so we will proceed as follows: first of all we will prove that if $\nu_n \to \nu$ in $\mathcal{P}(X)$, then

$$\varliminf_n H(\nu_n|\mu_n) \geqslant H(\nu|\mu)$$

whereupon we will show that given $\nu \in \mathcal{P}(X)$ it exists $(\nu_n)_n \subseteq \mathcal{P}(X)$ that holds

$$\varlimsup_n H(\nu_n|\mu_n) \leqslant H(\nu|\mu)$$

From hypothesis of (P2), let us choose $\varphi_n = \varphi \in C_b(X) \quad \forall n \in \mathbb{N}$ in (1.8). So $\mu_n(\varphi) \to \mu(\varphi)$: $\Gamma - \varliminf \varphi_n = \varphi$ and we get the convergence replacing $-\varphi$ to φ. Fixed $\nu \in \mathcal{P}(X)$, we consider an arbitrary sequence in $(\nu_n)_n \subseteq \mathcal{P}(X)$ such that $\nu_n \to \nu$ in $\mathcal{P}(X)$.

$$\begin{aligned}
\varliminf_n H(\nu_n|\mu_n) &= \varliminf_n \sup_{\varphi \in C_b(X)} \{\nu_n(\varphi) - \log \mu_n(e^\varphi)\} \geqslant \\
&\geqslant \sup_{\varphi \in C_b(X)} \varliminf_n \{\nu_n(\varphi) - \log \mu_n(e^\varphi)\} = \\
&= \sup_{\varphi \in C_b(X)} \{\nu(\varphi) - \log \mu(e^\varphi)\} = H(\nu|\mu)
\end{aligned}$$

thus the first inequality is proved. Without loss of generality the Γ-limsup inequality can be proved for $\nu \ll \mu$ i. e. with $H(\nu|\mu) < +\infty$. Let $(E^i_{\delta,l})_{i=0}^{N_{\delta,l}}$ be as in Lemma 8. We fix $\delta, l > 0$ and for n large enough we define $\nu_{n,\delta,l} \in \mathcal{P}(X)$ as

$$\nu_{n,\delta,l}(A) = \sum_{i=0}^{N_{\delta,l}} \nu(E^i_{\delta,l}) \frac{\mu_n(A \cap E^i_{\delta,l})}{\mu_n(E^i_{\delta,l})}$$

with $\nu(E^i_{\delta,l}) = 0$ if $\mu_n(E^i_{\delta,l}) = 0$. So

$$\lim_l \lim_\delta \lim_n \nu_{n,\delta,l} = \nu$$

We saw in the previous section that $H(\nu_{n,\delta,l}|\mu_n) = H_{\mathcal{G}_{\delta,l}}(\nu|\mu_n)$. We will prove

$$\varlimsup_n H(\nu_{n,\delta,l}|\mu_n) \leqslant H(\nu|\mu)$$

so the Γ-limsup inequality.

In fact

$$\begin{aligned}
\overline{\lim_n} H(\nu_{n,\delta,l}|\mu_n) &= \overline{\lim_n} H_{\mathcal{G}_{\delta,l}}(\nu|\mu_n) \\
&= \overline{\lim_n} \sum_{i=0}^{N_{\delta,l}} \int_{E^i_{\delta,l}} \log\left(\frac{d\nu_{n,\delta,l}}{d\mu_n}\right) d\nu_{n,\delta,l} \\
&= \overline{\lim_n} \sum_{i=0}^{N_{\delta,l}} \nu_n(E^i_{\delta,l}) \log\left(\frac{\nu(E^i_{\delta,l})}{\mu_n(E^i_{\delta,l})}\right) \\
&= H_{\mathcal{G}_{\delta,l}}(\nu|\mu) \leqslant H(\nu|\mu).
\end{aligned}$$

Thus we can state that there exist sequences $(\delta_n)_n$ and $(l_n)_n$ such that $\nu_{n,\delta_n,l_n} \to \nu$ and $\overline{\lim}_n H(\nu_n|\mu_n) \leqslant H(\nu|\mu)$. Thus the Γ-limsup inequality holds, so the thesis.

(H)⇒(P1) For Prokhorov's theorem if we show that the sequence $(\mu_n)_n$ is tight, so it is precompact in $\mathcal{P}(X)$ thus we can apply the result of convergence on minimizers. For hypothesis of $\Gamma-$convergence and precompactness, converging sequence of minimizers of H_n converge to minimizers of the $\Gamma-$limit H. In particular μ_n is the unique minimizer of H_n and μ is the unique minimizer of H, so we would have the result (P1). Let us prove tightness. (H) implies the existence of a recovery sequence of probability measures $(\nu_n)_n$ converging to μ and such that it holds the limsup inequality, i.e.

$$\overline{\lim_n} H(\nu_n|\mu_n) \leqslant H(\mu|\mu)$$

thus

$$\lim_n H(\nu_n|\mu_n) = 0$$

and $(\nu_n)_n$ is tight (for Prokhorov's theorem, in the inverse sense respect to the last use).

Reversing now inequality (1.7) we obtain (where A is a Borel set)

$$\frac{1}{\mu(A)} \leqslant e^{\left(\frac{\log 2 + H(\nu|\mu)}{\nu(A)}\right)} - 1$$

and so

$$\mu(A) \geqslant \frac{1}{e^{\left(\frac{\log 2 + H(\nu|\mu)}{\nu(A)}\right)} - 1}$$

Let us note with some easy reckoning that, given $\varepsilon > 0$, the right hand side of the last written inequality is $\geqslant 1 - \varepsilon$ if

$$e^{1+\frac{1}{1-\varepsilon}} \geqslant \frac{\log 2 + H(\nu|\mu)}{\nu(A)}$$

i.e.

$$\nu(A) \geqslant \frac{\log 2 + H(\nu|\mu)}{e^{1+\frac{1}{1-\varepsilon}}}$$

Setting $\nu = \nu_n$ and $\mu = \mu_n$ in the formulas due to inequality 1.7 we can say, for tightness of $(\nu_n)_n$ and for n large enough exists K_ε such that

$$\nu_n(K_\varepsilon) \geqslant \frac{\log 2 + H(\nu_n|\mu_n)}{e^{1+\frac{1}{1-\varepsilon}}}$$

thus

$$\mu_n(K_\varepsilon) \geqslant 1 - \varepsilon \quad \forall n \geqslant n_0$$

so the tightness of $(\mu_n)_n$ is showed and the thesis proved.

□

Chapter 2

The connection between Large Deviations and Γ-convergence

In this chapter we will state and prove first of all the Large Deviations version of Proposition 7 and Proposition 9.

An interesting observation to focus on is that the Proposition 9 will have two Large Deviations versions. The first one using a Large Deviations lower bound and the second one using a Large Deviations weak upper bound. At the same time the (H) point of Proposition 9 will be splint into (H1) and (H2), or rather a version fixing $x \in X$ and a version fixing $\nu \in \mathcal{P}(X)$.

The (P) condition will split into (P1) and (P2) in the second theorem of this chapter, and from (P1) to (P4) in the third one.

Hereafter let $a = (a_n)_n$ be a sequence of strictly positive real numbers such that

$$\lim_n a_n = +\infty$$

let $(\mu_n)_n$ be a sequence of probability measures in $\mathcal{P}(X)$ and $I : X \to [0, +\infty]$ a measurable lower semicontinuous functional.

In order to have a Large Deviations version, we have to define a *new*, indeed, *weighted functional* of entropy, starting from the one introduced in the previous chapter.

2.1 Exponential tightness and entropy functionals

Definition 11 (Weighted Relative Entropy)**.** Given $a = (a_n)_n$ and $(\mu_n)_n$ as above, we define $H_n^a : \mathcal{P}(X) \to [0, +\infty]$ as follows

$$H_n^a(\nu) := \frac{1}{a_n} H(\nu|\mu_n)$$

What we stated before with tightness, now will become *exponentially tightness*. The weak convergence in $\mathcal{P}(X)$ will be the realization of a Large Deviations Principle. The idea of exponential tightness is a strengthening of classical tightness.

Definition 12 (Exponential tightness)**.** Let X be as above. A sequence of probability measures in $\mathcal{P}(X)$, $(\mu_n)_n$ is said exponentially tight if $\forall\varepsilon > 0$ exists K_ε compact of X such that

$$\overline{\lim_n} \frac{1}{n} \log \mu_n(K_\varepsilon^c) \leqslant -\varepsilon.$$

We can also give a definition of exponential tightness with a speed $(a_n)_n$, more interesting of a *fixed* velocity as in the classical exponential tightness definition.

Definition 13 (Exponential tightness with speed $(a_n)_n$)**.** Let X and $(a_n)_n$ be as above. Let $(\mu_n)_n$ be a sequence of probability measures in $\mathcal{P}(X)$, so $(\mu_n)_n$ is said exponentially tight with speed $(a_n)_n$ if $\forall\varepsilon > 0$ exists K_ε compact of X such that

$$\overline{\lim_n} \frac{1}{a_n} \log \mu_n(K_\varepsilon^c) \leqslant -\varepsilon.$$

Let us start with the analogous of Proposition 7.

Theorem 10. *The following are equivalent:*

- (P) $(\mu_n)_n$ *is exponentially tight with speed* $(a_n)_n$.
- (H) (H_n^a) *is equicoercive on* $\mathcal{P}(X)$

Proof.

(P)⇒(H) If I show that $\forall M \geqslant 0$ the set

$$\{\nu \in \mathcal{P}(X) | H_n^a(\nu) \leqslant M\}$$

is tight $\forall n \in \mathbb{N}$, by Prokhorov's theorem is precompact, thus H_n^a is equicoercive by definition. We will prove that given $\eta > 0$ there exists a compact $\tilde{K}_\eta$ for which $\nu(\tilde{K}_\eta) < \eta$ for every ν in the set. Since

$$\begin{aligned}
&H_n^a(\nu) = \frac{1}{a_n} \sup_\varphi \{\nu(\varphi) - \log \mu_n(e^\varphi)\} \leqslant M \Rightarrow \\
&\forall\varphi \quad \nu(\varphi) - \log \mu_n(e^\varphi) \leqslant M a_n \\
&\forall\varphi \quad \nu(\varphi) \leqslant M a_n + \log \mu_n(e^\varphi)
\end{aligned}$$

A good choice of φ could be, depending on n,

$$\varphi_n(x) := \begin{cases} 0 & \text{if } x \in K \\ M_n & \text{if } x \in K^c \end{cases}$$

where $M_n \in \mathbb{R}^+$ and the compact set K (not depending on n) will be specified later.

