Large Deviations For Stochastic (Partial) Differential Equations

Annie Millet

1. The Cramer, Ellis-Gärtner and Schilder theorems

1.1 Introduction

Large deviations study "rare events", that is sets with exponentially small probability. These sets appear naturally in various applied settings. The aim of these lectures is to present the mathematical framework (due to Varadhan) and some methods that can be used in many settings.

A simple example: Let (X_n) be iid standard Gaussian (denoted $N(0,1)$). Then $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{a.s.} E(X_1) = 0$. Hence if A is a Borel set such that $0 \notin A$, $P(\overline{X}_n \in A) \to 0$. The question is "how fast" and how is the speed related to A?

Let $A = [t, +\infty)$ for $t > 0$. Since $\overline{X}_n \sim N(0, \frac{1}{n})$,

$$P(\overline{X}_n \geq t) = P\left(\frac{1}{\sqrt{n}} N(0,1) \geq t\right) = P(N(0,1) \geq t\sqrt{n})$$

These are usual "tail estimates" of the Gaussian: for $a > 0$,

$$P(N(0,1) \geq a) = \frac{1}{\sqrt{2\pi}} \int_a^{+\infty} e^{-\frac{x^2}{2}} dx \leq \frac{1}{\sqrt{2\pi}\, a} \int_a^{+\infty} x\, e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \frac{1}{a} e^{-\frac{a^2}{2}}$$

Hence $P(\overline{X}_n \geq t) \leq \frac{1}{\sqrt{2\pi}\, t\sqrt{n}} \exp\left(-n \frac{t^2}{2}\right)$

and $\overline{\lim_n} \frac{1}{n} \ln P(\overline{X}_n \geq t) \leq -\frac{t^2}{2}$

For lower estimates, we have

$$P(N(0,1) \geq a) \geq \frac{1}{\sqrt{2\pi}} \int_a^{2a} e^{-\frac{x^2}{2}} dx \geq \frac{1}{\sqrt{2\pi}} \frac{1}{2a} \int_a^{2a} x\, e^{-\frac{x^2}{2}} dx$$

$$\geq \frac{1}{\sqrt{2\pi}} \frac{1}{2a} \left(e^{-\frac{a^2}{2}} - e^{-2a^2}\right)$$

Hence $P(\overline{X}_n \geq t) \geq \frac{1}{\sqrt{2\pi}} \frac{1}{2t\sqrt{n}} \left(e^{-\frac{t^2 n}{2}} - e^{-2nt^2} \right)$ and

$$\underline{\lim} \frac{1}{n} \ln P(\overline{X}_n \geq t) \geq -\frac{t^2}{2}$$

This implies $\frac{1}{n} \ln P(\overline{X}_n \geq t) \to -\frac{t^2}{2}$, that is $P(\overline{X}_n \geq t) \asymp e^{-\frac{nt^2}{2}}$

This probability converges to 0 exponentially fast as $n \to \infty$

We want to extend this to

* more general set A and more general iid sequence (X_n) [Cramer]
* non iid sequences (such as Markov chains) [Ellis-Gärtner]
* infinite dimensional Gaussian random variables [Schilder]

Another (bad) example

Let (X_n) be iid symmetric, α-stable for $\alpha \in (0,1)$ [$\alpha = 0$ is a standard Gaussian r.v and $\alpha = 1$ is a Cauchy - non integrable - random variable]. Then the characteristic function is

$E[e^{itX_1}] = e^{-|t|^\alpha}$, while $P(|X| \geq t) \sim C t^{-\alpha}$ for "large" t.

For $\alpha \in (1,2)$, $X \in L^1$ and $X \notin L^2$ since

$$E(|X|) = \int_0^\infty P(|X| \geq t)\, dt \leq A + \int_A^{+\infty} C t^{-\alpha}\, dt < \infty$$

and $E(X^2) \geq \int_A^{+\infty} C\, t\, t^{-\alpha}\, dt = +\infty$

Furthermore

$$E[e^{it\overline{X}_n}] = \left(E(e^{i\frac{t}{n}X_1}) \right)^n = e^{-\left(\frac{|t|}{n}\right)^\alpha n}$$

$$= e^{-|t|^\alpha n^{1-\alpha}} = E[e^{itn^{\frac{1-\alpha}{\alpha}}X_1}]$$

Hence $\overline{X}_n \overset{\mathcal{L}}{=} n^{\frac{1-\alpha}{\alpha}} X_1$ and $P(|\overline{X}_n| \geq t) \sim C t^{-\alpha} n^{1-\alpha}$; $\frac{1}{n} \ln P(|\overline{X}_n| \geq t) \to 0$ (not related to t)

1.2 First definitions

Let X be a topological space, $(\mu_\varepsilon, \varepsilon>0)$ be a family of probabilities on a σ-field $\mathcal{F}$ on X.

Definition 1.1 (μ_ε) satisfies a Large Deviation Principle (LDP) with rate function I if

(1) $I : X \longrightarrow [0,+\infty]$ is l.s.c. (that is $\{I \leq a\}$ is a closed subset of X for any $a \geq 0$, or if X is metric, $I(x) \leq \underline{\lim}\, I(x_n)$ when $x_n \to x$.)

(2) For every $\Gamma \in \mathcal{F}$,

$$-\inf\{I(x) : x \in \mathring{\Gamma}\} \leq \underline{\lim}_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(\Gamma) \leq \overline{\lim}_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(\Gamma) \leq -\inf\{I(x) : x \in \overline{\Gamma}\}$$

Remarks (1) If one has a sequence (μ_n), one says that it satisfies a LDP for $\varepsilon_n = \frac{1}{n}$

(2) If $\mathcal{B}(X) \subset \mathcal{F}$, one checks

* lower bounds for open sets $\underline{\lim}_{\varepsilon} \varepsilon \ln \mu_\varepsilon(O) \geq -\inf\{I(x) : x \in O\}$
* upper bounds for closed sets $\overline{\lim}_{\varepsilon} \varepsilon \ln \mu_\varepsilon(F) \leq -\inf\{I(x) : x \in F\}$

(3) For $X_1 \sim N(0,1)$ and $F = [t,+\infty)$, $t>0$, we have shown upper bounds for the distribution of $\overline{X}_n$ with $I(x) = \frac{x^2}{2}$ and for $O = (t,+\infty)$, lower bounds with $I(x) = \frac{x^2}{2}$; $I : \mathbb{R} \to [0,+\infty]$ is continuous, hence l.s.c

(4) For X_1 α-stable, $1<\alpha<2$, we have shown upper bounds for $F = [t,+\infty)$ with $I(x) \equiv 0$.

(5) The above definition does NOT imply that the rate function is unique

However, for "good" topological spaces, the rate function is unique

Definition The topological space X is a regular Hausdorff space if for every closed set F and $x \notin F$, there exist disjoint open sets O_1 and O_2 of X such that $x \in O_1$ and $F \subset O_2$.

Examples metric spaces, Hausdorff topological vector spaces are regular

Proposition 1.2 La family (μ_ε) of probabilities defined on a σ-field $\mathcal{F}$ of a regular Hausdorff space has at most one rate function defining its LDP

1.3 Cramer's theorem in $\mathbb{R}$

Let μ be a probability on $\mathcal{B}(\mathbb{R})$, X a random variable with distribution μ

Definition 1.3 (i) The Laplace transform of μ (or of X) is defined by

$$\lambda \mapsto \int e^{\lambda x}\mu(dx) = E[e^{\lambda X}] \in [0,+\infty]$$

(2) The logarithmic moment generating function (Log Laplace) of μ, or of X, is

$$\Lambda : \mathbb{R} \longrightarrow (-\infty,+\infty]$$
$$\lambda \longmapsto \Lambda(\lambda) = \ln\left(\int e^{\lambda x}\mu(dx)\right)$$

The domain of Λ, $\mathcal{D}(\Lambda) = \{\lambda \in \mathbb{R} : \Lambda(\lambda) < \infty\}$

We at first prove the following results on Λ

Proposition 1.4 i) Λ is a convex, l.s.c. function, $\Lambda(0)=0$

(ii) For $\lambda \in \mathring{\mathcal{D}}(\Lambda)$, $\Lambda'(\lambda) = \dfrac{E(Xe^{\lambda X})}{E[e^{\lambda X}]}$

Hence if $0 \in \mathring{\mathcal{D}}(\Lambda)$, $\Lambda'(0) = E(X)$

(iii) If $P(X>0)>0$, $\Lambda(\lambda) \to +\infty$ as $\lambda \to +\infty$ (symmetric result on $(-\infty,0)$)

(iv) If $P(X<0)=0$, Λ is non decreasing and $\Lambda(\lambda) \to \ln P(X=0)$ as $\lambda \to -\infty$

Proof (i) Let $\alpha \in (0,1)$, λ_1 and λ_2 belong to $\mathcal{D}(\Lambda)$. Hölder's inequality with conjugate exponents $\frac{1}{\alpha}$ and $\frac{1}{1-\alpha}$ implies

$$E[e^{(\alpha\lambda_1 + (1-\alpha)\lambda_2)X}] \le \left(E(e^{\lambda_1 X})\right)^{\alpha}\left(E(e^{\lambda_2 X})\right)^{1-\alpha}.$$

Taking $\ln$, we deduce $\Lambda(\alpha\lambda_1 + (1-\alpha)\lambda_2) \le \alpha\Lambda(\lambda_1) + (1-\alpha)\Lambda(\lambda_2)$, and Λ is convex.

If $\lambda_n \to \lambda$, Fatou's lemma implies

$E[e^{\lambda X}] \le \varliminf_n E(e^{\lambda_n X})$, and hence $\Lambda(\lambda) \le \varliminf \Lambda(\lambda_n)$.

Furthermore $\Lambda(0) = \ln E(e^0) = 0$.

(ii) Let $\lambda_0 \in \overset{\circ}{\mathcal{D}(\Lambda)}$ and $\alpha > 0$ be such that $(\lambda_0 - 3\alpha, \lambda_0 + 3\alpha) \subset \mathcal{D}(\Lambda)$.

For $\lambda \in (\lambda_0 - \alpha, \lambda_0 + \alpha)$, we have

$$\left|\frac{\partial}{\partial \lambda} e^{\lambda X}\right| = |X e^{\lambda X}| \leq c\, [e^{(\lambda_0 + 2\alpha)X} + e^{(\lambda_0 - 2\alpha)X}].$$

Hence $\frac{\partial}{\partial \lambda}\left(E(e^{\lambda X})\right) = E\left[\frac{\partial}{\partial \lambda}(e^{\lambda X})\right] = E\,(X e^{\lambda X})$; the chain rule yields

$$\Lambda'(\lambda_0) = \frac{E(X e^{\lambda_0 X})}{E(e^{\lambda_0 X})}$$

The proofs of (iii) and (iv) are left as an exercise. $\square$

Definition 1.5 Let $\mathcal{J}$ be a probability on $\mathcal{B}(\mathbb{R})$. Its Cramer transform is the Fenchel Legendre transform of Λ, that is

$$\Lambda^*(x) = \sup\{\lambda x - \Lambda(\lambda) : \lambda \in \mathbb{R}\} \in [0, +\infty]$$

<u>Remark</u> $\Lambda(0) = 0$ implies that $\Lambda^*(x) \geq 0$ for every $x \in \mathbb{R}$.

Examples

(i) $X \sim N(0,1)$ $\quad E[e^{\lambda X}] = e^{\frac{\lambda^2}{2}}$ and $\Lambda(\lambda) = \frac{\lambda^2}{2}$

$$\Lambda^*(x) = \sup\{\lambda x - \tfrac{\lambda^2}{2};\ \lambda \in \mathbb{R}\} = \sup\{-\tfrac{1}{2}(\lambda - x)^2 + \tfrac{x^2}{2};\ \lambda \in \mathbb{R}\} = \frac{x^2}{2}.$$

(ii) $X \sim N(m, \sigma^2)$ $\quad X = m + \sigma Y$, where $Y \sim N(0,1)$

$$E[e^{\lambda X}] = e^{\lambda m + \frac{\lambda^2 \sigma^2}{2}} \quad \text{and} \quad \Lambda(\lambda) = \lambda m + \frac{\lambda^2 \sigma^2}{2} = \lambda m + \Lambda_0(\sigma \lambda)$$

where Λ_0 is the Log Laplace of Y

$$\Lambda^*(x) = \sup\{\lambda x - (\lambda m + \Lambda_0(\sigma\lambda));\ \lambda \in \mathbb{R}\} = \sup\{\lambda(x - m) - \Lambda_0(\sigma\lambda);\ \lambda \in \mathbb{R}\}$$

$$= \sup\{\lambda \tfrac{x-m}{\sigma} - \Lambda_0(\lambda);\ \lambda \in \mathbb{R}\} = \Lambda_0^*\left(\tfrac{x-m}{\sigma}\right) = \tfrac{1}{2}\left(\tfrac{x-m}{\sigma}\right)^2$$

Proposition 1.6 (i) The function Λ^* is convex, l.s.c

(ii) If $y = \Lambda'(\eta)$ for $\eta \in \overset{\circ}{\mathcal{D}(\Lambda)}$, then $\Lambda^*(y) = \eta y - \Lambda(\eta)$

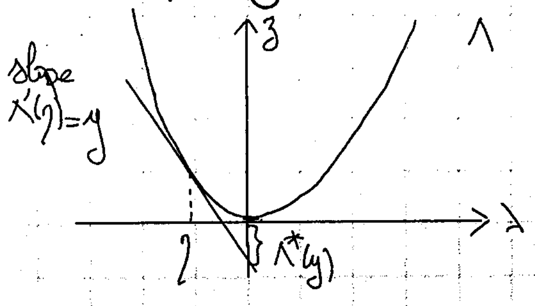
(iii) If $m = E(X) \in \mathbb{R}$, $\Lambda^*(m) = 0$

(iv) For $x \geq m$, $\Lambda^*(x) = \sup\{\lambda x - \Lambda(\lambda) \,;\, \lambda \geq 0\}$ and Λ^* is increasing on $(m, +\infty)$ (and symmetric results for $x \leq m$)

Proof (i) Λ^* is the sup of affine function, hence Λ^* is convex.