So

$$\begin{aligned}
M_n \nu(K^c) &\leqslant M a_n + \log(\mu_n(K) + \mu_n(K^c) e^{M_n}) \\
\nu(K^c) &\leqslant \frac{M a_n}{M_n} + \frac{\log(1 - \mu_n(K^c) + \mu_n(K^c) e^M)}{M_n} \\
\nu(K^c) &\leqslant \frac{M a_n}{M_n} + \frac{\log(1 + \mu_n(K^c)(e^{M_n} - 1))}{M_n}
\end{aligned}$$

Given $\eta > 0$ we set $M_n = \frac{2Ma_n}{\eta}$ so

$$\nu(K^c) \leqslant \frac{\eta}{2} + \frac{\log(1 + \mu(K^c)(e_n^M - 1))}{M_n}$$

Now we use the hypothesis of exponential tightness of $(\mu_n)_n$. Indeed

$$\forall \varepsilon > 0 \quad \exists K := K_\varepsilon : \mu_n(K^c) < e^{-\varepsilon a_n}$$

thus choosing ε such that

$$e^{-\varepsilon a_n} < \frac{e^{\frac{M_n \eta}{2}} - 1}{e^{M_n} - 1} = \frac{e^{Ma_n} - 1}{e^{\frac{2Ma_n}{\eta}} - 1}$$

($\exists n_0$ s.t. $\forall n > n_0$ it is possible) we will obtain

$$\frac{\log(1 + \mu(K^c)(e^{M_n} - 1))}{M_n} < \frac{\eta}{2}$$

and so the thesis for $n > n_0$, but the finite union with $n \leqslant n_0$ is also precompact, so the thesis. Of course we let $\tilde{K_\eta} = K_\varepsilon$.

(H)⇒(P) By the previous inequalities, for each $l > 0$ and integer $n_0 \geqslant 1$

$$\begin{aligned} G(n_0, l) := \bigcup_{n \geqslant n_0} \left\{ \mu_n^{K^c}, K \subseteq X \text{ is compact and } \mu_n(K^c) \geqslant e^{-la_n} \right\} \subseteq \\ \subseteq \bigcup_{n \geqslant n_0} \{ \nu \in \mathcal{P}(X) : H_n^a(\nu) \leqslant l \} \end{aligned}$$

thus $\forall l > 0 \quad \exists n_0(l)$ such that $G(n_0, l)$ is precompact in $\mathcal{P}(X)$, thus tight for Prokhorov's theorem. So $\forall l > 0$ we can find a compact $K_l \subseteq X$ such that $\mu_n^{K^c}(K_l^c) \leqslant \frac{1}{2} \quad \forall n \geqslant n_0(l)$ and for each K compact such that $\mu_n(K^c) \geqslant e^{-la_n}$, thanks to tightness.

Since $\mu_n(K^c | K^c) = 1 \forall K : \mu_n(K) > 0$ thus $K_l \neq K$ for each compact K (otherwise we would arrive to an absurd) with $\mu_n(K^c) \geqslant e^{-la_n}$ for some $n \geqslant n_0(l)$, so definitively in $n, \forall l > 0 \quad \mu_n(K_l^c) \leqslant e^{-la_n}$. So we have (P) for definition of exponentially tightness.

□

2.2 Γ− convergence and Large Deviations

Now we state and prove the equivalent versions in Large Deviations context of Proposition 9. Next theorem concerns about convergence satisfying a Large Deviations lower bound, related to an inequality for the $\Gamma - \overline{\lim}_n \frac{1}{a_n} H(\cdot | \mu_n)$.

Theorem 11. *The following are equivalent:*

(P1) $(\mu_n)_n$ *satisfies a Large Deviations lower bound with speed* $(a_n)_n$ *and rate* I.

(P2) For each sequence $(\varphi_n)_n$ of measurable functions $\varphi_n : X \to \mathbb{R} \cup \{+\infty, -\infty\}$ the inequality

$$\underline{\lim}_n \frac{1}{a_n} \log \mu_n(\exp(a_n \varphi_n)) \geqslant \sup_{x \in X} \{(\Gamma - \underline{\lim}_n \varphi_n)(x) - I(x)\} \tag{2.1}$$

holds, with $(\Gamma - \underline{\lim}_n \varphi_n)(x) - I(x) = -\infty$ if $I(x) = +\infty$.

(H1)

$$(\Gamma - \overline{\lim}_n H_n^a)(\delta_x) \leqslant I(x). \tag{2.2}$$

(H2) For each $\nu \in \mathcal{P}(X)$

$$(\Gamma - \overline{\lim}_n H_n^a)(\nu) \leqslant \nu(I) = \int_X I d\nu. \tag{2.3}$$

Proof. The implications will be proved as follows: (P1)⇒(H1)⇒(P2)⇒(P1) and (H1)⇔(H2).

(P1)⇒(H1) Fixing $x \in X$, inequality $(\Gamma - \overline{\lim}_n H_n^a)(\delta_x) \leqslant I(x)$ is proved by providing a sequence $\tilde{\nu_n} \to \delta_x$ such that $\overline{\lim}_n H_n^a(\tilde{\nu_n}) \leqslant I(x)$.

For $x \in X$ and $\delta > 0$ let $B_\delta(x)$ the open ball of radius δ centered in x. Fix n and define $\nu_{n,\delta} \in \mathcal{P}(X)$ as follows:

$$\nu_{n,\delta} := \begin{cases} \mu_n^{B_\delta(x)} & \text{if } \mu_n(B_\delta(x)) > 0 \\ \delta_x & \text{otherwise} \end{cases}$$

We observe that, in the case $\mu_n(B_\delta(x)) > 0$,

$$\frac{d\nu_{n,\delta}}{d\mu_n} = \frac{d\mu_n^{B_\delta(x)}}{d\mu_n} = \frac{\chi_{B_\delta(x)}}{\mu_n(B_\delta(x))}$$

So we can obtain a comfortable formula of $H(\nu_{n,\delta}|\mu_n)$:

$$\begin{aligned} H(\nu_{n,\delta}|\mu_n) &= \int_X \frac{d\nu_{n,\delta}}{d\mu_n} \log\left(\frac{d\nu_{n,\delta}}{d\mu_n}\right) d\mu_n \\ &= \int_{B_\delta(x)} \frac{d\nu_{n,\delta}}{d\mu_n} \log\left(\frac{d\nu_{n,\delta}}{d\mu_n}\right) d\mu_n + \int_{X \setminus B_\delta(x)} \frac{d\nu_{n,\delta}}{d\mu_n} \log\left(\frac{d\nu_{n,\delta}}{d\mu_n}\right) d\mu_n \\ &= \int_{B_\delta(x)} \frac{\chi_{B_\delta(x)}}{\mu_n(B_\delta(x))} \log\left(\frac{\chi_{B_\delta(x)}}{\mu_n(B_\delta(x))}\right) d\mu_n \\ &= -\log(\mu_n(B_\delta(x))) \end{aligned}$$

that we can extend also to case $\mu_n(B_\delta(x)) = 0$, where we understand $-\log(0) = +\infty$. Now, for Large Deviations lower bound definition $\forall \delta > 0$

$$\overline{\lim}_n H_n^a(\nu_{n,\delta}) = -\underline{\lim}_n \frac{1}{a_n} \log(\mu_n(B_\delta(x))) \leqslant \inf_{y \in B_\delta(x))} I(y) \leqslant I(x)$$

where we chose the open $O = B_\delta(x)$ At the same time

$$\lim_{\delta \to 0} \lim_n \nu_{n,\delta} = \delta_x \text{ in } \mathcal{P}(X)$$

thus by a diagonal argument there exists a sequence $(\delta_n)_n$ converging to 0 such that $\lim_n \nu_{n,\delta_n} = \delta_x$. For the initial observation we got the thesis.

(H1)⇒(H2) First of all recall the Jensen inequality for a convex function φ and X random variable.

$$\varphi(\mathbb{E}[X]) \leqslant \mathbb{E}[\varphi(X)]$$

Because of the convexity of H_n^a as a functional, also $\Gamma - \overline{\lim}_n H_n^a$ is convex and lower semicontinuous.

$$(\Gamma - \overline{\lim_n} H_n^a)(\nu) = \Gamma - \overline{\lim_n} H_n^a \left(\int_{\mathcal{P}(X)} \nu(dx)\delta_x \right)$$

and using Jensen inequality[1] and hypothesis (H1)

$$\leqslant \int_{\mathcal{P}(X)} \nu(dx)(\Gamma - \overline{\lim_n} H_n^a)(\delta_x) \leqslant \int_{\mathcal{P}(X)} \nu(dx) I(x) = \nu(I)$$

(H2)⇒(H1) We can trivially take $\nu = \delta_x$.

(H1)⇒(P2) Before starting with the real proof of the implication, let us do an observation. Let

$$Y = \left\{ x \in X : (\Gamma - \underline{\lim_n} \varphi_n)(x) > -\infty \right\}$$

so $\forall x \in Y \quad \exists \delta(x) > 0 \quad \exists n_0(x) = n_0 \in \mathbb{N}$ such that

$$\inf_{y \in B_\delta(x)} \left(\inf_{n \geqslant n_0} \varphi_n(y) \right) > -\infty$$

Indeed, we can show it with a proof by contradiction. Assume to the contrary that

$$\exists x \in Y : \forall \delta > 0 \quad \forall n_0 \in \mathbb{N} \qquad \inf_{y \in B_\delta(x)} \left(\inf_{n \geqslant n_0} \varphi_n(y) \right) = -\infty$$

So given x as in the contrary assumption we fix $\delta = \frac{1}{m}$ with $m \in \mathbb{N}^+$. By definition of infimum,

$$\forall M > 0 \quad \exists y'(M) = y' \in B_{\frac{1}{m}}(x) : \inf_{n \geqslant n_0} \varphi_n(y') < -M$$

Recall for clarity the definition of $\Gamma - \underline{\lim}_n$:

$$(\Gamma - \underline{\lim_n} \varphi_n)(x) = \inf \left\{ \underline{\lim_n} \varphi_n(x_n) : x_n \to x \right\}$$

Joining the last observations we can say:

$$\forall m \in \mathbb{N}^+ \exists y_m \in B_{\frac{1}{m}}(x) : \inf_{n \geqslant n_0} \varphi_n(y_m) < -m$$

thus $\forall M > 0$ we can extract a sequence $y_n \to x$ such that $\underline{\lim}_n \varphi_n(y_n) < -M$. This is absurd by definition of Y.