Λ^* is the sup of continuous functions hence Λ^* is l.s.c.

(ii) Let $y = \Lambda'(\eta)$ for $\eta \in \overset{\circ}{\widehat{\mathcal{D}(\Lambda)}}$. One has to check that the function g defined by $g(\lambda) = \lambda y - \Lambda(\lambda)$ achieves its maximum at η. The function g is concave and $g'(\eta) = y - \Lambda'(\eta) = 0$



tangent at point η $\quad z - \Lambda(\eta) = \Lambda'(\eta)(\lambda - \eta)$
$= y(\lambda - \eta)$

intersection with $Oz \to \lambda = 0$

$z = \Lambda(\eta) - \eta y = -\Lambda^*(y)$

(iii) Since ln is concave Jensen's inequality implies

$$\Lambda(\lambda) = \ln E(e^{\lambda X}) \geq E[\ln(e^{\lambda X})] = \lambda m$$

Hence $\Lambda^*(m) = \sup\{\lambda m - \Lambda(\lambda) : \lambda \in \mathbb{R}\} \leq 0$; since $\Lambda^* \geq 0$, we deduce $\Lambda^*(m) = 0$.

(iv) Let $x \geq m$.

Case 1 $m \in \mathbb{R}$, which implies $\Lambda^*(m) = 0$

We check that for $\lambda \leq 0$, $\lambda x - \Lambda(\lambda) \leq 0$, which implies $\Lambda^*(x) = \sup\{\lambda x - \Lambda(\lambda) ; \lambda \geq 0\}$

We have $\lambda x - \Lambda(\lambda) \leq \lambda m - \Lambda(\lambda) \leq \Lambda^*(m) = 0$

Case 2 $m = -\infty$, which implies $\Lambda(\lambda) = +\infty$ for $\lambda < 0$. Indeed, if $\Lambda(\lambda) < \infty$ for some $\lambda < 0$, we have $m > -\infty$. This is deduced from the fact that $E[e^{|\lambda| X^-}] < \infty$ and hence $1 + |\lambda| E(X^-) < \infty$. Hence $X^- \in L^1(\Omega)$.

Thus $\Lambda^*(x) = \sup\{\lambda x - \Lambda(\lambda) : \lambda \geq 0\} \vee (-\infty) = \sup\{\lambda x - \Lambda(\lambda) : \lambda \geq 0\}$

Since Λ is right-continuous at 0, we deduce

$$\Lambda^*(x) = \sup\{\lambda x - \Lambda(\lambda) ;\ \lambda > 0\} \text{ and } \Lambda^* \text{ increases on } (m, +\infty) . \qquad \square$$

<u>Exercises</u> 1.1 Prove that if $\mathcal{D}(\Lambda) = \{0\}$, $\Lambda^* \equiv 0$.

1.2 Compute Λ and Λ^* when $\mathcal{J} = \delta_a$. Prove that if $\mathcal{J}$ is not a Dirac mass, Λ is strictly convex.

1.3 Prove that if $\mathcal{D}(\Lambda) = \mathbb{R}$, $\lim_{|x| \to \infty} \frac{\Lambda^*(x)}{|x|} = +\infty$

1.4 Check the expressions of Λ and Λ^* when X has the following distribution:

- $\mathcal{B}(p)$, that is $P(X=1) = p$, $P(X=0) = 1-p$, $0 < p < 1$
- $\mathcal{P}(\theta)$ Poisson with parameter $\theta > 0$, $X\ \Omega \to \mathbb{N}$, $P(X=n) = e^{-\theta} \frac{\theta^n}{n!}$
- $\mathcal{E}(\theta)$ Exponential with parameter $\theta > 0$: $X : \Omega \to \mathbb{R}$ has the density $f(x) = \theta e^{-\theta x} 1_{(0,+\infty)}(x)$.

$\boxed{\text{Theorem 1.7}}$ (Cramer) Let $\{X_n\}_{n\geq 1}$ be an iid sequence of distribution μ. The sequence μ_n of distributions of $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ satisfies a LDP with rate function Λ^*.

<u>Remark</u> The upper estimate is more precise. For every closed set F,

$$\mu_n(F) \leq 2\exp(-n \inf\{\Lambda^*(x) : x\in F\}).$$

<u>Proof</u> : <u>Upper bound</u> Let F be a non empty closed set. If $\inf\{\Lambda^*(x), x\in F\}=0$, the upper bound is trivial. Suppose $\inf\{\Lambda^*(x) : x\in F\}>0$. Then Λ^* is not identically equal to 0, and $\mathcal{D}(\Lambda)\neq\{0\}$ (cf. Exercise 1.1).

Let $\lambda\geq 0$; then an exponential Markov inequality implies

$$P(\overline{X}_n\geq t)\leq E\left(e^{n\lambda(\overline{X}_n-t)}\right) = e^{-n\lambda t}\left(E\left(e^{\lambda X_1}\right)\right)^n$$

$$\leq \exp(-n\lambda t + n\Lambda(\lambda))$$

Suppose that $t>m$; then

$$P(\overline{X}_n\geq t)\leq \exp(-n\sup\{\lambda t-\Lambda(\lambda) : \lambda\geq 0\}) = \exp(-n\Lambda^*(t)),$$

where the last identity is deduced from Proposition 1.5 (iv).

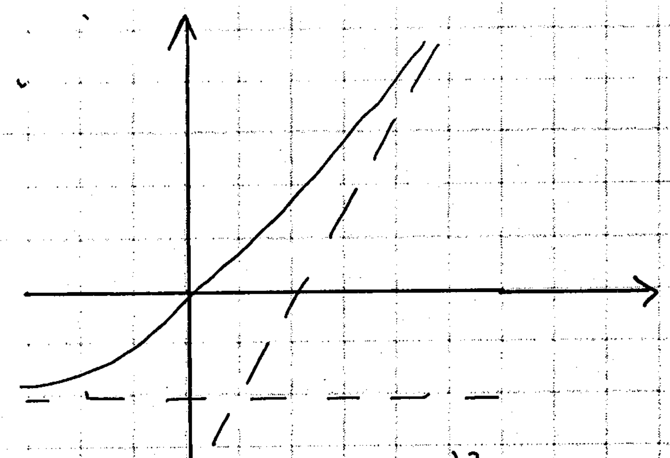
A similar computation shows that for $t<m$, $P(\overline{X}_n\leq t)\leq \exp(-n\Lambda^*(t))$.

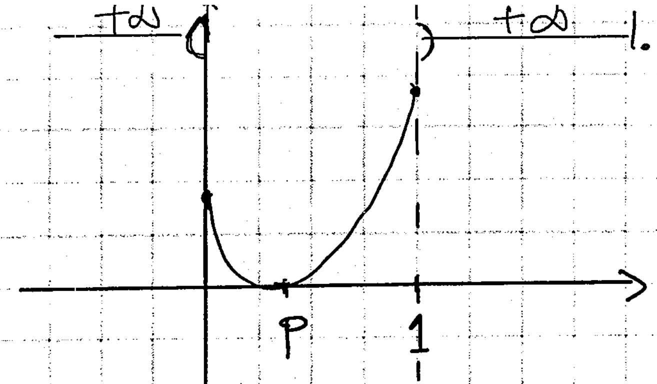
<u>Case 1</u> $m\in\mathbb{R}$; then $\Lambda^*(m)=0$. Let (a,b) be the connex component of the open set F^c containing m. Then $F\subset(-\infty,a]\cup[b,+\infty)$, $a,b\in F$ and

$$P(\overline{X}_n\in F)\leq P(\overline{X}_n\leq a)+P(\overline{X}_n\geq b)\leq \exp(-n\Lambda^*(a))+\exp(-n\Lambda^*(b))$$

$$\leq 2\exp(-n\inf\{\Lambda^*(x) : x\in F\}).$$

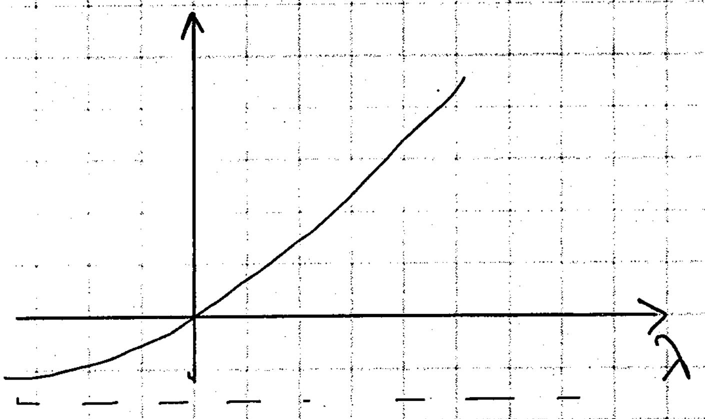
<u>Lower bound</u>. To focus on the main idea suppose that $\mathcal{D}(\Lambda)=\mathbb{R}$. Let G be a non empty open set, $x\in G$, $\delta>0$ such that $(x-\delta,x+\delta)\subset G$. Then

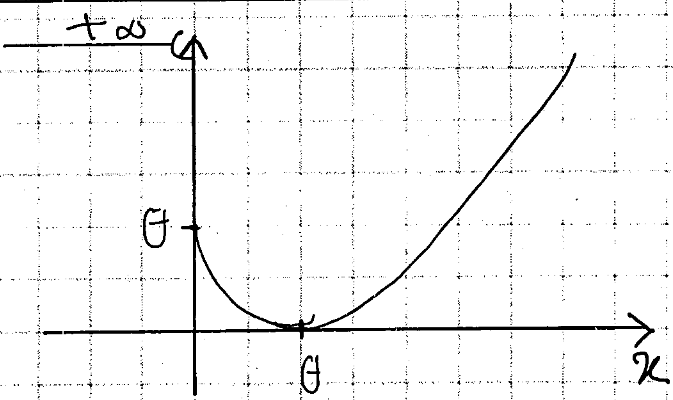




$B(p)$ $\quad \Lambda(\lambda) = \ln[(1-p)+pe^{\lambda}]$

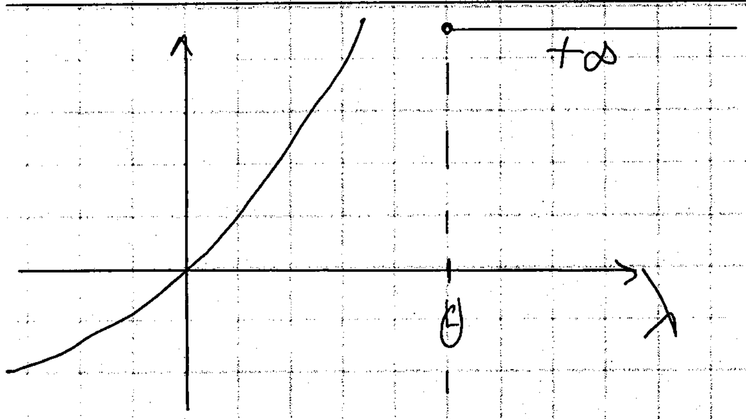
$$\Lambda^*(x) = \begin{cases} x\ln\left(\frac{x}{p}\right) + (1-x)\ln\left(\frac{1-x}{1-p}\right), & 0\le x\le 1 \\ +\infty & \text{otherwise} \end{cases}$$

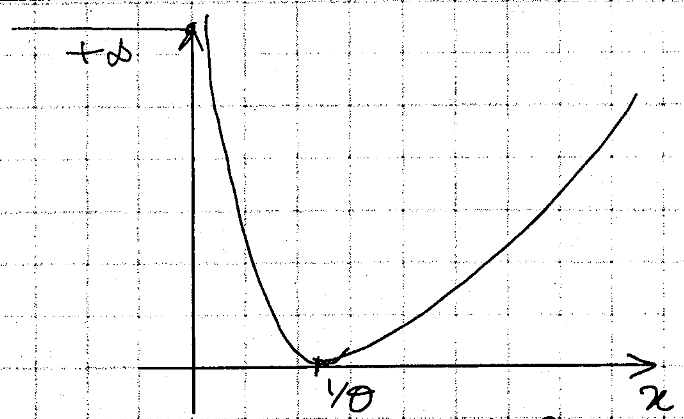




$\text{Poisson}(\theta)$ $\quad \Lambda(\lambda) = \theta(e^{\lambda}-1)$

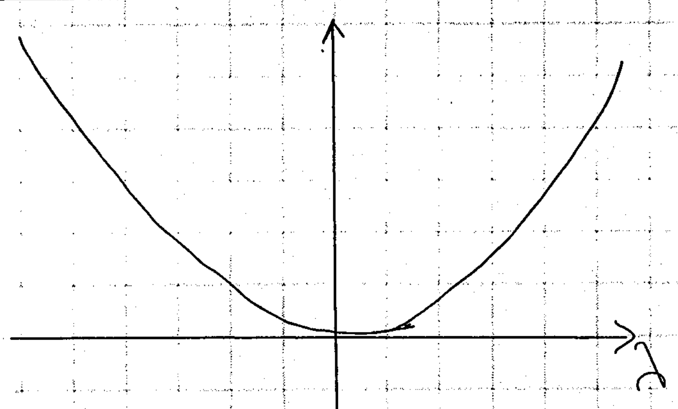
$$\Lambda^*(x) = \begin{cases} x\ln\left(\frac{x}{\theta}\right) + \theta - x, & x\ge 0 \\ +\infty, & x<0 \end{cases}$$

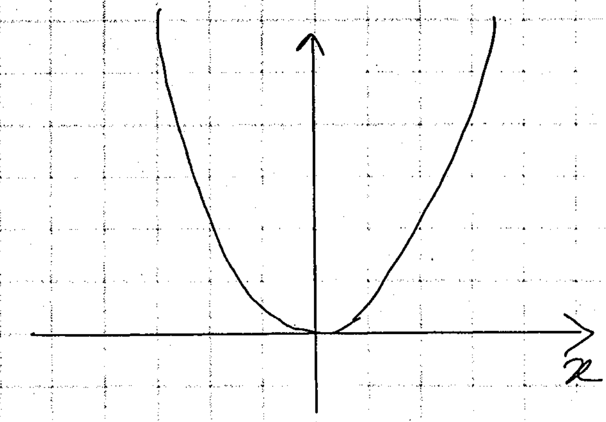




$\mathcal{E}(\theta)$ $\quad \Lambda(\lambda) = \begin{cases} \ln\left(\frac{\theta}{\theta-\lambda}\right), & \lambda<\theta \\ +\infty & \lambda\ge\theta \end{cases}$

$$\Lambda^*(x) = \begin{cases} \theta x - 1 - \ln(\theta x), & x>0 \\ +\infty & x\le 0 \end{cases}$$





$N(0,\sigma^2)$ $\quad \Lambda(\lambda) = \frac{\lambda^2\sigma^2}{2}$ $\quad \Lambda^*(x) = \frac{x^2}{2\sigma^2}$

$$\underline{\lim} \frac{1}{n} P(\bar{X}_n \in G) \geq \underline{\lim} \frac{1}{n} \ln P(|\bar{X}_n - x| < \delta)$$

We prove $\underline{\lim} \frac{1}{n} \ln P(|\bar{X}_n - x| < \delta) \geq -\Lambda^*(x)$; this will complete the proof.

Set $Y_1 = X_1 - x$; then $\Lambda_Y(\lambda) = \Lambda(\lambda) - \lambda x$ and $\Lambda_Y^*(y) = \Lambda^*(x+y)$.

We have to check that $\underline{\lim} \frac{1}{n} \ln P(|\bar{Y}_n| < \delta) \geq -\Lambda_Y^*(0)$.

Case 1. $P(Y>0)>0$ and $P(Y<0)>0$. Then Proposition 1.3 (iii) and (iv) imply $\lim_{|\lambda| \to \infty} \Lambda_Y(\lambda) = +\infty$. There exists $\eta \in \mathbb{R}$ such that $\Lambda_Y(\eta) = \inf\{\Lambda_Y(\lambda) : \lambda \in \mathbb{R}\}$.

Thus $\Lambda_Y'(\eta) = 0$ and $\Lambda_Y^*(0) = -\Lambda_Y(\eta)$

Let μ denote the distribution of Y_1; let $\tilde{\mu}$ be the probability defined on $\mathcal{B}(\mathbb{R})$ by $\frac{d\tilde{\mu}}{d\mu}(x) = \exp(\eta x - \Lambda_Y(\eta))$ and let (Z_n) be iid with distribution $\tilde{\mu}$. Then Z_1 is integrable and $E(Z_1) = \frac{E[Y_1 e^{\eta Y_1}]}{E[e^{\eta Y_1}]} = \Lambda_Y'(\eta) = 0$.

The law of large numbers implies $P(|\bar{Z}_n| < \varepsilon) \to 1$ as $n \to \infty$ for any $\varepsilon > 0$.

Let $\varepsilon \in (0, \delta)$; then

$$P(|\bar{Y}_n| < \varepsilon) = \int_{|\sum_{i=1}^n x_i| < n\varepsilon} d\mu(x_1) \cdots d\mu(x_n)$$

$$\geq e^{-n\varepsilon|\eta|} \int_{|\sum_{i=1}^n x_i| < n\varepsilon} \exp(\eta \sum_{i=1}^n x_i)\, d\mu(x_1) \cdots d\mu(x_n)$$

$$\geq e^{-n\varepsilon|\eta| + n\Lambda_Y(\eta)} \int_{|\sum_{i=1}^n x_i| < n\varepsilon} d\tilde{\mu}(x_1) \cdots d\tilde{\mu}(x_n)$$

$$\geq \exp(-n\varepsilon|\eta| + n\Lambda_Y(\eta))\, P(|\bar{Z}_n| < \varepsilon)$$

Hence for $\varepsilon \in (0, \delta)$,

$$\underline{\lim} \frac{1}{n} \ln P(|\bar{Y}_n| < \delta) \geq \underline{\lim} \frac{1}{n} \ln P(|\bar{Y}_n| < \varepsilon) \geq \Lambda_Y(\eta) - \varepsilon|\eta|$$

As $\varepsilon \to 0$, we deduce $\underline{\lim} \frac{1}{n} P(|\bar{Y}_n| < \delta) \geq \Lambda_Y(\eta) = -\Lambda^*(0)$.

<u>Case 2</u> $P(Y_1<0)=0$. Then Proposition 1.3 (iv) implies that Λ_Y is non decreasing and $\inf\{\Lambda_Y(\lambda):\lambda\in\mathbb{R}\}=\lim_{\lambda\to-\infty}\Lambda_Y(\lambda)=\ln P(Y_1=0)$

Then $P(|\overline{Y}_n|<\delta)\geq P(Y_1=0)^n$

and $\Lambda_Y^*(0)=\lim_{\lambda\to-\infty}-\Lambda_Y(\lambda)=-\ln P(Y_1=0)$.

This concludes the proof.

The case $P(Y_1>0)=0$ is dealt with in a similar way.

The general case $(\mathcal{D}(\Lambda)\neq\mathbb{R})$ is proven conditioning by the events $\{|X_1|\leq M\}$ and letting $M\to+\infty$. $\square$

<u>Exercise</u> (Hoeffding inequality) Suppose that $a\leq X_1\leq b$.

(a) Prove that for any $\lambda\in\mathbb{R}$,

$$E[e^{\lambda X_1}]\leq\frac{b-\bar{x}}{b-a}e^{\lambda a}+\frac{\bar{x}-a}{b-a}e^{\lambda b},\quad\text{where } \bar{x}=E(X)$$

(b) Deduce that for $x\in[a,b]$, $\Lambda^*(x)\geq H\left(\frac{x-a}{b-a}\,\Big|\,\frac{\bar{x}-a}{b-a}\right)$, where for $p_0\in(0,1)$, $p\in[0,1]$, $H(p|p_0)=p\ln\left(\frac{p}{p_0}\right)+(1-p)\ln\left(\frac{1-p}{1-p_0}\right)$.

Find a distribution of X_1, with support $[a,b]$, such that the above inequality is an equality

<u>Definition</u> Let μ and ν be probabilities on $\mathcal{F}$. The conditional entropy $H(\nu|\mu)$ is

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

<u>Exercise</u> Prove that

(i) If $X\sim B(p)$, $0<p<1$, $\Lambda^*(x)=x\ln\left(\frac{x}{p}\right)+(1-x)\ln\left(\frac{1-x}{1-p}\right)=H(B(x)\,|\,B(p))$ for $x\in[0,1]$

(ii) If $X\sim\mathcal{P}(\theta)$, $\theta>0$, $\Lambda^*(x)=\theta-x+x\ln\left(\frac{x}{\theta}\right)=H(\mathcal{P}(x)|\mathcal{P}(\theta))$ if $x\geq0$

(iii) If $X\sim\mathcal{E}(\theta)$, $\theta>0$, $\Lambda^*(x)=\theta x-1-\ln(\theta x)=H\left(\mathcal{E}\left(\frac{1}{x}\right)\,|\,\mathcal{E}(\theta)\right)$ if $x>0$

(iv) If $X\sim N(0,\sigma^2)$, $\Lambda^*(x)=\frac{x^2}{2\sigma^2}=H(N(x,\sigma^2)\,|\,N(0,\sigma^2))$.

1.4 Cramer's theorem in $\mathbb{R}^d$

We extend the previous theorem to the average $\overline{X}_n$ of an i.i.d sequence $(X_k, k\geq 1)$ of random vectors in $\mathbb{R}^d$.

To focus on the main arguments, we assume that $\mathcal{D}(\Lambda)=\mathbb{R}^d$, where $(x,y)=\sum_{i=1}^d x_i y_i$ for $x,y\in\mathbb{R}^d$, and

$\Lambda(\lambda)=\ln E[e^{(\lambda,X_1)}]$

For $x\in\mathbb{R}^d$, the Cramer transform of X is the Legendre transform of Λ, defined by

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

If $\mathcal{D}(\Lambda)=\mathbb{R}^d$, Λ is differentiable on $\mathbb{R}^d$, and if $y=\nabla\Lambda(\eta)$, $\Lambda^*(y)=(\eta,y)-\Lambda(\eta)$.

Furthermore, Λ is a convex function, $\Lambda^*:\mathbb{R}^d\longrightarrow[0,+\infty]$ is lsc; for $a\geq 0$, the level sets $\{\Lambda^*\leq a\}$ is a compact subset of $\mathbb{R}^d$. One says that Λ^* is a good rate function

Theorem 1.8 (Cramer in $\mathbb{R}^d$) Suppose that $\mathcal{D}(\Lambda)=\mathbb{R}^d$. Then the sequence of distributions of $\overline{X}_n$ satisfies a LDP with the good rate function Λ^*

The proof of upper bounds was based on inequality. It has to be changed

The idea is to replace upper bounds for closed sets by upper bounds for compact sets.

This requires the following notion.

Definition 1.9 (Exponential tightness) A family (μ_ε) of probabilities on a σ-field $\mathcal{G}\supset\mathcal{B}(X)$ is exponentially tight if for every $M>0$ there exists a compact set $K_M\subset X$ such that

$$\overline{\lim_{\varepsilon}}\ \varepsilon\ln\mu_\varepsilon(K_M^c)\leq -M.$$

We check that the sequence μ_n of distributions of $\overline{X}_n$ is exponentially tight.

For every $A>0$, $[-A,+A]^c = \bigcup_{j=1}^{d} \{ |x^j| > A \}$, where $x=(x^1,\dots,x^d)$

Let μ_n denote the distribution of $\overline{X}_n$ and μ_n^j the distribution of $\overline{X}_n^j$, the j^{th} component of $\overline{X}_n$. Then

$$\mu_n([-A,+A]^c) \leq \sum_{j=1}^{d} \mu_n^j([A,+\infty)) + \sum_{j=1}^{d} \mu_n^j((-\infty,A])$$
$$\leq \sum_{j=1}^{d} \{ e^{-n\Lambda_j^*(A)} + e^{-n\Lambda_j^*(-A)} \}$$

Using Exercise 1.3, we deduce that $\Lambda_j^*(z) \to +\infty$ as $|z| \to \infty$. Hence, given $M>0$ we can chose A such that for $j=1,\dots,d$, $\Lambda_j^*(A) > M$ and $\Lambda_j^*(-A) > M$

Hence $\mu_n([-A,+A]^c) \leq 2d\, e^{-nM}$, and $\overline{\lim} \frac{1}{n} \ln \mu_n([-A,+A]^c) \leq -M$.