For $x \in Y$ by condition 2.2 there exists a sequence $(\nu_{n,x})_n$ converging to δ_x in $\mathcal{P}(X)$ and such that

[1]The previous and the following passages are quite technical and need a level of accuracy that would divert the reader's attention from the proof, so it will be better specified in the appendix. In particular Pettis integral and a generalized form of Jensen inequality are necessary to formalize these steps.

$$\overline{\lim_n} H_n^a(\nu_{n,x}) \leqslant I(x)$$

that can be assumed concentrated on $B_\delta(x)$. By definition or Relative Entropy 1.5 for each measurable $\varphi : X \to \mathbb{R} \cup \{+\infty, -\infty\}$ we obtain:

$$H(\nu_{n,x}|\mu_n) \geqslant \nu_{n,x}(\varphi) - \log \mu_n(e^\varphi) \tag{2.4}$$

provided we understand the right hand side as $-\infty$ whenever $H(\nu_{n,x}|\mu_n) = +\infty$ or $\nu_{n,x}(\varphi^-) = +\infty$. Choosing $\varphi = a_n\varphi_n$ we obtain

$$\begin{aligned}
&H(\nu_{n,x}|\mu_n) \geqslant \nu_{n,x}(a_n\varphi_n) - \log \mu_n(e^{a_n\varphi_n})\\
&\Rightarrow \log \mu_n(e^{a_n\varphi_n}) \geqslant (-H(\nu_{n,x}|\mu_n) + \nu_{n,x}(a_n\varphi_n))\\
&\Rightarrow \frac{1}{a_n} \log \mu_n(e^{a_n\varphi_n}) \geqslant (-H_n^a(\nu_{n,x}) + \nu_{n,x}(\varphi_n))\\
&\Rightarrow \underline{\lim_n} \frac{1}{a_n} \log \mu_n(e^{a_n\varphi_n}) \geqslant \underline{\lim_n} (-H_n^a(\nu_{n,x}) + \nu_{n,x}(\varphi_n))\\
&\Rightarrow \underline{\lim_n} \frac{1}{a_n} \log \mu_n(e^{a_n\varphi_n}) \geqslant \underline{\lim_n} (-H_n^a(\nu_{n,x})) + \underline{\lim_n} \nu_{n,x}(\varphi_n)\\
&\Rightarrow \underline{\lim_n} \frac{1}{a_n} \log \mu_n(e^{a_n\varphi_n}) \geqslant -\overline{\lim_n} H_n^a(\nu_{n,x}) + \underline{\lim_n} \nu_{n,x}(\varphi_n)\\
&\Rightarrow \underline{\lim_n} \frac{1}{a_n} \log \mu_n(e^{a_n\varphi_n}) \geqslant \sup_{x\in Y} \left\{ -\overline{\lim_n} H_n^a(\nu_{n,x}) + \underline{\lim_n} \nu_{n,x}(\varphi_n) \right\}
\end{aligned}$$

Because of the concentration of $\nu_{n,x}$ on $B_\delta(x)$ and $y \in Y$, we have that φ_n for $n \geqslant n_0(x)$ is bounded from below. Using (H1) hypothesis (2.2) and setting $\mu = \delta_x$ in (1.8) of point (P2) of Proposition 9, we can write the following inequality:

$$\underline{\lim_n} \nu_{n,x}(\varphi_n) \geqslant \delta_x(\Gamma - \underline{\lim_n} \varphi_n) = \Gamma - \underline{\lim_n} \varphi_n(x)$$

Moreover the definition of Γ-limsup and (H1) yield

$$I(x) \geqslant \overline{\lim_n} H_n^a(\nu_{n,x})$$

and joining the inequalities

$$\begin{aligned}
&-\overline{\lim_n} H_n^a(\nu_{n,x}) + \underline{\lim_n} \nu_{n,x}(\varphi_n) \geqslant -I(x) + (\Gamma - \underline{\lim_n} \varphi_n)(x)\\
&\Rightarrow \sup_{x\in Y} \left\{ -\overline{\lim_n} H_n^a(\nu_{n,x}) + \underline{\lim_n} \nu_{n,x}(\varphi_n) \right\} \geqslant \sup_{x\in Y} \left\{ -I(x) + (\Gamma - \underline{\lim_n} \varphi_n)(x) \right\}\\
&\Rightarrow \underline{\lim_n} \frac{1}{a_n} \log \mu_n(e^{a_n\varphi_n}) \geqslant \sup_{x\in Y} \left\{ -I(x) + (\Gamma - \underline{\lim_n} \varphi_n)(x) \right\}\\
&\Rightarrow \underline{\lim_n} \frac{1}{a_n} \log \mu_n(e^{a_n\varphi_n}) \geqslant \sup_{x\in X} \left\{ (\Gamma - \underline{\lim_n} \varphi_n)(x) - I(x) \right\}
\end{aligned}$$

that is the thesis.

(P2)⇒(P1) Given $O \subseteq X$ open and $M > 0$, we define $\varphi_n \equiv M\chi_O \quad \forall n \in \mathbb{N}$ that is lower semicontinuous, so it coincides with its Γ-limit. Let us do some reckoning.

First of all we will proceed with an estimate on $\mu_n(e^{a_n\varphi_n})$.

$$\begin{aligned}\mu_n(e^{a_n\varphi_n}) &= \int_X e^{a_n M\chi_O} d\mu_n = \int_O e^{a_n M\chi_O} d\mu_n + \int_{O^c} e^{a_n M\chi_O} d\mu_n = \\ &= \int_O e^{a_n M\chi_O} d\mu_n + \int_{O^c} 1 d\mu_n = 1 - \mu_n(O) + \mu_n(O)e^{a_n M} \leqslant \\ &\leqslant 1 + \mu_n(O)e^{a_n M}\end{aligned}$$

At this point we note some trivial inequalities: given $r \in \mathbb{R}^+$, if $r \geqslant 1 \Rightarrow 2r \geqslant r+1 \Rightarrow \log(r+1) \leqslant \log(2r) \leqslant \log 2 + \log r$. Otherwise $r+1 < 2 \Rightarrow \log(r+1) < \log(2)$. Choosing $r = \mu_n(O)e^{a_n M}$ follows that $\log(1+\mu_n(O)e^{a_n M}) \leqslant \log 2 + \max\{0, Ma_n + \log(\mu_n(O))\}$. Thus

$$\frac{1}{a_n}\log \mu_n(e^{a_n\varphi_n}) \leqslant \frac{\log 2}{a_n} + \max\left\{0, M + \frac{\log(\mu_n(O))}{a_n}\right\} \tag{2.5}$$

Now we will use the (P2) hypothesis (2.1). Because of lower semicontinuity of chosen φ_n, it follows

$$\Gamma - \varliminf_n \varphi_n(x) - I(x) = \varphi_n(x) - I(x) = M\chi_O(x) - I(x) \tag{2.6}$$

By (2.5) taking the liminf in n

$$\varliminf_n \frac{1}{a_n}\log \mu_n(e^{a_n\varphi_n}) - M \leqslant \max\left\{-M, \varliminf_n \frac{\log(\mu_n(O))}{a_n}\right\}$$

but for (2.1) in (P2) and (2.6) we got

$$\varliminf_n \frac{1}{a_n}\log \mu_n(e^{a_n\varphi_n}) - M \geqslant \sup_{x\in X}\{M\chi_O(x) - I(x)\} - M$$

Moreover

$$\begin{aligned}-\inf_{x\in O} I(x) &= \sup_{x\in O}\{-I(x)\} = \\ &= \sup_{x\in O}\{(M\chi_O(x) - I(x)) + (M - M\chi_O(x)) - M\} \leqslant \\ &\leqslant \sup_{x\in O}\{M\chi_O(x) - I(x)\} + \sup_{x\in O}\{M - M\chi_O(x)\} - M = \\ &= \sup_{x\in O}\{M\chi_O(x) - I(x)\} - M\end{aligned} \tag{2.7}$$

so

$$\begin{aligned}-\inf_{x\in O} I(x) &\leqslant \sup_{x\in O}(M\chi_O(x) - I(x)) - M \leqslant \\ &\leqslant \varliminf_n \frac{1}{a_n}\log \mu_n(e^{a_n\varphi_n}) - M \leqslant \\ &\leqslant \max\left\{-M, \varliminf_n \frac{\log(\mu_n(O))}{a_n}\right\}\end{aligned}$$

All these inequalities yield to

$$-\inf_{x\in O} I(x) \leqslant \max\left\{-M, \varliminf_n \frac{\log(\mu_n(O))}{a_n}\right\} \tag{2.8}$$

that implies the thesis when taking $M \to +\infty$.

□

Theorem 12. *The following are equivalent.*

(P1) $(\mu_n)_n$ *satisfies a Large Deviations weak upper bound with speed* $(a_n)_n$ *and rate* I.

(P2) For each sequence $(\varphi_n)_n$ *of measurable functions* $\varphi_n : X \to \mathbb{R} \cup \{+\infty, -\infty\}$ *bounded from below and such that*

$$\sup_{K \subseteq X \text{ compact}} \overline{\lim_n} \frac{\mu_n(1_{K^c} \exp(-a_n \varphi_n))}{\mu_n(\exp(-a_n \varphi_n))} = 0 \tag{2.9}$$

then

$$\overline{\lim_n} \frac{1}{a_n} \log \mu_n(\exp(-a_n \varphi_n)) \leqslant \sup_{x \in X} \{-(\Gamma - \underline{\lim_n} \varphi_n)(x) - I(x)\} \tag{2.10}$$

holds, where $-(\Gamma - \underline{\lim}_n \varphi_n)(x) - I(x) = -\infty$ *if* $I(x) = +\infty$.