We next prove that if (μ_ε) is exponentially tight, upper-estimates can be obtained dealing only with compact spaces, and that the rate function is "good"

We at first prove some technical lemma

Lemma 1.10 (i) Let $(a_i(\varepsilon),\ 1\leq i\leq N)$ be a family of strictly positive real numbers. Then

$$\overline{\lim}\ \varepsilon \ln\Big(\sum_{i=1}^{N} a_i(\varepsilon)\Big) = \max_{1\leq i\leq N} \overline{\lim}\ \varepsilon \ln(a_i(\varepsilon))$$

(ii) Let (a_ε) and (b_ε) be families of strictly positive real numbers. Then

$$\underline{\lim}\ \varepsilon \ln(a_\varepsilon + b_\varepsilon) \leq \max\big(\underline{\lim}\ \varepsilon \ln a_\varepsilon,\ \overline{\lim}\ \varepsilon \ln b_\varepsilon\big).$$

Proof (i) Since $a_i(\varepsilon) \leq \sum_j a_j(\varepsilon)$, $\overline{\lim}\ \varepsilon \ln(\sum_{i=1}^N a_i(\varepsilon)) \geq \max_{1\leq i\leq N} \overline{\lim}\ \varepsilon \ln(a_i(\varepsilon))$.

For $i=1,\dots,N$, let $l_i = \overline{\lim}\ \varepsilon \ln(a_i(\varepsilon))$. Fix $\delta>0$ and let $\varepsilon_0>0$ be such that for $\varepsilon \in (0,\varepsilon_0]$, $a_i(\varepsilon) \leq \exp\left(\frac{l_i+\delta}{\varepsilon}\right)$ for $i=1,\dots,N$.

Let $l = \max_{1\leq i\leq N} l_i$; then for $\varepsilon\in(0,\varepsilon_0]$

$$\sum_{i=1}^{N} a_i(\varepsilon) \leq \sum_{i=1}^{N} \exp\left(\frac{l_i+\delta}{\varepsilon}\right) \leq N \exp\left(\frac{l+\delta}{\varepsilon}\right).$$

This implies $\overline{\lim}\, \varepsilon \ln \left(\sum_{i=1}^{N} a_i(\varepsilon)\right) \leq \ell + \delta$. This proves (i) as $\delta \to 0$.

(ii) <u>Case 1</u> $\overline{\lim}\, \varepsilon \ln(b_\varepsilon) \leq \gamma \leq \underline{\lim}\, \varepsilon \ln(a_\varepsilon)$.

For any $\alpha > 0$, one can choose $\varepsilon_0 > 0$ such that for $\varepsilon \in (0, \varepsilon_0]$,

$\varepsilon \ln(b_\varepsilon) \leq \gamma + \alpha$ and $\varepsilon \ln(a_\varepsilon) \geq \gamma - \alpha$. This yields $b_\varepsilon \leq e^{\frac{\gamma+\alpha}{\varepsilon}}$, $a_\varepsilon \geq e^{\frac{\gamma-\alpha}{\varepsilon}}$

Hence $a_\varepsilon + b_\varepsilon \leq a_\varepsilon + a_\varepsilon e^{\frac{2\alpha}{\varepsilon}} \leq 2 a_\varepsilon e^{\frac{2\alpha}{\varepsilon}}$ for $\varepsilon \in (0, \varepsilon_0]$ and

$\varepsilon \ln(a_\varepsilon + b_\varepsilon) \leq \varepsilon \ln 2 + \varepsilon \ln a_\varepsilon + 2\alpha$

Hence $\underline{\lim}\, \varepsilon \ln(a_\varepsilon + b_\varepsilon) \leq \underline{\lim}\, \varepsilon \ln(a_\varepsilon) + 2\alpha$, which implies (ii) as $\alpha \to 0$

<u>Case 2</u> $\underline{\lim}\, \varepsilon \ln(a_\varepsilon) \leq \gamma = \overline{\lim}\, \varepsilon \ln(b_\varepsilon)$

Let $\varepsilon_n \to 0$ be such that $\lim_n \varepsilon_n \ln(a_{\varepsilon_n}) = \underline{\lim}\, \varepsilon \ln(a_\varepsilon) \leq \gamma$. Then

$\overline{\lim}\, \varepsilon_n \ln(b_{\varepsilon_n}) \leq \overline{\lim}\, \varepsilon \ln(b_\varepsilon) = \gamma$

Fix $\alpha > 0$; there exists N such that for $n \geq N$, we have $\varepsilon_n \ln(a_{\varepsilon_n}) \leq \gamma + \varepsilon$ and

$\varepsilon_n \ln(b_{\varepsilon_n}) \leq \gamma + \varepsilon$. This implies

$\varepsilon_n \ln(a_{\varepsilon_n} + b_{\varepsilon_n}) \leq \varepsilon_n \ln\left(2 e^{\frac{\gamma+\alpha}{\varepsilon_n}}\right) \leq \varepsilon_n \ln(2) + \gamma + \alpha$

and hence

$\underline{\lim}\, \varepsilon \ln(a_\varepsilon + b_\varepsilon) \leq \underline{\lim}_n\, \varepsilon_n \ln(a_{\varepsilon_n} + b_{\varepsilon_n}) \leq \gamma + \alpha$.

This implies (ii) as $\alpha \to 0$. □

We next prove the following

<u>Proposition 1.11</u> Let (μ_ε) be an exponentially tight family of probabilities on $\mathcal{B}(X)$. Then

(i) If the upper estimate holds for compact sets, then it holds for closed sets.

(ii) If the lower estimate holds for open sets, then I is a good rate function. (that is I has compact level sets.)

Proof (i) Let F be closed. If $\inf\{I(x) : x\in F\} = 0$, the upper bound is trivial.
Suppose that $\inf\{I(x) : x\in F\} = R > 0$ (it can be $+\infty$). Let $M\in(0,R)$ and let K_M be a compact set such that $\overline{\lim}\, \varepsilon \ln \mu_\varepsilon(K_M^c) \leq -M$. The set $K_M\cap F$ is compact and $\inf\{I(x) : x\in K_M\cap F\} \geq \inf\{I(x) : x\in F\}$. Lemma 1.8 (i) implies

$$\overline{\lim}\, \varepsilon \ln \mu_\varepsilon(F) \leq \max\left(\overline{\lim}\, \varepsilon \ln \mu_\varepsilon(F\cap K_M),\ \overline{\lim}\, \varepsilon \ln \mu_\varepsilon(K_M^c)\right)$$

$$\leq \max(-R, -M) = -M.$$

Letting M increase to R, we conclude the proof of (i).

(ii) Fix $a>0$, $M>a$ and let K_M be a compact set such that $\overline{\lim}\, \varepsilon \ln \mu_\varepsilon(K_M^c) < -M$.
Then K_M^c is open, and the lower bound implies

$$-M \geq \underline{\lim}\, \varepsilon \ln \mu_\varepsilon(K_M^c) \geq -\inf\{I(x) : x\in K_M^c\}$$

This implies $\inf\{I(x) : x\in K_M^c\} \geq M > a$, that is $\{I\leq a\} \subset K_M$. Since I is l.s.c, $\{I\leq a\}$ is closed and hence compact. $\square$

The lower bounds are proved using a change of probability similar to that in dimension 1. It is enough to prove that if $\Lambda^*(y) < \infty$, for any $\delta > 0$,

$$\underline{\lim}\, \frac{1}{n} \ln P(\overline{X}_n \in B(y,\delta)) \geq -\Lambda^*(y).$$

If $y = \nabla\Lambda(\eta)$, one uses a change of probability similar to that in dimension 1.
If $y \notin \{\nabla\Lambda(\lambda) : \lambda\in\mathbb{R}^d\}$ let (G_n) be iid $N(0, Id)$ such that (X_n) and (G_n) are independent. Fix $M>0$ and set $Y_n = X_n + \frac{G_n}{\sqrt{M}}$. Then the logarithmic moment generating function Λ_M of Y_1 is

$$\Lambda_M(\lambda) = \Lambda(\lambda) + \frac{|\lambda|^2}{2M} \geq \Lambda(\lambda), \text{ hence } \Lambda^*(y) = \sup\{(\lambda,y) - \Lambda(\lambda)\} \geq \sup\{(\lambda,y) - \Lambda_M(\lambda)\} = \Lambda_M^*(y)$$

Set $g(\lambda) = (\lambda, y) - \Lambda_M(\lambda)$. Then $\lim_{\rho \to \infty} \sup_{|\lambda| \geq \rho} g(\lambda) = -\infty$. Hence g achives its supremum at some point $\eta \in \mathbb{R}^d$ such that $0 = \nabla g(\eta) = y - \nabla \Lambda_M(\eta)$.

The preceeding argument shows that $\underline{\lim} \frac{1}{n} \ln P(\overline{Y}_n \in B(y,\delta)) \geq -\Lambda^*_M(y)$

Furthermore $\overline{Y}_n = \overline{X}_n + \overline{G}_n \overset{(d)}{=} \overline{X}_n + \frac{G}{\sqrt{Mn}}$, where G is $N(0, I_d)$ independent of $\overline{X}_n$. Hence $\{\overline{Y}_n \in B(y,\delta)\} \cap \{\frac{G}{\sqrt{Mn}} \in B(0,\delta)\} \subset \{\overline{X}_n \in B(y, 2\delta)\}$

and $P(\overline{X}_n \in B(y, 2\delta)) \geq P(\overline{Y}_n \in B(y,\delta)) - P(\|G\| \geq \delta\sqrt{Mn})$

Since $G \sim N(0, I_d)$, $\overline{\lim} \frac{1}{n} \ln P(\|G\| \geq \delta\sqrt{Mn}) \leq -\frac{M\delta^2}{2}$, that is

for n large enough $P(\|G\| \geq \delta\sqrt{Mn}) \leq \exp(-\frac{nM\delta^2}{2})$. Hence

Let $\alpha > 0$; for n large enough

$P(\overline{Y}_n \in B(y,\delta)) \geq \exp(-n[\Lambda^*_M(y) + \alpha]) \geq \exp(-n[\Lambda^*(y) + \alpha])$; Lemma 1.10(ii) implies

$-\Lambda^*(y) - \alpha \leq \max(\underline{\lim} \frac{1}{n} \ln P(\overline{X}_n \in B(y, 2\delta)), -\frac{n\delta^2}{2})$.

Let $M \to \infty$, then $\alpha \to 0$; we conclude that $\underline{\lim} \frac{1}{n} \ln P(\overline{X}_n \in B(y, 2\delta)) \geq -\Lambda^*(y)$.

The proof of the upper bound for compact sets will be done in the following theorem, in a more general framework.

<u>Exercise</u> Let $Y = (Y_1, \dots, Y_d)$ be a Gaussian vector with mean 0 and Covariance matrix I_d. Prove that $(\frac{1}{\sqrt{n}} Y)$ satisfies a LDP with a good rate function to be given explicitely.

1.5 The Ellis Gärtner theorem

Let (μ_n) be a sequence of probabilities on $\mathcal{B}(\mathcal{X})$, where $\mathcal{X}$ is a topological vector space. Let $\mathcal{X}'$ denote the dual of $\mathcal{X}$. For $\lambda \in \mathcal{X}'$ set

$$\Lambda_n(\lambda) = \ln\left(\int_{\mathcal{X}} e^{\langle \lambda, x\rangle} \mu_n(dx) \right)$$

where $\langle \cdot, \cdot \rangle$ denote the duality between $\mathcal{X}$ and $\mathcal{X}'$. Proofs similar to that in Section 1.3 prove that Λ_n is convex, l.s.c., takes values in $(-\infty, +\infty]$, and $\Lambda_n(0)=0$.

In this section, we make the following assumption

Condition (C) For every $\lambda \in \mathcal{X}'$, $\Lambda(\lambda) = \lim_n \frac{1}{n} \Lambda_n(n\lambda) \in [-\infty, +\infty]$ exists.

Λ is convex and l.s.c.

<u>Example</u> $\mathcal{X} = \mathbb{R}^d$, μ_n is the distribution of $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$, where $(X_k, k\geq 1)$ is i.i.d. Then $\Lambda_n(n\lambda) = \ln E\left[e^{n\langle \lambda, \frac{1}{n}\sum_{k=1}^{n} X_k\rangle}\right] = n \ln E[e^{\langle \lambda, x\rangle}] = n\Lambda(\lambda)$

Hence $\frac{1}{n}\Lambda_n(n\lambda) = \Lambda(\lambda)$ for every n

<u>Remark</u> The aim of condition (C) is to generalize averages of i.i.d random variables to averages of more general sequences such as Markov Chains

Let $\Lambda^* : \mathcal{X} \to [0, +\infty]$ denote the Legendre transform of Λ, that is

$$\Lambda^*(x) = \sup\{ \langle \lambda, x\rangle - \Lambda(\lambda) : \lambda \in \mathcal{X}' \}$$

Once more, the arguments used in section 1.3 prove that Λ^* is convex, l.s.c.

Theorem 1.12 (<u>Upper estimates for compact sets</u>)

Let (μ_n) be a sequence of probabilities on $\mathcal{B}(\mathcal{X})$ such that the condition (C) is satisfied. Then for every compact set $K \subset \mathcal{X}$, $\overline{\lim} \frac{1}{n} \ln \mu_n(K) \leq -\inf\{\Lambda^*(x) : x \in K\}$

Proof. Case 1. Suppose that there exists $\lambda_0 \in \mathcal{X}'$ such that $\Lambda(\lambda_0) = -\infty$ 1.19

Then $\Lambda^* \equiv +\infty$ and since $x \in \mathcal{X} \mapsto \langle \lambda_0, x\rangle$ is continuous and K is compact, there exists $M \in \mathbb{R}$ such that $K \subset \{x \in \mathcal{X} : \langle \lambda_0, x\rangle \geq -M\}$. This implies

$$\mu_n(K) \leq \mu_n(\{x \in \mathcal{X} : \langle \lambda_0, x\rangle \geq -M\}) \leq e^{nM} \int e^{\langle n\lambda_0, x\rangle} \mu_n(dx) = e^{nM + \Lambda_n(n\lambda_0)}$$

and using condition (c),

$$\overline{\lim} \frac{1}{n} \ln \mu_n(K) \leq M + \overline{\lim} \frac{1}{n} \Lambda_n(n\lambda_0) = -\infty$$

This proves the upper estimate for K

Case 2. The limiting map $\Lambda : \mathcal{X}' \longrightarrow (-\infty, +\infty]$. For $\delta > 0$, let

$$I^\delta(x) = \inf\{\Lambda^*(x) - \delta, \tfrac{1}{\delta}\} \nearrow \Lambda^*(x) \quad \text{as} \quad \delta \searrow 0.$$

We prove that $\overline{\lim} \frac{1}{n} \ln \mu_n(K) \leq \delta - \inf\{I^\delta(x) : x \in K\}$; this will conclude the proof as $\delta \to 0$.

Given $x \in K$, let $\lambda_x \in \mathcal{X}'$ be such that

$$\langle \lambda_x, x\rangle - \Lambda(\lambda_x) > I^\delta(x), \text{ which implies } \Lambda(\lambda_x) < \infty.$$

Set $O_x = \{y \in \mathcal{X} : \langle \lambda_x, y\rangle - \Lambda(\lambda_x) > I^\delta(x)\}$. Then O_x is an open set containing x and $\bigcup_{x \in K} O_x \supset K$. We can extract finitely many open sets $O_{x_1}, \dots, O_{x_N}$ such that $K \subset \bigcup_{i=1}^N O_{x_i}$. Then

$$\mu_n(K) \leq \sum_{i=1}^N \mu_n(\{y : e^{n\langle \lambda_{x_i}, y\rangle - n\Lambda(\lambda_{x_i})} > e^{nI^\delta(x_i)}\})$$

$$\leq \sum_{i=1}^n e^{-nI^\delta(x_i) - n\Lambda(\lambda_{x_i})} \int e^{\langle n\lambda_{x_i}, y\rangle} d\mu_n(y)$$

$$\leq \sum_{i=1}^N e^{-nI^\delta(x_i) - n\Lambda(\lambda_{x_i}) + \Lambda_n(n\lambda_{x_i})}$$

Lemma 1.8 (i) implies that

$$\overline{\lim} \frac{1}{n} \ln \mu_n(K) \leq \max_{1 \leq i \leq N} \left\{ -\underset{\delta}{I}(x_i) - \Lambda(\lambda_{x_i}) + \overline{\lim} \frac{1}{n} \Lambda_n(n\lambda_{x_i}) \right\}$$

$$\leq \max_{1 \leq i \leq N} \{ -\underset{\delta}{I}(x_i) \} \leq -\inf \{ \underset{\delta}{I}(x) : x \in K \}. \qquad \square$$

<u>Corollary 1.13</u> If (μ_n) is exponentially tight and satisfies the condition (c), then the upper estimate holds for closed sets, and $\Lambda : \mathcal{X}' \longrightarrow (-\infty, +\infty]$

<u>Proof</u> The upper estimate for closed sets is a consequence of Proposition 1.9 (i) and of Theorem 1.10. Furthermore, let $M > 0$ and K_M be a compact set of $\mathcal{X}$ such that $\overline{\lim} \frac{1}{n} \ln \mu_n(K_M^c) \leq -M$. Using case 1 in the proof of Theorem 1.10, we deduce that if $\Lambda(\lambda_0) = -\infty$ for some $\lambda_0 \in \mathcal{X}'$, $\overline{\lim} \frac{1}{n} \ln \mu_n(K) = -\infty$.

Lemma 1.8 (i) implies that

$$0 = \overline{\lim} \frac{1}{n} \ln \mu_n(\mathcal{X}) \leq \max(-M, -\infty) = -M$$

which yields a contradiction. $\square$

We next have to prove lower bounds for open sets. Since we cannot apply the law of large numbers, this requires a new idea. We will still use a change of probability.

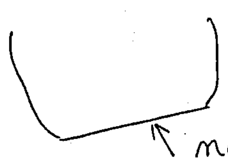
<u>Definition 1.14</u> An element x of $\mathcal{X}$ is an exposed point of Λ^* if there exists $\eta \in \mathcal{X}'$ such that

$$\Lambda^*(y) - \langle \eta, y \rangle > \Lambda^*(x) - \langle \eta, x \rangle \ , \quad \forall y \neq x$$

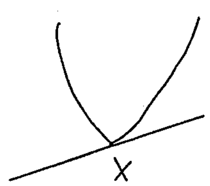
(that is for $y \neq x$, $\Lambda^*(y) > \Lambda^*(x) + \langle \eta, y - x \rangle$.)

The function Λ^* is strictly convex at x and η is an exposing hyperplane for x

<u>Remark</u> One want to avoid "flat" curves



↑ non exposed points



x

x exposed point (η not unique)

Proposition 1.15 Let (μ_n) be an exponentially tight sequence satisfying condition (c).

Let $\mathcal{E} = \{x \in \mathcal{X} : x$ exposed point of Λ^*, $\exists \eta \in \mathcal{X}'$ exposing hyperplane for x, $\exists s > 1 : \Lambda(s\eta) < \infty\}$

Then for any open set G of $\mathcal{X}$, $\underline{\lim}\, \frac{1}{n} \ln \mu_n(G) \geq -\inf\{\Lambda^*(x) : x \in G \cap \mathcal{E}\}$

Proof Let $x \in \mathcal{E} \cap G$, $\eta \in \mathcal{X}'$ an exposing hyperplane for x and $s_0 > 1$ be such that $\Lambda(s_0 \eta) < \infty$. For $\delta > 0$, let

$B_\delta = G \cap \{y \in \mathcal{X} : |\langle \eta, y - x \rangle| < \delta\}$ be an open neighborhood of x.

To prove Proposition 1.13, it suffices to check that

$$\underline{\lim}\, \frac{1}{n} \ln \mu_n(B_\delta) \geq -\Lambda^*(x) - \delta$$

Indeed, $\underline{\lim}\, \frac{1}{n} \ln \mu_n(G) \geq \underline{\lim}\, \frac{1}{n} \ln \mu_n(B_\delta) \geq -\inf\{\Lambda^*(x) : x \in G \cap \mathcal{E}\} - \delta$

and the proof is complete as $\delta \to 0$.

Note that since Λ is convex, $\Lambda(0) = 0$ and $\Lambda(s_0\eta) < \infty$, we have $\Lambda(s\eta) < \infty$ for $s \in [0, s_0]$

Hence $\Lambda(\eta) < \infty$ and Condition (c) implies the existence of n_0 such that $\Lambda_n(n\eta) < \infty$ for $n \geq n_0$.

Step 1 Change of probability For $n \geq n_0$, let $\tilde{\mu}_n$ be defined by

$\frac{d\tilde{\mu}_n}{d\mu_n}(z) = \exp(n\langle z, \eta\rangle - \Lambda_n(n\eta))$; then for any $\delta > 0$,

$$\begin{aligned}
\frac{1}{n} \ln \mu_n(B_\delta) &= \frac{1}{n} \ln \int_{B_\delta} e^{-n\langle z, \eta\rangle + \Lambda_n(n\eta)}\, d\tilde{\mu}_n(z) \\
&= \frac{1}{n} \ln \int_{B_\delta} e^{\Lambda_n(n\eta) - n\langle x, \eta\rangle - n\langle z - x, \eta\rangle}\, d\tilde{\mu}_n(z) \\
&\geq \frac{1}{n} \Lambda_n(n\eta) - \langle x, \eta\rangle + \ln \int_{B_\delta} e^{-n\delta}\, d\tilde{\mu}_n(z) \\
&\geq \frac{1}{n} \Lambda_n(n\eta) - \langle x, \eta\rangle - \delta + \frac{1}{n} \ln \tilde{\mu}_n(B_\delta)
\end{aligned}$$

This implies $\underline{\lim}\, \frac{1}{n} \ln \mu_n(B_\delta) \geq \Lambda(\eta) - \langle x, \eta\rangle - \delta + \underline{\lim}\, \frac{1}{n} \ln \tilde{\mu}_n(B_\delta)$

$$\geq -\Lambda^*(x) - \delta + \underline{\lim}\, \frac{1}{n} \ln \tilde{\mu}_n(B_\delta).$$

The proof will be complete if we can check that $\lim \tilde{\mu}_n(B_\delta) = 1$ (which will replace the law of large numbers.) $\boxed{1.22}$

<u>Step 2</u> We prove that $(\tilde{\mu}_n)$ is exponentially tight. The Laplace transform of $\tilde{\mu}_n$ is

$$\mathcal{L}(\tilde{\mu}_n)(\lambda) = \int e^{\langle \lambda, x\rangle + \langle \eta, x\rangle} \mu_n(dx)\, e^{-\Lambda_n(\eta)}$$
$$= e^{\Lambda_n(\lambda+\eta) - \Lambda_n(\eta)}$$

Hence $\tilde{\Lambda}_n(\lambda) = \ln \mathcal{L}(\tilde{\mu}_n)(\lambda) = \Lambda_n(\lambda+\eta) - \Lambda_n(\eta)$, and $\frac{1}{n}\Lambda_n(n\lambda) \to \tilde{\Lambda}(\lambda) = \Lambda(\lambda+\eta) - \Lambda(\eta)$.

Let ν_n be the image distribution of $\tilde{\mu}_n$ by the map $y \mapsto \langle \eta, y\rangle$. Its Laplace transform $\mathcal{L}(\nu_n)(t) = \mathcal{L}(\tilde{\mu}_n)(t\eta) = \exp\left(\Lambda_n[(t+1)\eta] - \Lambda_n(\eta)\right)$ satisfies

$$\lim \frac{1}{n} \ln[\mathcal{L}(\nu_n)(t)] = \tilde{\Lambda}_\eta(t) = \Lambda((t+1)\eta) - \Lambda(\eta)$$

Since $\Lambda(s\eta) < \infty$ for $0 \le s < s_0$ with $s_0 > 1$, $0 \in \mathring{\mathcal{D}}(\tilde{\Lambda}_\eta)$

We next prove that (ν_n) is exponentially tight. Let $\theta > 0$ be such that

$\tilde{\Lambda}_\eta(\theta) < \infty$ and $\tilde{\Lambda}_\eta(-\theta) < \infty$. For any $t > 0$,

$$\nu_n((t,+\infty)) = \nu_n\{y : e^{n\theta y} \ge e^{n\theta t}\} \le e^{-n\theta t}\,\mathcal{L}(\nu_n)(n\theta)$$

Hence $\overline{\lim_n} \frac{1}{n} \ln \nu_n((t,+\infty)) \le -\theta t + \tilde{\Lambda}_\eta(\theta) \le -M$

for fixed M and t large enough. A similar computation implies $\overline{\lim} \frac{1}{n} \ln \nu_n((-\infty,-t)) \le -M$ for t large enough.