(H1) $\forall x \in X$,

$$(\Gamma - \underline{\lim_n} H_n^a)(\delta_x) \geqslant I(x) \tag{2.11}$$

(H2) $\forall \nu \in \mathcal{P}(X)$,

$$(\Gamma - \underline{\lim_n} H_n^a)(\nu) \geqslant \nu(I) \tag{2.12}$$

Moreover assuming that $(\mu_n)_n$ *satisfies the conditions of Theorem 10, then the statements are also equivalent to*

(P3) $(\mu_n)_n$ *satisfies a Large Deviations upper bound with speed* (a_n) *and rate* I.

(P4) Given a sequence $(\varphi_n)_n$ *like in (P2) the following inequality*

$$\overline{\lim_n} \frac{1}{a_n} \log \mu_n(\exp(-a_n \varphi_n)) \leqslant \sup_{x \in X} \{-(\Gamma - \underline{\lim_n} \varphi_n)(x) - I(x)\} \tag{2.13}$$

holds, where $-(\Gamma - \underline{\lim}_n \varphi_n)(x) - I(x) = -\infty$ *if* $I(x) = +\infty$.

Proof. The implications will be proved as follows:

$$\text{(P1)} \Rightarrow \text{(H1)} \Leftrightarrow \text{(H2)} \Rightarrow \text{(P2)} \Rightarrow \text{(P1)}$$

for the first part. Assuming also Theorem 11 we will proceed with (Theorem 11-(P) & (P1))⇒(P3)⇒(P1) and (Theorem 11-(P) & (P2))⇒(P4)⇒(P2). The graph built up by putting conditions together comes out strongly connected.

(P1)⇒(H1) First of all we do some elementary reckoning:

$$\begin{aligned} \forall x \in \mathbb{R}^+ \quad 0 \leqslant \log(x+1) = \log\left(x\left(1+\frac{1}{x}\right)\right) = \\ = \log(x) + \log\left(1+\frac{1}{x}\right) \Rightarrow \\ \Rightarrow \log\left(1+\frac{1}{x}\right) \geqslant -\log(x) \end{aligned} \tag{2.14}$$

Moreover from inequality (1.7) $\forall A \in \mathcal{B}(X)$ we get

$$H(\nu|\mu) \geqslant \nu(A)\log\left(1+\frac{1}{\mu(A)}\right) - \log 2$$

and applying 2.14 with $x = \mu(A)$ we get

$$H(\nu|\mu) \geqslant -\nu(A)\log\mu_n(A) - \log 2 \tag{2.15}$$

We also recall that given two sequences $(a_n)_n \subseteq \mathbb{R}^+$ and $(b_n)_n \subseteq \mathbb{R}^+$, they hold:

$$\overline{\lim_n}(a_n b_n) \leqslant (\overline{\lim_n}\, a_n)(\overline{\lim_n}\, b_n) \text{ and } \underline{\lim_n}(a_n b_n) \geqslant (\underline{\lim_n}\, a_n)(\underline{\lim_n}\, b_n)$$

Let us start with the actual proof. Let $x \in X$ and $(\nu_n)_n \subseteq \mathcal{P}(X)$ a generic sequence be such that $\lim_n \nu_n = \delta_x$ in $\mathcal{P}(X)$. Showing

$$\underline{\lim_n}\, H_n^a(\nu_n) \geqslant I(x)$$

yields to (H1) by definition of Γ-liminf. Given $\varepsilon > 0$, the sequence ν_n is tight for Prokhorov's theorem so there exists K compact of X such that $\nu_n(K) \geqslant 1-\varepsilon \quad \forall n \in \mathbb{N} \Rightarrow \underline{\lim}_n\, \nu_n(K) \geqslant 1-\varepsilon$. Just take now $A = K \cap \overline{B_\varepsilon(x)}$, $\nu = \nu_n$ and $\mu = \mu_n$ in 2.15 and dividing by a_n

$$\begin{aligned}
&\frac{1}{a_n}H(\nu_n|\mu_n) \geqslant -\frac{1}{a_n}\nu_n(A)\log\mu_n(A) - \frac{\log 2}{a_n} \Rightarrow \\
&\Rightarrow \underline{\lim_n}\, H_n^a(\nu_n) \geqslant \underline{\lim_n}\left(-\frac{1}{a_n}\nu_n(A)\log\mu_n(A) - \frac{\log 2}{a_n}\right) \geqslant \\
&\qquad \geqslant \underline{\lim_n}\left(-\frac{1}{a_n}\nu_n(A)\log\mu_n(A)\right) \geqslant \\
&\qquad \geqslant -\underline{\lim_n}\,(\nu_n(A))\left(\overline{\lim_n}\,\frac{1}{a_n}\log\mu_n(A)\right)
\end{aligned} \tag{2.16}$$

Noting that $\underline{\lim}_n\, \nu_n(A) \geqslant 1-\varepsilon$, joining this with 2.16 and using (P1) we obtain

$$\begin{aligned}
\underline{\lim_n}\, H_n^a(\nu_n) &\geqslant -(1-\varepsilon)\left(\overline{\lim_n}\,\frac{1}{a_n}\log\mu_n(A)\right) \\
&\geqslant (1-\varepsilon)\inf_{y\in A} I(y) \geqslant (1-\varepsilon)\inf_{y\in\overline{B_\varepsilon(x)}} I(y)
\end{aligned}$$

Taking the limit $\varepsilon \to 0$, because of lower semicontinuity of I the right hand side converges to $I(x)$. For the initial observation we got the thesis.

(H1)$\Rightarrow$(H2) Before starting the proof, it's necessary to focus on some elementary calculations in order to use them below.

Let N be a positive integer number, let $p_1, \dots, p_N$ be positive real numbers such that $p_i \in (0,1) \quad \forall i \in \{1, \dots, N\}$ and $\sum_{i=1}^N p_i = 1$, thus

$$\sum_{i=1}^N p_i \log p_i \geqslant -\log N \tag{2.17}$$

Indeed this is equivalent to prove that

$$\sum_{i=1}^N p_i \log p_i + \log N \geqslant 0$$

as follows:

$$\begin{aligned}
\left(\sum_{i=1}^N p_i \log p_i\right) + \log N &= \left(\sum_{i=1}^N p_i \log p_i\right) - \log \frac{1}{N} = \\
&= \left(\sum_{i=1}^N p_i \log p_i\right) - \left(\sum_{i=1}^N p_i\right) \log \frac{1}{N} = \\
&= \sum_{i=1}^N p_i \left(\log p_i - \log \frac{1}{N}\right) = \\
&= \sum_{i=1}^N p_i \left(-\log \left(\frac{1}{p_i}\frac{1}{N}\right)\right) \geqslant \\
&\geqslant -\log \left(\sum_{i=1}^N p_i \frac{1}{p_i}\frac{1}{N}\right) = 0
\end{aligned}$$

where we used Jensen's inequality thanks to convexity of $-\log x$ and because of $\sum_{i=1}^N p_i = 1$. We proceed now with an observation about partitions.

Let $(E^i)_{i=1}^N$ a finite partition of the metric space X, let $\mu, \nu \in \mathcal{P}(X)$, we want to write the Relative Entropy functional in a more comfortable form. In the following sums the terms are understood to vanish for all i such that $\nu_n(E^i) = 0$. From 1.6

$$\begin{aligned}
H(\nu|\mu) &= \int_X \frac{d\nu}{d\mu} \log\left(\frac{d\nu}{d\mu}\right) d\mu = \sum_{i=1}^N \int_{E^i} \frac{d\nu}{d\mu} \log\left(\frac{d\nu}{d\mu}\right) d\mu = \\
&= \sum_{i=1}^N \int_{E^i} \frac{d\nu}{d\mu}\left(\log\left(\frac{d\nu}{d\mu}\right) - \log \nu(E^i) + \log \nu(E^i)\right) d\mu = \\
&= \sum_{i=1}^N \int_{E^i} \frac{d\nu}{d\mu}\left(\log\left(\frac{1}{\nu(E^i)}\frac{d\nu}{d\mu}\right) + \log \nu(E^i)\right) d\mu = \\
&= \sum_{i=1}^N \nu(E^i)\left(\int_{E^i} \frac{1}{\nu(E^i)}\frac{d\nu}{d\mu}\left(\log\left(\frac{1}{\nu(E^i)}\frac{d\nu}{d\mu}\right) + \log \nu(E^i)\right) d\mu\right) = \\
&= \sum_{i=1}^N \nu(E^i)\left(\int_{E^i} \frac{1}{\nu(E^i)}\frac{d\nu}{d\mu}\left(\log\left(\frac{1}{\nu(E^i)}\frac{d\nu}{d\mu}\right)\right) d\mu\right) + \\
&+ \sum_{i=1}^N \int_{E^i} \frac{d\nu}{d\mu} \log \nu(E^i) d\mu
\end{aligned}$$

Noting that

$$H(\nu^{E^i}|\mu) = \int_{E^i} \frac{1}{\nu(E^i)} \frac{d\nu}{d\mu} \left(\log \left(\frac{1}{\nu(E^i)} \frac{d\nu}{d\mu} \right) \right) d\mu$$

Relative Entropy becomes

$$H(\nu|\mu) = \sum_{i=1}^{N} \nu(E^i) H(\nu^{E^i}|\mu) + \sum_{i=1}^{N} \nu(E^i) \log \nu(E^i) \tag{2.18}$$

Applying 2.17 to the second term of 2.18 we get

$$H(\nu|\mu) \geqslant \sum_{i=1}^{N} \nu(E^i) H(\nu^{E^i}|\mu) - \log N$$

Now we can start with the proof of (H1)⇒(H2). To prove (H2) it suffices to show that, given $\nu \in \mathcal{P}(X)$ and $(\nu_n)_n$ a sequence converging to ν in $\mathcal{P}(X)$.