Set $C_{\eta,t} = \{x \in \mathcal{X} : |\langle \eta, x\rangle| \le t\}$; using Lemma 1.8 (i) and the definition of ν_n we deduce that given $M > 0$, for t large enough $\overline{\lim} \ln \tilde{\mu}_n(C_{\eta,t}^c) \le -M$.

Recall that (μ_n) is exponentially tight. Hence for any $R > 0$, there exists a compact set K_R of $\mathcal{X}$ such that $\overline{\lim} \frac{1}{n} \ln \mu_n(K_R^c) \le -R$.

We upper estimate $\tilde{\mu}_n(K_R^c)$.

For any $R>0$,

$\tilde{\mu}_n(K_R^c) \leq \tilde{\mu}_n(K_R^c \cap C_{\eta,t}) + \tilde{\mu}_n(C_{\eta,t}^c)$, and

$\tilde{\mu}_n(K_R \cap C_{\eta,t}) = \int 1_{K_R^c \cap C_{\eta,t}}(x)\, e^{n\langle \eta, x\rangle - \Lambda_n(n\eta)}\, dx \leq e^{nt - \Lambda_n(n\eta)}\, \mu_n(K_R^c)$

Hence Lemma 1.10(i) implies

$\overline{\lim} \frac{1}{n} \ln \tilde{\mu}_n(K_R^c) \leq \max(-M, t - \Lambda(\eta) - R)$

Fix M; choose t and then let $R = M + t - \Lambda(\eta)$. Then $\overline{\lim} \frac{1}{n}\ln\tilde{\mu}_n(K_R^c) \leq -M$.

<u>Step 3</u> We prove that $\lim_n \tilde{\mu}_n(B_\delta) = 1$. Since $(\tilde{\mu}_n)$ satisfies condition (c) and is exponentially tight, Corollary 1.11 implies that it satisfies upper bounds for closed sets with the rate function $\tilde{\Lambda}^*$, where

$\tilde{\Lambda}^*(y) = \sup\{\langle\lambda, y\rangle - \tilde{\Lambda}(\lambda);\ \lambda \in \mathcal{X}'\} = \sup\{\langle \lambda, y\rangle - \Lambda(\lambda+\eta) + \Lambda(\eta):\ \lambda \in \mathcal{X}'\}$

$= \sup\{\langle \lambda+\eta, y\rangle - \Lambda(\lambda+\eta) - \langle \eta, y\rangle + \Lambda(\eta):\ \lambda \in \mathcal{X}'\}$

$= \Lambda^*(y) - \langle \eta, y\rangle + \Lambda(\eta)$

If $y \neq x$ we have $\tilde{\Lambda}^*(y) > 0$. Indeed x is exposed with exposing hyperplane η; hence $\Lambda^*(y) - \langle \eta, y\rangle > \Lambda^*(x) - \langle \eta, x\rangle \geq \langle \eta, x\rangle - \Lambda(\eta) - \langle \eta, x\rangle = -\Lambda(\eta)$ which implies $\Lambda^*(y) - \langle \eta, y\rangle + \Lambda(\eta) > 0$.

Step 2 implies that given any $M>0$, there exists a compact set K_M such that $\overline{\lim} \frac{1}{n} \ln \tilde{\mu}_n(K_M^c) \leq -M$, and hence $\overline{\lim} \frac{1}{n} \ln \tilde{\mu}_n(B_\delta^c \cap K_M^c) \leq -M$.

Using the upper estimate for the compact set $B_\delta^c \cap K_M$, we deduce that $\overline{\lim} \frac{1}{n} \ln \tilde{\mu}_n(B_\delta^c \cap K_M) \leq -\inf\{\tilde{\Lambda}^*(x):\ x \in B_\delta^c \cap K_M\}$. Since $x \notin B_\delta^c \cap K_M$, and since $\tilde{\Lambda}^*$ is l.s.c, strictly positive on the compact set $B_\delta^c \cap K_M$ it achieves its minimum on that set, which is hence a strictly positive number c.

Hence using once more Lemma 1.10 (i), we deduce

$\overline{\lim} \frac{1}{n} \ln \tilde{\mu}_n (B_\delta^c) \leq \max(-M, -c) < 0$. This concludes the proof. $\square$

1.6 The Ellis-gärtner theorem in finite dimension

We state here sufficient conditions implying exponential tightness of (μ_n) and lower bounds for open sets. In this section, we suppose that $\mathcal{X} = \mathbb{R}^d$.

Condition (H) Suppose $\lim \frac{1}{n} \Lambda_n(n\lambda) = \Lambda(\lambda) \in (-\infty, +\infty]$ and $0 \in \overset{\circ}{\overline{\mathcal{D}(\Lambda)}}$, $\forall \lambda \in \mathbb{R}^d$

Proposition 1.16 Suppose that (μ_n) satisfies condition (H). Then (μ_n) is exponentially tight

Proof Let $(e_j, j = 1, .., d)$ denote the canonical basis of $\mathbb{R}^d$. There exists $\alpha > 0$ such that for $j = 1, .., d$, $\Lambda(\alpha e_j) < \infty$ and $\Lambda(-\alpha e_j) < \infty$. Let μ_n^j denote the j^{th} marginal of μ_n, $j = 1, .., d$. Then for any $t > 0$

$$\mu_n^j([t, +\infty)) \leq \int_t^\infty e^{n\alpha(x-t)} \, d\mu_n^j(x) \leq \exp(-n\alpha t + \Lambda_n(n\alpha e_j))$$

and $\overline{\lim} \frac{1}{n} \ln \mu_n^j([t, +\infty)) \leq -\alpha t + \Lambda(\alpha e_j)$

A similar argument proves that $\overline{\lim} \frac{1}{n} \ln \mu_n^j((-\infty, -t]) \leq -\alpha t + \Lambda(-\alpha e_j)$

Hence $\overline{\lim} \frac{1}{n} \ln \mu_n(([-t, +t]^d)^c) \leq -\alpha t + \max_{1 \leq j \leq d} (\Lambda(\alpha e_j) \vee \Lambda(-\alpha e_j))$

Taking t large enough, we conclude the proof. □

Lemma 1.17 Suppose that Λ defined by condition (H) is differentiable on $\overset{\circ}{\overline{\mathcal{D}(\Lambda)}}$. Let $x = \nabla\Lambda(\eta)$, $\eta \in \overset{\circ}{\overline{\mathcal{D}(\Lambda)}}$. Then $\Lambda^*(x) = \langle \eta, x \rangle - \Lambda(\eta)$, $x \in \mathcal{E}$ and η is an exposing hyperplane such that $\Lambda(s\eta) < \infty$ for some $s > 1$.

Proof Let $x = \nabla\Lambda(\eta)$, $\eta \in \overset{\circ}{\overline{\mathcal{D}(\Lambda)}}$ and let $g(\lambda) = \langle x, \lambda \rangle - \Lambda(\lambda)$. Then g is concave and reaches its maximal at a critical point (that is $\nabla g(\lambda) = 0$).

Clearly η is critical and $\Lambda^*(x) = \langle \eta, x \rangle - \Lambda(\eta)$.

Let $y \in \mathcal{X}$ be such that $\Lambda^*(y) \le \Lambda^*(x) + \langle \eta, y - x \rangle$; we prove that $y = x$.

We have $\Lambda^*(y) \le \langle \eta, x\rangle - \Lambda(\eta) + \langle \eta, x - y\rangle = \langle \eta, y\rangle - \Lambda(\eta)$

By definition of the Legendre transform $\Lambda^*(y) \ge \langle \eta, y\rangle - \Lambda(\eta)$. Hence $\Lambda^*(y) = \langle \eta, y\rangle - \Lambda(\eta)$. The map $\theta \mapsto \langle \theta, y\rangle - \Lambda(\theta)$ achieves its maximum at η which is a critical point. Hence $y = \nabla\Lambda(\eta) = x$

Since $\eta \in \mathring{\mathcal{D}}(\Lambda)$, there exists $s > 1$ such that $\Lambda(s\eta) < \infty$. □

This yields the following

Theorem 1.18 Let $\mathcal{X} = \mathbb{R}^d$, (μ_n) be a sequence of probabilities on $\mathcal{B}(\mathbb{R}^d)$ which satisfies the condition (C), and suppose that $\mathcal{D}(\Lambda) = \mathbb{R}^d$, and that Λ is differentiable. Then (μ_n) is exponentially tight, and satisfies a LDP with the good rate function Λ^*.

The proof is left as an exercise. Using Proposition 1.14 and Lemma 1.15, it suffices to check lower bounds for open sets. This can be done using the perturbation by a small multiple of an independent Gaussian vector as in the proof of Theorem 1.8 (Cramer's theorem in $\mathbb{R}^d$).

<u>Exercise</u> The aim of this example is to prove that, because of possible variations of the domains, the first terms in an average may not be disregarded

Let $(X_n, n \ge 1)$ be a sequence of independent r.v such that

(i) X_1 has an expontial distribution of parameter $\theta_1 > 0$,

(ii) For $n \ge 2$, X_n has an exponential distribution of parameter $\theta_2 > \theta_1$.

Set $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\overline{Y}_n = \frac{1}{n-1}\sum_{i=2}^{n} X_i$.

1) Compute $\tilde{\Lambda}(\lambda) = \ln E[e^{\lambda X_2}]$ for $\lambda \in \mathbb{R}$ and its Legendre transform $\tilde{\Lambda}^*$.

2) Compute $\Lambda_n(\lambda) = \ln E[e^{\lambda \overline{X}_n}]$ for $\lambda \in \mathbb{R}$, $n \geq 2$.

3) Prove that $\Lambda(\lambda) = \lim \frac{1}{n} \Lambda_n(n\lambda)$ exists and compute this limit explicitely.

4) Compute the Legendre transform Λ^* of Λ and draw Λ^* and $\tilde{\Lambda}^*$ on the same picture.

5) Let F be a closed subset of $\mathbb{R}$. Give an upper limit of $\overline{\lim} \frac{1}{n} \ln \mu_n(F)$.

6) Prove the existence of $x_0 \in \mathbb{R}$ such that for any $x \in (0, x_0)$ and any $\delta > 0$,

$$\underline{\lim} \frac{1}{n} \ln P(|\overline{X}_n - x| < \delta) \geq -I(x)$$

for some function I to be determined.

7) Let $x > x_0$ and $\delta_0 > 0$ be such that $x - \delta_0 > x_0$.

a) Prove that for all $\delta \in (0, \delta_0)$

$$\sup_{\varepsilon \in (0,\, x - \delta - \frac{1}{\theta_2})} -\theta_1 \varepsilon - \tilde{\Lambda}^*(x - \delta - \varepsilon) = -\Lambda^*(x - \delta).$$

b) Prove that for $0 < \delta < \delta_0$ and $\eta \in (0, x - \delta_0 - x_0)$

$$\underline{\lim} \frac{1}{n} \ln P(\overline{X}_n > x - \delta) \geq -\tilde{\Lambda}^*(x - \delta - \eta) - \theta_1 \eta$$

c) Prove that for $\delta \in (0, \delta_0]$, $\underline{\lim} \frac{1}{n} \ln P(|\overline{X}_n - x| < \delta) \geq -\Lambda^*(x - \delta)$

8) Deduce that the sequence of distributions of $(\overline{X}_n)$ satisfies a LDP with a good rate function to be described.

9) Suppose now that $(X_n, n \geq 1)$ is independent, that there exist an integer k such that the random variables X_i, $1 \leq i \leq k$ have an exponential distribution with parameter $\theta_i > 0$, and that for $n > k$ the random variables X_n have the same exponential distribution with parameter θ_{k+1}. Does the sequence $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ satisfy a LDP?

1.7 Schilder's Theorem

In this section, we prove a LDP for the distribution of a one-dimensional Brownian motion. In the next section, this will be extended to any Gaussian process.

Let $\Omega = \mathcal{C}_0([0,T];\mathbb{R})$ be endowed with the topology of uniform convergence, P denote the Wiener measure on its Borel sets. Under P, the process $W : [0,T]\times\Omega\to\mathbb{R}$ defined by $W(t,\omega)=\omega_t$ is a standard Brownian motion, that is, $W_0=0$ and for $0=t_0<t_1<\dots<t_n\leq T$, the random variables $W_{t_k}-W_{t_{k-1}}$, $k=1,\dots,n$ are independent Gaussian random variables with mean 0 and variance t_k-t_{k-1}.

We have to describe the logarithmic moment generating function of P, and its Legendre transform.

The dual Ω^* is the set of Radon measures λ s.t. $\lambda(\{0\})=0$. For $f\in\Omega$, $\lambda\in\Omega^*$, $\langle\lambda,\omega\rangle=\int_0^T\omega_s\lambda(ds)$.

If $\lambda = a_1\delta_{t_1}+a_2\delta_{t_2}+\dots+a_n\delta_{t_n}$, $\langle\lambda,\omega\rangle=(a_1W_{t_1}+\dots+a_nW_{t_n})(\omega)$

$$=\Big(\sum_{i=1}^n a_i\Big)W_{t_1}+\Big(\sum_{i=2}^n a_i\Big)(W_{t_2}-W_{t_1})+\dots+\Big(\sum_{i=k}^n a_i\Big)(W_{t_k}-W_{t_{k-1}})+\dots+a_n(W_{t_n}-W_{t_{n-1}})$$

$$=\int_0^T\lambda([s,1))\,dW_s$$

This extends to any Borel measure λ: $\langle\lambda,\omega\rangle=\int_0^T\lambda((s,1])\,dW_s$

Hence $E[\langle\lambda,\omega\rangle^2]=E\Big(\Big|\int_0^T\lambda([s,1))\,dW_s\Big|^2\Big)=\int_0^T\lambda((s,1])^2\,ds.$

Set $H=\{f\in\Omega : f(t)=\int_0^t\dot f_s\,ds,\ \dot f\in L^2(0,T;\mathbb{R})\}$

and for $f\in H$, set $\|f\|_H=\|\dot f\|_{L^2(0,1)}$.

Then H is a Hilbert space, and $\begin{array}{c}\Omega^*\hookrightarrow H\\ \lambda\mapsto F_\lambda\end{array}$, where $\dot F_\lambda(s)=\lambda(s,T]$.

Lemma 1.19 The logarithmic moment generating function of P is
$$\Lambda(\lambda)=\ln E[e^{\langle\lambda,\omega\rangle}]=\frac{1}{2}\|F_\lambda\|_H^2.$$

Proof For $\lambda \in \Omega^*$, $\langle \lambda, \omega \rangle = \int \lambda(s,T]\, dW_s$ is a Gaussian random variable with mean 0 and variance $\frac{1}{2}\int_0^T \lambda(s,1]^2 ds = \frac{1}{2}\|F_\lambda\|_H^2$. $\square$

For $\varepsilon > 0$, let $\mu_\varepsilon = P(\sqrt{\varepsilon} W)^{-1}$ denote the distribution of $\sqrt{\varepsilon}W$. We prove that $(\mu_\varepsilon, \varepsilon > 0)$ satisfies the condition (C), where $\frac{1}{n}$ is replaced by $\varepsilon > 0$.

For $\lambda \in \Omega^*$, $\Lambda_\varepsilon\left(\frac{\lambda}{\varepsilon}\right) = \ln E\left[e^{\langle \frac{\lambda}{\varepsilon}, \sqrt{\varepsilon}W\rangle}\right] = \Lambda\left(\frac{\lambda}{\sqrt{\varepsilon}}\right)$. The expression of Λ obtained in Lemma 1.19 shows that $\Lambda(\alpha\lambda) = \alpha^2 \Lambda(\lambda)$, for $\alpha \in \mathbb{R}$.

Hence $\varepsilon \Lambda_\varepsilon\left(\frac{\lambda}{\varepsilon}\right) = \varepsilon \left(\frac{1}{\sqrt{\varepsilon}}\right)^2 \Lambda(\lambda) = \Lambda(\lambda)$.

Using Theorem 1.12, we deduce that upper bounds hold for compact subsets of $C_0([0,1];\mathbb{R})$ with rate function Λ^*. We next compute Λ^*. To get heuristically the expression let $0 < t_1 < t_2 < \dots < t_k \le T$ and $X = (W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_k} - W_{t_{k-1}})$. For $\lambda \in \mathbb{R}^k$,

$$\tilde{\Lambda}(\lambda) = \ln E\left[e^{\langle \lambda, X\rangle}\right] = \frac{\lambda_1^2}{2} t_1 + \lambda_2^2 \frac{(t_2 - t_1)}{2} + \dots + \lambda_k^2 \frac{(t_k - t_{k-1})}{2},$$

and for $x \in \mathbb{R}^k$, $\tilde{\Lambda}^*(x) = \sup\{\langle \lambda, x\rangle - \tilde{\Lambda}(\lambda) : \lambda \in \mathbb{R}^k\} = \frac{x_1^2}{2t_1} + \frac{x_2^2}{2(t_2 - t_1)} + \dots + \frac{x_k^2}{2(t_k - t_{k-1})}$

We reformulate $\tilde{\Lambda}^*$ in terms of a function f such that $x_j = f(t_j) - f(t_{j-1})$, $1 \le j \le k$.

$$\tilde{\Lambda}^*(x) = \frac{1}{2}\int_0^{t_1} \left(\frac{f(t_1)}{t_1}\right)^2 dx + \frac{1}{2}\int_{t_1}^{t_2} \left(\frac{f(t_2) - f(t_1)}{t_2 - t_1}\right)^2 dx + \dots + \frac{1}{2}\int_{t_{k-1}}^{t_k} \left(\frac{f(t_k) - f(t_{k-1})}{t_k - t_{k-1}}\right)^2 dx$$

Let $\tilde{f}$ denote the linear interpolation of f on the points $t_1, \dots, t_k$, letting $f(0) = 0$.

Then $\dot{\tilde{f}}(x) = \frac{f(t_i) - f(t_{i-1})}{t_i - t_{i-1}}$ on (t_{i-1}, t_i), and $\tilde{\Lambda}^*(x) = \frac{1}{2}\int_0^T |\dot{\tilde{f}}(x)|^2 dx = \frac{1}{2}\|\tilde{f}\|_H^2$

The following result shows the link between Λ^* and H.

Lemma 1.20 (i) $h \in H$ if and only if $\lim 2^n \sum_{i=1}^{2^n} |h(Ti2^{-n}) - h(T(i-1)2^{-n})|^2 < \infty$

If $h \in H$, this limit is equal to $\int_0^T |\dot{h}_s|^2 ds = \|h\|_H^2$

(ii) For $f \in \Omega$, $\Lambda^*(f) = \sup\{\langle \lambda, f\rangle - \Lambda(\lambda) : \lambda \in \Omega^*\} = \begin{cases} \frac{1}{2}\|f\|_H^2 & \text{if } f \in H \\ +\infty & \text{otherwise} \end{cases}$

<u>Proof</u> (i) Let $h \in H$; then $\dot h \in L^2(0,T)$ and for $\mathcal{F}_n = \sigma\left([\frac{(i-1)T}{2^n}, \frac{iT}{2^n}),\ i=1,..,n \right)$,

To ease notations, let $T=1$.

$$\dot h_n(t) = E_{\text{Lebe}}(\dot h \mid \mathcal{F}_n) = 2^n \sum_{i=1}^{2^n} [h(i2^{-n}) - h((i-1)2^{-n})]\, 1_{[\frac{i-1}{2^n}, \frac{i}{2^n})}(t)$$

1.30

The martingale $\dot h_n$ converges to $\dot h$ in $L^2(0,1;\text{Lebesgue})$ and $\|\dot h_n\|_2 \le \|\dot h\|_2$

Furthermore $\|\dot h_n\|_2^2 = \sum_{i=1}^{2^n} 2^{2n} |h(i2^{-n}) - h((i-1)2^{-n})|^2 2^{-n} = 2^n \sum_{i=1}^{n} |h(i2^{-n}) - h((i-1)2^{-n})|^2$

Conversely, if $\sup_n 2^n \sum_{i=1}^{2^n} |h(i2^{-n}) - h((i-1)2^{-n})|^2 < \infty$, the sequence $(\dot h_n, \mathcal{F}_n)$ is a martingale bounded in $L^2(\text{Lebesgue})$. Hence it is closed, converges in L^2 to some function φ and $\dot h_n = E_{\text{Lebe}}(\varphi \mid \mathcal{F}_n)$.

Since $h_n(t) = \int_0^t \dot h_n(s)ds$ is the linear interpolation of h on the dyadic grid $(i2^{-n}, i=0,..,2^n)$, h_n converges to h uniformly on $[0,1]$. Hence $h_n(t) = \int_0^t \dot h_n(s)ds \to \int_0^t \varphi(s)ds$, that is $h(t) = \int_0^t \varphi(s)ds$ and $\varphi \in L^2(0,1)$. Thus $h \in H$, $\|h\|_H^2 = \|\varphi\|_2 = \lim_m \|\dot h_m\|_{L^2}^2$

(ii) Let $f \in H$, $\lambda \in \Omega^*$; then integration by parts implies

$$\langle \lambda, f\rangle = \int_0^1 f(t)\,\lambda(dt) = -\int_0^1 f(t)\, d(\lambda(t,1]) = \int_0^1 \lambda(t,1]\, \dot f(t)dt$$

Hence for $f \in H$,

$$\Lambda^*(f) = \sup\{\langle \lambda, f\rangle - \Lambda(\lambda) : \lambda \in \Omega^*\} = \sup\left\{ \int_0^1 \dot f(t)\lambda(t,1]dt - \frac12 \int_0^1 |\lambda(t,1]|^2 dt : \lambda \in \Omega^* \right\}$$

$$= \sup\left\{ -\frac12 \int_0^1 |\dot f(t) - \lambda(t,1]|^2 dt + \frac12 \int_0^1 |\dot f(s)|^2 dt : \lambda \in \Omega^* \right\}$$

$$= \frac12 \int_0^1 |\dot f(s)|^2 dt = \frac12 \|f\|_H^2.$$

If $f \notin H$, let $\lambda = 2^n \sum_{i=1}^{2^n} [f(i2^{-n}) - f((i-1)2^{-n})] (\delta_{i2^{-n}} - \delta_{(i-1)2^{-n}}) \in \Omega^*$. Then

$\Lambda(\lambda) = \frac12 2^n \sum_{i=1}^{2^n} |f(i2^{-n}) - f((i-1)2^{-n})|^2$, and

$$\Lambda^*(f) \ge \langle \lambda, f\rangle - \Lambda(\lambda) = \frac12 \sum_{i=1}^{2^n} |f(i2^{-n}) - f((i-1)2^{-n})|^2 \longrightarrow \infty \text{ as } n \to \infty$$

Hence $\Lambda^*(f) = +\infty$, which completes the proof. $\square$

Remark 1.21 The function Λ^* is a good rate function. Indeed, let $a > 0$. Then $\{f \in \Omega : \Lambda^*(f) \leq a\} = \{f \in H : \int_0^1 |\dot{f}(t)|^2 dt \leq 2a\}$ is a bounded, equicontinuous subset of Ω (since $|f(t_2) - f(t_1)| \leq |t_2 - t_1|^{\frac{1}{2}} \left(\int_{t_1}^{t_2} |\dot{f}(s)|^2 ds\right)^{\frac{1}{2}} \leq \sqrt{2a}\,(t-s)^{\frac{1}{2}}$.

Ascoli's theorem implies that is is a relatively compact subset of Ω (for the uniform topology). Since $\{\Lambda^* \leq a\}$ is closed, it is a compact subset of Ω.

We next prove that the sequence (P^ε) of distributions of $(\sqrt{\varepsilon} W)$ satisfies a LDP.

Theorem 1.22 (Schilder's theorem) The family (P^ε) of distributions of $\sqrt{\varepsilon} W$ satisfies on $C_0([0,1]; \mathbb{R})$ a LDP with the good rate function Λ^*.

Proof We already have shown that (P_ε) satisfies the condition (C) of the Ellis-Gärtner theorem, and hence that upper bounds are true for compact sets. Hence, using Proposition 1.11, we only have to prove that (P^ε) is exponentially tight on $C_0([0,1]; \mathbb{R})$, and that (P^ε) satisfies lower bounds for open sets.

Lower bounds We use a direct argument based on a change of probability.

Let G be a non-empty open set of Ω and $f \in G$. We prove that

$\varliminf_{\varepsilon} \varepsilon \ln P(\sqrt{\varepsilon} W \in G) \geq -\Lambda^*(f)$. This will imply the lower bound.

If $\Lambda^*(f) = +\infty$, this lower estimate is trivial.

Let $f \in H$, $\delta > 0$ be such that $B(f, \delta) \subset G$. Let $\tilde{W}_t = W_t - \frac{f_t}{\sqrt{\varepsilon}}$.