$$\Gamma - \varliminf_n H_n^a(\nu_n) \geqslant \nu(I)$$

For $\delta > 0$ and $l > 0$ let $(E_{\delta,l}^i)_{i=0}^{N_{\delta,l}}$ be a partition as in Lemma 8 with $\mu = \nu$. For $i \in \{0, \dots, N_{\delta,l}\}$ such that $\nu_n(E_{\delta,l}^i)$, we define $\nu_{n,\delta,l}^i := \nu_n^{E_{\delta,l}^i} \in \mathcal{P}(X)$. Considering the above observations we get:

$$H(\nu_n|\mu_n) \geqslant \sum_{i=1}^{N_{\delta,l}} \nu_n(E^i) H(\nu_{n,\delta,l}^i|\mu_n) - \log N_{\delta,l}$$

and dividing by a_n

$$H_n^a(\nu_n) \geqslant \sum_{i=1}^{N_{\delta,l}} \nu_n(E^i) H_n^a(\nu_{n,\delta,l}^i) - \frac{1}{a_n} \log N_{\delta,l}$$

We take the liminf to both sides

$$\begin{aligned} \varliminf_n H_n^a(\nu_n) &\geqslant \varliminf_n \sum_{i=1}^{N_{\delta,l}} \nu_n(E^i) H_n^a(\nu_{n,\delta,l}^i) + \varliminf_n \left(-\frac{1}{a_n} \log N_{\delta,l} \right) \geqslant \\ &\geqslant \sum_{i=1}^{N_{\delta,l}} \nu(E^i) \varliminf_n H_n^a(\nu_{n,\delta,l}^i) = \int_X J_{\delta,l}(x) d\nu(x) \end{aligned}$$

where

$$J_{\delta,l}(x) = \varliminf_n H_n^a(\nu_{n,\delta,l}^i) \text{ if } x \in E_{\delta,l}^i.$$

Moreover by Fatou's lemma on the right hand side of the last inequality

$$\varliminf_n H_n^a(\nu_n) \geqslant \int_X \lim_l \lim_\delta J_{\delta,l}(x) d\nu(x)$$

but since

$$\lim_l \lim_\delta \lim_n \nu_{n,\delta,l}^i = \delta_x$$

(H1) for definition of Γ-liminf implies

$$\lim_l \lim_\delta J_{\delta,l}(x) \geqslant I(x)$$

therefore the thesis for the initial observation.

(H2)⇒(H1) We can trivially take $\nu = \delta_x$.

(H2)⇒(P2) Let $(\varphi_n)_n$ be a sequence of measurable functions as in hypothesis of (P2) statements. First of all we consider the sequence of probability measures $(\nu_n)_n \subseteq \mathcal{P}(X)$ defined as follows:

$$\nu_n(dx) = \frac{e^{-a_n\varphi_n(x)}}{\mu_n(e^{-a_n\varphi_n})}\mu_n(dx) \tag{2.19}$$

that implies the following relative density:

$$\frac{d\nu_n}{d\mu_n} = \frac{e^{-a_n\varphi_n(x)}}{\mu_n(e^{-a_n\varphi_n})} \tag{2.20}$$

Let us note that in the statement of implication (H1)⇒(H2) the measures ν_n are not involved so we can choose them. Moreover $\mu_n(e^{-a_n\varphi_n})$ is a real *fixed* number, thus we can treat it as a constant inside an integral.

We start from the alternative definition of Relative Entropy (1.6) and (2.20) to do some reckoning. Choosing $\nu = \nu_n$ and $\mu = \mu_n$ in (1.6) we have:

$$\begin{aligned}
H(\nu_n|\mu_n) &= \int_X \frac{d\nu_n}{d\mu_n}\log\left(\frac{d\nu_n}{d\mu_n}\right)d\mu_n = \int_X \log\left(\frac{d\nu_n}{d\mu_n}\right)d\nu_n = \\
&= \int_X \log\left(\frac{e^{-a_n\varphi_n(x)}}{\mu_n(e^{-a_n\varphi_n})}\right)d\nu_n = \\
&= \int_X \log\left(e^{-a_n\varphi_n(x)}(\mu_n(e^{-a_n\varphi_n}))^{-1}\right)d\nu_n \\
&= \int_X \log\left(e^{-a_n\varphi_n(x)}\right)d\nu_n - \int_X \log\left(\mu_n(e^{-a_n\varphi_n})\right)d\nu_n = \\
&= \int_X -a_n\varphi_n(x)d\nu_n - \log\mu_n(e^{-a_n\varphi_n}) = \\
&= -a_n\nu_n(\varphi_n) - \log\mu_n(e^{-a_n\varphi_n})
\end{aligned} \tag{2.21}$$

At this point dividing by a_n and rearranging terms ensue that

$$\frac{1}{a_n}\log\mu_n(e^{-a_n\varphi_n}) = -\nu_n(\varphi_n) - H_n^a(\nu_n)$$

The last equality will be quite useful for our goal. In the following lines we show how the hypothesis of (P2) ensues a classical tightness of the sequence $(\nu_n)_n$. We observe indeed that given A Borel set of X, using (2.19)

$$\begin{aligned}
\nu_n(A) &= \int_X \chi_A d\nu_n = \int_X \chi_A \frac{e^{-a_n\varphi_n(x)}}{\mu_n(e^{-a_n\varphi_n})}d\mu_n = \\
&= \mu_n\left(\frac{\chi_A e^{-a_n\varphi_n(x)}}{\mu_n(e^{-a_n\varphi_n})}\right)
\end{aligned} \tag{2.22}$$

From (2.9) formula, we easily deduce:

$$\forall \varepsilon > 0 \quad \exists K_\varepsilon \text{ compact s.t. } \overline{\lim_n} \frac{\mu_n(1_{K_\varepsilon^c} e^{-a_n\varphi_n})}{\mu_n(e^{-a_n\varphi_n})} \leqslant \varepsilon$$

and as a consequence, setting $A = K_\varepsilon$ in (2.22) we get

$$\overline{\lim_n} \, \nu_n(K_\varepsilon) \leqslant \varepsilon$$

This proves tightness for $(\nu_n)_n \Rightarrow (\nu_n)_n$ is precompact for Prokhorov's theorem. Let $\nu \in \mathcal{P}(X)$ such that exists a subsequence $(\nu_{k_n})_n$ of $(\nu_n)_n$ and $\nu_{k_n} \to \nu$ in $\mathcal{P}(X)$, existing for precompactness.

$$\begin{aligned}
&\overline{\lim_n} \frac{1}{a_n} \log \mu_n(e^{-a_n\varphi_n}) \leqslant \overline{\lim_n} \, (-\nu_n(\varphi_n) - H_n^a(\nu_n)) \leqslant \\
&\leqslant \overline{\lim_n} \, (-\nu_n(\varphi_n)) + \overline{\lim_n} \, (-H_n^a(\nu_n)) \leqslant -\underline{\lim_n} \, \nu_n(\varphi_n) - \underline{\lim_n} \, H_n^a(\nu_n)
\end{aligned}$$

for (H2) hypothesis (2.12) of Theorem 12 and formula (1.8), so optimizing on x the right hand side, we arrive to

$$\begin{aligned}
\overline{\lim_n} \frac{1}{a_n} \log \mu_n(e^{-a_n\varphi_n}) &\leqslant -\nu(\Gamma - \underline{\lim_n} \, \varphi_n) - \nu(I) \leqslant \\
&\leqslant \int_X \sup_{x\in X} \{-\Gamma - \underline{\lim_n} \, \varphi_n(x) - I(x)\} d\nu \leqslant \\
&\leqslant \sup_{x\in X} \{-\Gamma - \underline{\lim_n} \, \varphi_n(x) - I(x)\} \int_X d\nu \leqslant \\
&\leqslant \sup_{x\in X} \{-\Gamma - \underline{\lim_n} \, \varphi_n(x) - I(x)\}.
\end{aligned}$$

We proved the thesis.

(P2)⇒(P1) Let K be any compact in X and M any positive real fixed. We fix the following sequence of measurable functions: $\varphi_n(x) \equiv 1_{K^c}(x) = \varphi(x) \quad \forall x \in X \quad \forall n \in \mathbb{N}$. We can start verifying that φ_n satisfies (2.9). Let us do some trivial reckoning.

$$\begin{aligned}
\mu_n(1_{K^c} e^{-a_n\varphi_n}) &= \int_X 1_{K^c}(x) e^{-Ma_n 1_{K^c}(x)} d\mu_n = \int_{K^c} e^{-Ma_n 1_{K^c}(x)} d\mu_n = \\
&= \mu_n(K^c) e^{-Ma_n} \\
\mu_n(e^{-a_n\varphi_n}) &= \int_X e^{-Ma_n 1_{K^c}(x)} d\mu_n = \\
&= \int_K 1 d\mu_n + \int_{K^c} e^{-Ma_n} d\mu_n = \mu_n(K) + \mu_n(K^c) e^{-Ma_n}
\end{aligned}$$

Joining the previous equalities we get:

$$\begin{aligned}
0 &\leqslant \frac{\mu_n(1_{K^c} e^{-a_n\varphi_n})}{\mu_n(e^{-a_n\varphi_n})} = \frac{\mu_n(K^c) e^{-Ma_n}}{\mu_n(K) + \mu_n(K^c) e^{-Ma_n}} = \\
&= \mu_n(K^c) e^{-Ma_n} \left(\frac{1}{1 + e^{Ma_n} \frac{\mu_n(K)}{\mu_n(K^c)}} \right) \leqslant \mu_n(K^c) e^{-Ma_n} \leqslant e^{-Ma_n}.
\end{aligned}$$

But $a_n \to +\infty$ thus

$$\overline{\lim_n} \frac{\mu_n(1_{K^c} e^{-a_n \varphi_n})}{\mu_n(e^{-a_n \varphi_n})} \leqslant \lim_n e^{-M a_n} = 0$$

that implies (2.9). So we can use (P2) with the chosen sequence $(\varphi_n)_n$. Furthermore we know that $\varphi = \Gamma - \lim_n \varphi_n$ because φ is lower semicontinuous. Moreover

$$1_K(x) \leqslant e^{-a_n M 1_{K^c}(x)} \Rightarrow \int_X 1_K(x) d\mu_n \leqslant \int_X e^{-a_n M 1_{K^c}(x)} d\mu_n$$

whereupon dividing by a_n and using (P2), with the result $\varphi = \Gamma - \lim_n \varphi_n$, yield to

$$\overline{\lim_n} \frac{1}{a_n} \log \mu_n(K) \leqslant \overline{\lim_n} \frac{1}{a_n} \log \mu_n(e^{-M a_n} 1_{K^c}) \leqslant \sup_{x \in X} \{-I(x) - M 1_{K^c}(x)\}$$

(P1) follows by the limit $M \to +\infty$.

(P3)⇒(P1) Trivial.

(P4)⇒(P2) Assuming (P4), (P2) is already proved.

From now on we will assume Theorem 11.

(P1)⇒(P3) Trivial as above.

(P2)⇒(P4) The idea to prove the implication is the following: remembering the proof of implication (H2) ⇒ (P2), we note that the hypothesis (2.9) is used *only* for existence of a subsequence of $(\nu_n)_n$ that converges to a limit point ν. Let us note that we already proved the equivalence (H2)⇔(P2), so assuming (P2) we also assume (H2). Moreover assuming Theorem 10 we also assume point (H) i.e. equicoercivity of H_n^a in $\mathcal{P}(X)$. We define as above some new probability measures ν_n:

$$\nu_n(dx) = \frac{e^{-a_n \varphi_n(x)}}{\mu_n(e^{-a_n \varphi_n})} \mu_n(dx) \tag{2.23}$$

Showing that exists, up to subsubsequences, a convergent subsequence of $(\nu_n)_n$, will allow to reproduce the same final steps of (H2) ⇒ (P2), that yield to thesis. For existence we can use equicoercivity of H_n^a, so we need to find constant $l > 0, n_0 \in \mathbb{N}$ such that $H_n^a(\nu_n) \geqslant l \quad \forall n \geqslant n_0$. We also note that if (2.9) holds, we already get the thesis because of (P2). By 2.21,

$$H_n^a(\nu_n) = -\nu_n(\varphi_n) - \frac{1}{a_n} \log \mu_n(e^{-a_n \varphi_n})$$

where, being $-\varphi_n(x) \leqslant B \in \mathbb{R}$ because bounded from below,

$$-\nu_n(\varphi_n) = \int_X \frac{-\varphi_n(x) e^{-a_n \varphi_n(x)}}{\mu_n(e^{-a_n \varphi_n})} d\mu_n \leqslant \frac{B}{\mu_n(e^{-a_n \varphi_n})} \int_X e^{-a_n \varphi_n} d\mu_n \leqslant B$$

whereupon of course there exists n_0 s.t. $a_n \geqslant 1 \quad \forall n \geqslant n_0 \Rightarrow \frac{1}{a_n} \leqslant 1 \quad \forall n \geqslant n_0$. We conclude that

$$H_n^a(\nu_n) \leqslant B - \log \mu_n(e^{-a_n \varphi_n})$$

The final step is therefore obtain a bound, up to subsequences, on the term $-\log \mu_n(e^{-a_n \varphi_n})$.

If (2.9) does not hold, it means that, given K compact of X,

$$\overline{\lim_n} \frac{\mu_n(1_{K^c} e^{-a_n \varphi_n})}{\mu_n(e^{-a_n \varphi_n})} = \lambda \in (0, 1]$$

and there exists a subsequence $(\nu_{k_n})_n$ that realizes limsup. For sake of simplicity of notation, we will use ν_n in place of ν_{k_n}. From

$$\lim_n \frac{\mu_n(1_{K^c} e^{-a_n \varphi_n})}{\mu_n(e^{-a_n \varphi_n})} = \lambda \in (0, 1]$$

we can find $n_0 \in \mathbb{N}$ and $\eta > 0$ such that $\lambda - \eta > 0$ and

$$\forall n \geqslant n_0 \quad \frac{\mu_n(1_{K^c} e^{-a_n \varphi_n})}{\mu_n(e^{-a_n \varphi_n})} \geqslant \lambda - \eta$$

At this point we estimate $\mu_n(1_{K^c} e^{-a_n \varphi_n})$ as follows:

$$\mu_n(1_{K^c} e^{-a_n \varphi_n}) \leqslant e^{Ba_n} \mu_n(K^c)$$

but we can also use the hypothesis of exponential tightness of $(\mu_n)_n$. Indeed

$$\forall \varepsilon > 0 \quad \exists K := K_\varepsilon : \mu_n(K^c) < e^{-\varepsilon a_n}$$

and setting $\varepsilon = B$

$$\mu_n(1_{K^c} e^{-a_n \varphi_n}) \leqslant e^{Ba_n} \mu_n(K^c) \leqslant 1$$

Putting the last inequality in one of the previous ones yields to

$$\forall n \geqslant n_0 \quad \frac{1}{\lambda - \eta} \geqslant \mu_n(e^{-a_n \varphi_n})$$

arriving to what we really need to use equicoercivity:

$$H_n^a(\nu_n) \leqslant B - \log \mu_n(e^{-a_n \varphi_n}) \leqslant B + \log(\lambda - \eta)$$

□

Chapter 3

Some results in Large Deviations theory

In this chapter we will prove some classical results of Large Deviations theory using the theorems of chapter 2. In particular we will prove an analogous of Schilder's theorem in a special case, using the probability laws given by Gaussian random variables. It is not of course the easiest way to catch on to the theorem, but applying the results to a *relatively straightforward* case could be clarifying. First of all we state Schilder's theorem with a quite general formulation, in order to understand the setting with a view from above.

Theorem 13 (Schilder's Theorem)**.** *Let B be a standard Brownian motion in d-dimensional Euclidean space $\mathbb{R}^d$ starting at the origin. Given $\varepsilon > 0$ let W_ε be the law of $\sqrt{\varepsilon}B$. Fixing $T \in \mathbb{R}^+$ let us denote*

$$C_0 = C_0([0,T]) := \{f : [0,T] \to \mathbb{R}^d | f \textit{ is continuous on } [0,T] \textit{ and } f(0) = 0\}$$

that is a Banach space equipped with the supremum norm $||\cdot||_\infty$. Then we define the rate function $J : C_0 \to \mathbb{R} \cup \{+\infty\}$ as

$$J(\omega) = \frac{1}{2}\int_0^T |\dot{\omega}(t)|^2 dt$$

if $\omega \in W^{1,2}([0,T]) \cap C_0$ and $+\infty$ otherwise.

Thus the probability measures W_ε satisfy the Large Deviations principle with good rate function J, i.e. for every open set $O \subseteq C_0$ and every closed set $C \subseteq C_0$ the following inequalities hold:

$$\overline{\lim_{\varepsilon \to 0}} \, \varepsilon \log W_\varepsilon(C) \leqslant -\inf_{\omega \in C} J(\omega)$$

and

$$\overline{\lim_{\varepsilon \to 0}} \, \varepsilon \log W_\varepsilon(O) \geqslant -\inf_{\omega \in G} J(\omega).$$

3.1 A classical result in finite dimension

Now we will apply the results we found to prove a finite dimensional version of Schilder's theorem in a special case. We are going to use the implications (H2)⇒(P1) of Theorem 11 and (H2)⇒(P3) of Theorem 12.

Proposition 14. *Let $V : \mathbb{R}^d \to \mathbb{R}_0^+$ be a continuous function such that*

$$\liminf_{x \to +\infty} \frac{V(x)}{||x||^2} > 0$$

Let $(\mu_n)_n$ be the sequence of probability measures in $\mathcal{P}(\mathbb{R}^d)$ such that

$$\mu_n(dx) = c_n e^{-nV(x)} dx,$$

where

$$c_n = \left(\int_X e^{-nV(x)} dx \right)^{-1}.$$

Moreover, denoting with λ the Lebesgue measure, we assume that

$$\lim_n \frac{1}{n} \log \lambda \left(\left\{ x \in \mathbb{R}^d | V(x) \leqslant \frac{1}{n} \right\} \right) = 0$$

Then μ_n satisfies a Large Deviations lower bound and a Large Deviations upper bound with speed $a_n = n$ and with rate $I : \mathbb{R}^d \to \mathbb{R}$ where $I(x) = V(x)$.

Proof. By theorems 11 and 12 we use the implications (H2)⇒(P1) of Theorem 11 and (H2)⇒(P3) of Theorem 12 to prove our claim. Thus we need to prove that, for any $\nu \in \mathcal{P}(\mathbb{R}^d)$ the following inequalities hold:

$$\inf \left\{ \quad \overline{\lim_n} \frac{1}{n} H(\nu_n|\mu_n) \quad \middle| \quad \nu_n \to \nu \right\} \leqslant \int_{\mathbb{R}^d} I d\nu$$

for Theorem 11 and

$$\inf \left\{ \quad \underline{\lim_n} \frac{1}{n} H(\nu_n|\mu_n) \quad \middle| \quad \nu_n \to \nu \right\} \geqslant \int_{\mathbb{R}^d} I d\nu$$

for Theorem 12. Let $(\nu_n)_n$ be a sequence such that $\nu_n \to \nu$ in $\mathcal{P}(\mathbb{R}^d)$. We denote $\rho_n := \frac{d\nu_n}{dx}$ if $\nu_n \ll \mu_n$ and $\rho := \frac{d\nu}{dx}$. First of all we do some reckoning useful both for *liminf inequality* and for *limsup inequality.* Noting that

$$\frac{d\nu_n}{d\mu_n} = \frac{\rho_n(x)}{e^{-nV(x)}}$$

it is possible write down the following:

$$\begin{aligned} H(\nu_n|\mu_n) &= H(\rho_n(x)dx|e^{-nV(x)+\log c_n}dx) = \int_{\mathbb{R}^d} \frac{d\nu_n}{d\mu_n} \log \left(\frac{d\nu_n}{d\mu_n} \right) d\mu_n = \\ &= \int_{\mathbb{R}^d} \frac{\rho_n(x)}{e^{-nV(x)} c_n} \log \left(\frac{\rho_n(x)}{e^{-nV(x)+\log c_n}} \right) e^{-nV(x)} c_n dx \\ &= \int_{\mathbb{R}^d} \rho_n(x) \log \left(\frac{\rho_n(x)}{e^{-nV(x)+\log c_n}} \right) dx = \\ &= \int_{\mathbb{R}^d} \rho_n(x) \log \left(\rho_n(x) e^{nV(x)} \right) dx - \log c_n = \\ &= \int_{\mathbb{R}^d} \rho_n(x) \log \left(\rho_n(x) \right) dx + \int_{\mathbb{R}^d} \rho_n(x) nV(x) dx + \log \frac{1}{c_n} \end{aligned}$$

and dividing by n we get

$$H_n(\nu_n) = \frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\rho_n(x)\right)dx + \int_{\mathbb{R}^d} V(x)d\nu_n + \frac{1}{n}\log\frac{1}{c_n} \tag{3.1}$$

Inter alia from 3.1 we get

$$\overline{\lim_n}\, H_n(\nu_n) \leqslant \overline{\lim_n}\,\frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\rho_n(x)\right)dx + \overline{\lim_n}\int_{\mathbb{R}^d} V(x)d\nu_n + \overline{\lim_n}\,\frac{1}{n}\log\frac{1}{c_n}$$

and

$$\underline{\lim_n}\, H_n(\nu_n) \geqslant \underline{\lim_n}\,\frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\rho_n(x)\right)dx + \underline{\lim_n}\int_{\mathbb{R}^d} V(x)d\nu_n + \underline{\lim_n}\,\frac{1}{n}\log\frac{1}{c_n}$$

From hypothesis we get

$$e^{-V(x)} \in L^1(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)).$$

Moreover

$$\frac{1}{c_n} = \int_X e^{-nV(x)}dx = \frac{1}{c_1}\int_X e^{-(n-1)V(x)}e^{-V(x)}c_1 dx \leqslant \frac{1}{c_1}\int_X e^{-V(x)}c_1 dx \leqslant \frac{1}{c_1},$$

therefore

$$\overline{\lim_n}\,\frac{1}{n}\log\frac{1}{c_n} \leqslant 0.$$

On the other hand

$$\frac{1}{c_n} = \int_X e^{-nV(x)}dx \geqslant \int_{\{x|V(x)\leqslant\frac{1}{n}\}} e^{-nV(x)}dx = \int_{\{x|V(x)\leqslant\frac{1}{n}\}} e^{-1}dx,$$

therefore

$$\underline{\lim_n}\,\frac{1}{n}\log\frac{1}{c_n} \geqslant 0.$$

Because of these comments we can go on without loss of generality starting from

$$H_n(\nu_n) = \frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\rho_n(x)\right)dx + \int_{\mathbb{R}^d} \rho_n(x)V(x)dx \tag{3.2}$$

instead of (3.1).

After this heavy but essential introduction we are going to prove the liminf inequality and the limsup inequality.

Liminf inequality In the following paragraph $(\nu_n)_n$ is a generic sequence of probability measures converging to ν in $\mathcal{P}(\mathbb{R}^d)$. Without loss of generality we are going to assume $H_n(\nu_n) < +\infty \quad \forall n$.

Starting from (3.2), for Fatou's lemma

$$\underline{\lim_n}\int_{\mathbb{R}^d} \rho_n(x)V(x)dx \geqslant \int_{\mathbb{R}^d} \rho(x)V(x)dx = \nu(I)$$

Showing

$$\underline{\lim_n}\,\frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\rho_n(x)\right)dx \geqslant 0$$

yields to our thesis because of

$$\underline{\lim}_n H_n(\nu_n) \geqslant \underline{\lim}_n \frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\rho_n(x)\right)dx + \nu(I)$$

Let us denote $d\gamma = \mu_1$ and $d\gamma(dx) = \gamma(x)dx = c_1 e^{-V(x)}dx$.

$$\begin{aligned}
&\frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\rho_n(x)\right)dx = \frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\rho_n(x)\frac{\gamma(x)}{\gamma(x)}\right)dx \\
&= \frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\frac{\rho_n(x)}{\gamma(x)}\right)dx + \frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\gamma(x)\right)dx \\
&= \frac{1}{n}\int_{\mathbb{R}^d} \frac{\rho_n(x)}{\gamma(x)}\log\left(\frac{\rho_n(x)}{\gamma(x)}\right)d\gamma + \frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\left(-V(x)+\log c_1\right)dx \\
&\Rightarrow \underline{\lim}_n \frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\rho_n(x)\right)dx \geqslant \underline{\lim}_n \frac{1}{n}\int_{\mathbb{R}^d} \frac{\rho_n(x)}{\gamma(x)}\log\left(\frac{\rho_n(x)}{\gamma(x)}\right)d\gamma = \\
&= \underline{\lim}_n \frac{1}{n}\int_{\mathbb{R}^d} \frac{\nu_n}{d\gamma}\log\left(\frac{\nu_n}{d\gamma}\right)d\gamma = \\
&= H(\nu_n|d\gamma) \geqslant 0
\end{aligned}$$

Limsup inequality In order to obtain the limsup inequality we are going to show the existence of a recovery sequence that satisfies it. We fix $\nu \in \mathcal{P}(\mathbb{R}^d)$. Let X be a ν-measurable random variable and N be a Gaussian random variable where $N \sim N(0,\mathbb{I})$. Let ε be a real positive number that will be specified later, and will be n-dependent, so $\varepsilon = \varepsilon(n)$, in particular $\varepsilon(n) \to 0$ when $n \to +\infty$. We assume X and N independent.

As first observation, we have

$$\mathbb{E}[|X+\varepsilon N|^2] = \mathbb{E}[X^2] + \varepsilon^2\mathbb{E}[N^2] \to \mathbb{E}[X^2] = \int_{\mathbb{R}^d} X^2 d\nu \text{ if } n \to +\infty$$

In particular we choose as recovery sequence, the sequence of laws given by the random variables $X+\varepsilon N$, that is the convolution $\rho_\varepsilon * \nu$. The sequence is indicized on n, because $\varepsilon = \varepsilon(n)$. We observe that

$$||\rho_\varepsilon * \nu||_\infty \leqslant \int_{\mathbb{R}^d} \frac{e^{-\frac{||x||^2}{2}}}{(2\pi\varepsilon)^{\frac{d}{2}}}\nu(x)dx \leqslant (2\pi\varepsilon)^{-\frac{d}{2}}$$

Moreover

$$\begin{aligned}
\int_{\mathbb{R}^d} (\rho_\varepsilon * \nu)(x)\log((\rho_\varepsilon * \nu)(x))dx &\leqslant \int_{\mathbb{R}^d} \log ||\rho_\varepsilon * \nu||_\infty dx \\
&\leqslant -\frac{d}{2}\log(2\pi\varepsilon)
\end{aligned}$$

Therefore using Fatou's lemma and (3.2) we get

$$\begin{aligned}
\overline{\lim_n} H_n(\nu_n) &\leqslant \overline{\lim_n} \frac{1}{n}\int_{\mathbb{R}^d} \rho_n(x)\log\left(\rho_n(x)\right)dx + \overline{\lim_n}\int_{\mathbb{R}^d} \rho_n(x)V(x)dx \leqslant \\
&\leqslant -\frac{d}{2}\log(2\pi\varepsilon)\frac{1}{n} + \nu(I)
\end{aligned}$$

and choosing

$$\varepsilon(n) = \frac{1}{n}$$

we get the thesis.

$\square$

Appendix A

Partitions of a metric space

In the following propositions to focus on some properties of the partitions of the space X proposed in Remark 3.9 of the article [6] by M.Mariani. The construction of the partition is a little bit technical but not so much sophisticated. Let X be a complete metric separable space and $\{C_k\}_{k\in\mathbb{N}^+}$ be a sequence of compacts of X such that exists $\mu \in \mathcal{P}(X)$ for which

$$\lim_{k\to+\infty} \mu(C_k) = 1$$

and let $\varphi, \quad \{\varphi_n\}_{n\in\mathbb{N}^+}$ be a sequence of real-valued Borel functions such that $\varphi_n \to \varphi$ pointwise μ-a.e. . Moreover we assume that

$$\forall k \geq 1 \quad \exists N_k \in \mathbb{N} \quad \exists\{E_l^k\}_{l=1}^{N_k} : C_k \subseteq \bigcup_{l=1}^{N_k} E_l^k$$

where E_k^l are subsets of X with the following properties[1] :

- $\mathcal{E}_k := \sigma(\{E_l^k\}_{l=1}^{N_k})$, and $\lim\limits_{k\to+\infty} \sup\limits_{1\leqslant l\leqslant k} diam(\mathcal{E}_l) \longrightarrow 0$
- $\forall k \in \mathbb{N}^+ \quad E_i^k \cap E_j^k = \emptyset \quad \forall i, j \in \{1, ..., N_k\}$ if $j \neq i$
- $\forall k \in \mathbb{N}^+ \quad \forall l \in \{1, ..., N_k\} \quad \mathring{E}_l^k \neq \emptyset$

Given

$$x \in X \backslash \left(\bigcup_{k=1}^{+\infty} \bigcup_{l=1}^{N_k} \partial E_l^k \right) \text{ s.t. } \mu(\{x\}) \neq 0$$

is trivial to say that

$$\exists k \in \mathbb{N}^+ \quad \exists! l_k(x) \in \{1, ..., N_k\} : x \in E_{l_k(x)}^k$$

For sake of simplicity when we fix k, we call $l(x) = l_k(x)$. Now we fix $k \in \mathbb{N}^+$ and define an equivalence relation. We say

$$\forall x, y \in \bigcup_{l=1}^{N_k} E_l^k, \quad x \sim_k y \Leftrightarrow l(x) = l(y)$$

To verify it is an equivalence relation is trivial.

[1] This is the setting of Lemma 8.

Proposition 15. *Given x as above we call*

$$\varphi_{n,k}(x) = \inf_{y \sim_k x} \varphi_n(y) \tag{A.1}$$

and

$$\psi_k(x) = \varliminf_n \varphi_{n,k}(x) \tag{A.2}$$

thus

$$\sup_k \psi_k(x) = \Gamma - \varliminf_n \varphi_n(x) \tag{A.3}$$

Proof. As a first step we want to prove

$$\sup_k \psi_k(x) \leq \inf \left\{ \varliminf_n \varphi_n(x_n) \mid x_n \to x \right\} = \Gamma - \varliminf_n \varphi_n(x) \tag{A.4}$$

Fixed k, given $x_n \to x$ we have $x_n \sim_k x$ definitively in n. So

$$\varphi_{n,k}(x) = \varphi_{n,k}(x_n) \leq \varphi_n(x_n)$$

passing now to liminfs

$$\varliminf_n \varphi_{n,k}(x) \leq \varliminf_n \varphi_n(x_n)$$

and considering the supremum in k at the left hand side and the infimum varying $x_n \longrightarrow x$ at the right hand side we obtain the initial inequality.

Next, we see that

$$\sup_k \psi_k(x) \geq \inf \left\{ \varliminf_n \varphi_n(x_n) \mid x_n \to x \right\} \tag{A.5}$$

To do this, we will choose a particular sequence $x_n \to x$ whenever $n \to +\infty$.

Fixing x we have that $\psi_k(x)$ are increasing in k, so calling $s(x) = \sup_k \psi_k(x) = \lim_k \psi_k(x)$. By definition of sup.

$$\forall \varepsilon > 0 \quad \exists \quad k(\varepsilon) = k_\varepsilon \in \mathbb{N} | \forall k \geqslant k_\varepsilon$$

we can write

$$\psi_k(x) \leqslant s(x) + \varepsilon$$

and so

$$\varliminf_n \varphi_{n,k}(x) \leqslant s(x) + \varepsilon$$

By definition of liminf $\forall \delta > 0$

$$\varliminf_n \varphi_{n,k}(x) \leqslant s(x) + \varepsilon + \delta$$

frequently in n thus exists a subsequence $h_n \to +\infty$ for $n \to +\infty$ such that

$$\varphi_{h_n,k}(x) \leqslant s(x) + \varepsilon + \delta$$

for any n. Recalling the initial definitions

$$\inf_{y \sim_k x} \varphi_{h_n,k}(y) \leqslant s(x) + \varepsilon + \delta$$

therefore $\forall \eta > 0$ exists $y = y(k, h_n, \eta) \in E_k^i$ such that

$$\varphi_{h_n}(y(k, h_n, \eta)) \leqslant s(x) + \varepsilon + \delta + \eta$$

At this point we choose $(x_n)_n$ as follows:

$$x_i = \begin{cases} y(k_\varepsilon + 1, h_n, \frac{1}{n}) & \text{if } i = h_n \\ x & \text{otherwise} \end{cases}$$

with $\varepsilon = \frac{1}{k}$ we get the thesis taking $\underline{\lim}_n$ and $\sup_k$.

□

The following proposition is necessary to prove point (vi) of Lemma 8. For sake of simplicity and understanding of the proof it is appropriate change notations and work in a little bit different environment from that presented in the lemma.

Proposition 16. *Let $(\mathcal{A}_n)_n$ be a sequence of σ-algebras and $\mathcal{A}$ a σ-algebra such that:*

- $\mathcal{A}_n \subset \mathcal{A}_{n+1} \quad \forall n \in \mathbb{N}$
- $\sigma\left(\bigcup_{n\in\mathbb{N}} \mathcal{A}_n\right) = \mathcal{A}$

Then, given $\nu, \mu \in \mathcal{P}(X)$ s.t. $\nu << \mu$

$$H_{\mathcal{A}}(\nu|\mu) = \lim_{n\to+\infty} H_{\mathcal{A}_n}(\nu|\mu)$$

Proof. As a first step, we say that $H_{\mathcal{A}_n}(\nu|\mu)$ is increasing in n. In fact

$$\begin{aligned} H_{\mathcal{A}_n}(\nu|\mu) &= \sup_{\varphi\in L^\infty(X,\mathcal{A}_n,\mu)} \{\nu(\varphi) - \log\mu(e^\varphi)\} \leqslant \\ &\leqslant \sup_{\varphi\in L^\infty(X,\mathcal{A}_{n+1},\mu)} \{\nu(\varphi) - \log\mu(e^\varphi)\} = H_{\mathcal{A}_{n+1}}(\nu|\mu) \end{aligned}$$

Now, given $\varphi \in L^\infty(X, \mathcal{A}, \mu)$, defined

$$\varphi_n = \mathbb{E}_\mu[\varphi|\mathcal{A}_n] \text{ and } \varphi = \mathbb{E}_\mu[\varphi|\mathcal{A}]$$

that is $\mathcal{A}_n$-measurable for definition. At this point,

$$\forall A \in \mathcal{A} \quad \mathbb{E}_\mu(A|\mathcal{A}_n) \to I_A \ \mu\text{-almost surley} \Rightarrow \varphi_n \to \varphi \ \mu\text{-a.s.}$$

by martingale convergence theorem (see [2]). Moreover $\exists M \in \mathbb{R}^+$ such that $|\varphi_n| < M \quad \forall n \in \mathbb{N}$. So by dominated convergence we obtain that

$$\int_X \varphi_n d\nu = \int_X \varphi_n \frac{d\nu}{d\mu} d\mu \longrightarrow \int_X \varphi \frac{d\nu}{d\mu} d\mu = \int_X \varphi d\nu$$

In the same way

$$\int_X e^{\varphi_n} d\mu \longrightarrow \int_X e^\varphi d\mu$$

and so the thesis with sup and entropy definition. □

Appendix B

Jensen's inequality in locally convex spaces

In this appendix we are going to justify some steps of implication (H1)⇒(H2) of Theorem 11. In particular we will refer to the paper [8], where the author proves Jensen's integral inequality for proper convex functions, defined on a convex set in a topological vector space. The rigorous formalization of how to use the inequality is quite advanced and requires more theory of that proposed in this thesis, thus we are going only to outline a sketch of the former, without focusing on details.

The steps that we would like to justify are the following:

- first step:

$$(\Gamma - \overline{\lim_n} H_n^a)(\nu) = \Gamma - \overline{\lim_n} H_n^a \left(\int_{\mathcal{P}(X)} \nu(dx)\delta_x \right)$$

- second step:

$$\Gamma - \overline{\lim_n} H_n^a \left(\int_{\mathcal{P}(X)} \nu(dx)\delta_x \right) \leqslant \int_{\mathcal{P}(E)} \nu(dx)(\Gamma - \overline{\lim_n} H_n^a)(\delta_x)$$

The problem is to set up hypothesis and setting of theorems proposed in [8] .

General setting Let E be a Hausdorff locally convex topological vector space, and (C, Σ, μ) a convex probability space (i.e. a positive measure space such that $\mu(C) = 1$ and C is a nonempty convex subset of E) such that $\mathcal{B}(C) \subseteq \Sigma$ and the barycenter x_μ of μ exists in E. Let $f : C \to (-\infty, +\infty]$ be a proper convex function. We assume that C is evenly convex and f is lower semicontinuous. The topological dual of E is denoted by E^*.

By $\mathrm{eco}(C)$ we denote the *evenly convex hull* of C, that is the intersection od all open halfspaces containing C, saying that C is *evenly convex* if $\mathrm{eco}(C) = C$. The Hahn-Banach separation theorem [1] implies that $\mathrm{eco}(C) \subseteq \overline{C}$ and if C is either open or closed, thus C is evenly convex.

[1]See [1].

Barycenter of a probability measure Let (C, Ω, μ) be a convex space in E such that $\mathcal{B}(C) \subseteq \Sigma$. The barycenter x_μ is defined as the Pettis integral[2] of the identity mapping $I : C \to E, I(y) = y$ or equivalently

$$x^*(x_\mu) = \int_C x^* d\mu, \quad \forall x^* \in E^*$$

In [8] we find the following theorem:

Theorem 17 (Jensen's integral inequality)**.** *Let E be a Hausdorff locally convex topological vector space, and $(C, \Sigma), \mu)$ a convex probability space such that $\mathcal{B}(C) \subseteq \Sigma$ and the barycenter x_μ of μ exists in E. Let $f : C \to (-\infty, +\infty]$ be a proper convex function. We assume that C is evenly convex and f is lower semicontinous. Then:*

(a) $x_\mu \in C$

(b) the Lebesgue integral $\int_C f d\mu$ exists (finite or infinite);

(c) the Jensen's inequality holds:

$$f(x_\mu) \leqslant \int_C f d\mu \tag{B.1}$$

We would like to apply the previous theorem with $C = \mathcal{P}(X)$, that is not a topological vector space, but it is a convex set. The idea is to immerge $\mathcal{P}(X)$ in a bigger space that is a topological vector space, equipped with an appropriate structure of probability space (and so determine an appropriate μ). Moreover $f = (\Gamma - \overline{\lim}_n H_n^a)$, defined of $\mathcal{P}(X)$.

The topological vector space we could choose is the space of finite signed measures over X. The sum of two finite signed measures is a finite signed measure, as is the product of a finite signed measure by a real number: they are closed under linear combination. Thus the set of finite signed measures on a measurable space $(X, \mathcal{B}(X))$ is a real vector space; this is in contrast to positive measures, which are only closed under conical combination, and therefore form a convex cone but not a vector space. Moreover the total variation defines a norm in respect to which the space of finite signed measures becomes a Banach space.

[2]See Pettis, B. J. On integration in vector spaces. Trans. Amer. Math. Soc. 44 (1938), no. 2, 277–304.

Bibliography

[1] Haim Brezis (auth.). *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Universitext. Springer-Verlag New York, 1 edition, 2010.

[2] V. I. Bogachev. *Measure theory. Vol. I, II.* Springer-Verlag, Berlin, 2007.

[3] Andrea Braides. Chapter 2 a handbook of γ-convergence. volume 3 of *Handbook of Differential Equations: Stationary Partial Differential Equations*, pages 101 – 213. North-Holland, 2006.

[4] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[5] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22(1):79–86, 03 1951.

[6] Mauro Mariani. A γ-convergence approach to large deviations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2018.

[7] Gianni Dal Maso. *An introduction to G-convergence.* Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser Boston, 1 edition, 1993.

[8] Libor Vesely. Jensen's integral inequality in locally convex spaces. *Annals of the Academy of Romanian Scientists: Series on Mathematics and its Applications*, 9:136–144, 12 2017.