Then Cameron Martin's theorem implies that $\tilde{W}$ is a $\tilde{P}$ Brownian motion, with

$$\frac{d\tilde{P}}{dP} = \exp\left(\int_0^T \frac{\dot{f}_s}{\sqrt{\varepsilon}} dW_s - \frac{1}{2\varepsilon} \|f\|_H^2\right), \text{ that is } \frac{dP}{d\tilde{P}} = \exp\left(-\int_0^T \frac{\dot{f}_s}{\sqrt{\varepsilon}} d\tilde{W}_s - \frac{1}{2}\|f\|_H^2\right)$$

Hence

$$P(\|\sqrt{\varepsilon}W-f\|_\infty<\delta)=P(\|\tilde W\|_\infty<\tfrac{\delta}{\sqrt{\varepsilon}})$$
$$=\int_{\{\|\tilde W\|_\infty<\frac{\delta}{\sqrt{\varepsilon}}\}}\exp\Big(-\int_0^1\frac{\dot f_s}{\sqrt{\varepsilon}}\,d\tilde W_s-\frac{1}{2\varepsilon}\|f\|_H^2\Big)\,d\tilde P$$

Since W (resp $\tilde W$) is a P (resp. $\tilde P$) Brownian motion, and since $(-W)$ is a P Brownian motion, using the inequality $\frac{1}{2}(e^x+e^{-x})\geq 1$, we deduce

$$P(\|\sqrt{\varepsilon}W-f\|_\infty<\delta)=E\Big(1_{\{\|W\|_\infty<\delta\}}\tfrac{1}{2}\Big[\exp\Big(-\tfrac{1}{\sqrt{\varepsilon}}\int_0^T\dot f_s\,dW_s\Big)+\exp\Big(+\tfrac{1}{\sqrt{\varepsilon}}\int_0^T\dot f_s\,dW_s\Big)\Big]\Big)\exp\Big(-\frac{\|f\|_H^2}{2\varepsilon}\Big)$$
$$\geq P(\|W\|_\infty<\tfrac{\delta}{\sqrt{\varepsilon}})\exp\Big(-\frac{1}{2\varepsilon}\|f\|_H^2\Big).$$

Therefore

$$\underline{\lim}\ \varepsilon\ln P(\sqrt{\varepsilon}W\in G)\geq-\frac{1}{2}\|f\|_H^2+\underline{\lim}\ \varepsilon\ln P(\|W\|_\infty<\tfrac{\delta}{\sqrt{\varepsilon}})$$

For fixed $\delta>0$, as $\varepsilon\to 0$, $P(\|W\|_\infty<\frac{\delta}{\sqrt{\varepsilon}})\to 1$ and $\varepsilon\ln P(\|W\|_\infty<\frac{\delta}{\sqrt{\varepsilon}})\to 0$.

Hence $\underline{\lim}\ \varepsilon\ln P(\sqrt{\varepsilon}W\in G)\geq-\frac{1}{2}\|f\|_H^2$.

This proves the lower estimate for open sets

__Exponential tightness__

To give an intuition, let us prove that if $X\sim N(0,1)$, the family of distributions of $(\sqrt{\varepsilon}X)$ is exponentially tight. For any $M>0$, $c>0$

$$P(|\sqrt{\varepsilon}X|>M)=P(|X|>\tfrac{M}{\sqrt{\varepsilon}})=P(e^{cX^2}>e^{c\frac{M^2}{\varepsilon}})\leq e^{-\frac{cM^2}{\varepsilon}}E[e^{cX^2}]$$

If $c<\frac{1}{2}$, $\int e^{cx^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=E(e^{cX^2})<\infty$

This proves that for $c<\frac{1}{2}$, $\overline{\lim}\ \varepsilon\ln P(|\sqrt{\varepsilon}X|>M)\leq -cM^2$, and hence that $(\sqrt{\varepsilon}X)$ is exponentially tight.

We transpose this idea to $\sqrt{\varepsilon}W$ on $\Omega=C_0([0,1],\mathbb{R})$

The proof relies on the following

Theorem 1.23 (Fernique) Let $\mathcal{X}$ be a separable Banach space, $\Phi : \mathcal{X} \longrightarrow [0,+\infty]$ be measurable such that $\Phi(x+y) \leq \Phi(x) + \Phi(y)$, $\Phi(\alpha x) = |\alpha| \Phi(x)$ for $x, y \in \mathcal{X}$, $\alpha \in \mathbb{R}$

Let μ be a probability on $\mathcal{B}(\mathcal{X})$ such that $\mu \otimes \mu$ is invariant by the map

$$\mathcal{X}^2 \longrightarrow \mathcal{X}^2$$
$$(x_1, x_2) \longmapsto \left(\frac{1}{\sqrt{2}}(x_1+x_2), \frac{1}{\sqrt{2}}(x_1-x_2)\right)$$

If $\mu(\{x \in \mathcal{X} : \Phi(x) < \infty\}) = 1$, there exists $c > 0$ such that

$$\int_{\mathcal{X}} e^{c\Phi^2(x)} \mu(dx) < \infty$$

Proof of Fernique's theorem. Fix s, t such that $0 < s < t$. Then

$$\mu(\{x : \Phi(x) \leq s\}) \mu(\{x : \Phi(x) \geq t\}) = \mu \otimes \mu(\{(x,y) : \Phi(x) \leq s, \Phi(y) \geq t\})$$
$$= \mu \otimes \mu(\{(x,y) : \Phi(x-y) \leq \sqrt{2}s, \Phi(x+y) \geq \sqrt{2}t\})$$
$$\leq \mu \otimes \mu(\{(x,y) : |\Phi(x) - \Phi(y)| \leq \sqrt{2}s, \Phi(x) + \Phi(y) \geq \sqrt{2}t\})$$
$$\leq \mu \otimes \mu\left(\left\{(x,y) : \Phi(x) \wedge \Phi(y) \geq \frac{t-s}{\sqrt{2}}\right\}\right) = \mu\left(\left\{x : \Phi(x) \geq \frac{t-s}{\sqrt{2}}\right\}\right)^2$$

Define by induction the sequence t_n as follows: $t_0 = s$, $t_n = s + \sqrt{2}\, t_{n-1}$, $n \geq 1$

Then $t_n = s \sum_{i=0}^{n} (\sqrt{2})^i \leq s \dfrac{(\sqrt{2})^n}{\sqrt{2}-1}$, and $t_{n-1} = \dfrac{t_n - s}{\sqrt{2}}$

We deduce by induction that

$$\frac{\mu(\{x : \Phi(x) \geq t_n\})}{\mu(\{x : \Phi(x) \leq s\})} \leq \left[\frac{\mu(\Phi(x) \geq t_{n-1})}{\mu(\Phi(x) \leq s)}\right]^2 \leq \left[\frac{\mu(\Phi(x) \geq s)}{\mu(\Phi(x) \leq s)}\right]^{2^n}$$
$$\leq \exp(-2^n \ln \sigma),$$

Where $\sigma = -\ln\left(\dfrac{\mu(\Phi(x) \geq s)}{\mu(\Phi(x) \leq s)}\right) > 0$, if we choose s such that $\mu(\Phi \leq s) > \frac{1}{2}$

Then if $b = 2\left(\dfrac{s}{\sqrt{2}-1}\right)^2$, we have $t_n^2 \leq 2^n b$. For any $a > 0$

$$\int_{\{\Phi^2(x) \geq b\}} \exp(a\Phi^2(x)) \mu(dx) = \sum_{n \geq 0} \int_{\{b2^n \leq \Phi(x) < b2^{n+1}\}} e^{a\Phi^2(x)} \mu(dx) \leq \sum_{n \geq 0} \exp(ab2^{n+1}) \mu(\{x : \Phi(x) \geq b2^n\})$$
$$\leq \sum_{n \geq 0} \exp(ab2^{n+1} - \sigma 2^n) = \sum_{n \geq 0} \exp((2ab - \sigma)2^n)$$

This series converges if $2ab < \sigma$. $\square$

Let $\|\cdot\|_\alpha$ denote the α-Hölder norm, defined by $\|f\|_\alpha = \sup_{s<t} \frac{|f(t)-f(s)|}{|t-s|^\alpha}$

The trajectories of (W_t) are a.s. α Hölder continuous if $\alpha < \frac{1}{2}$.

Fernique's theorem implies that for $\alpha \in [0, \frac{1}{2})$, there exists $c > 0$ such that

$$E[e^{c\|W\|_\alpha^2}] < \infty$$

Indeed, let $\Phi(\omega) = \|\omega\|_\alpha$. Then if W^1 and W^2 are independent Brownian motions, $\frac{1}{\sqrt{2}}(W^1+W^2)$ and $\frac{1}{\sqrt{2}}(W^1-W^2)$ are again independent Brownian motions

Indeed, these are Gaussian processes with independent increments, and $\langle \frac{1}{\sqrt{2}}(W^1+W^2), \frac{1}{\sqrt{2}}(W^1-W^2)\rangle_t = 0$; the Ray Knight theorem proves that $\Phi(\omega) = \|\omega\|_\alpha$ satisfies the requirements of the Fernique Theorem for $\alpha < \frac{1}{2}$.

Finally, let $K_M = \{\omega \in \Omega : \|\omega\|_\alpha \leq M\}$ for $\alpha < \frac{1}{2}$

Using Ascoli's theorem, we deduce that K_M is bounded, equicontinuous, and hence relatively compact in Ω. Furthermore

$$P(\sqrt{\varepsilon} W \notin \overline{K_M}) \leq P(\|W\|_\alpha > \frac{M}{\sqrt{\varepsilon}}) = P(e^{c\|W\|_\alpha^2} > e^{\frac{cM^2}{\varepsilon}})$$

$$\leq \exp(-\frac{cM^2}{\varepsilon}) E[e^{c\|W\|_\alpha^2}]$$

For c small enough we deduce that $\overline{\lim}\, \varepsilon \ln P(\sqrt{\varepsilon} W \notin \overline{K_M}) \leq -cM^2$; this completes the proof of exponential tightness. $\square$

General Gaussian processes

Recall that a vector $X : \Omega \to \mathbb{R}^d$ is Gaussian if for every $a \in \mathbb{R}^d$, $\langle a, X\rangle = \sum_{k=1}^d a_k X_k$ is a real-valued Gaussian random variable. Its distribution is characterized by vector $m = E(X) = (E(X_1), \dots, E(X_d))$ such that $E(\langle a, X\rangle) = \langle a, E(X)\rangle$ for any $a \in \mathbb{R}^d$ and its covariance matrix $\Gamma_X(j,k) = \mathrm{Cov}(X_j, X_k) = E(X_j X_k) - E(X_j)E(X_k)$. Furthermore given $a \in \mathbb{R}^d$, $\mathrm{Var}(\langle a, X\rangle) = \tilde{a}\,\Gamma_X a$, where $a = \begin{pmatrix} a_1 \\ \vdots \\ a_d \end{pmatrix}$ and $\tilde{a} = (a_1, \dots, a_d)$ is the transposed vector of a. Furthermore, given $\lambda \in \mathbb{R}^d$,

$$\Lambda(\lambda) = \ln E\left(e^{\lambda X}\right) = \langle \lambda, m\rangle + \frac{1}{2}\tilde{\lambda}\,\Gamma_X \lambda.$$

The computation of $\Lambda^*(x)$, $x \in \mathbb{R}^d$, is left as an exercise.

Recall that the components X_k, $k = 1, \dots, d$ are independent if and only if Γ_X is diagonal.

We next extend this notion replacing $\mathbb{R}^d$ by a Hilbert space H

Definition 1.24 Let H be a Hilbert space. A measure μ on $\mathcal{B}(H)$ is Gaussian, if for every $h \in H$, the pushforward measure of μ by the map Π_h defined by $\Pi_h : \begin{array}{l} H \longrightarrow \mathbb{R} \\ x \longmapsto \langle h, x\rangle \end{array}$ is a Gaussian measure on $\mathcal{B}(\mathbb{R})$.

We next define the mean and covariance operator.

The maps $h \in H \mapsto \int_H \langle h, x\rangle \mu(dx) = \int_{\mathbb{R}} y \; \mu\circ\Pi_h^{-1}(dy)$ and

$(h_1, h_2) \in H \times H \mapsto \int_H \langle h_1, x\rangle\langle h_2, x\rangle \mu(dx) \leq \left(\int_{\mathbb{R}} y^2 d\mu\circ\Pi_{h_1}^{-1}(y)\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} y^2 d\mu\circ\Pi_{h_2}^{-1}(y)\right)^{\frac{1}{2}}$

are well defined. Furthermore they are continuous. Hence there exists

* $m \in H$ such that for any $h \in H$, $\langle m, h\rangle = \int_H \langle x, h\rangle \mu(dx)$
* $Q \in \mathcal{L}(H, H)$ such that for any $h_1, h_2 \in H$,

$$\langle Q h_1, h_2\rangle = \int_H \langle x - m, h_1\rangle\langle x - m, h_2\rangle \mu(dx)$$

The element m is called the mean and the operator Q the covariance of μ.

The operator Q is symmetric, non negative (that is $\langle Qh, h\rangle \geq 0$, $\forall h \in H$)

$\boxed{\text{Lemma 1.25}}$ If μ is a Gaussian measure on $\mathcal{B}(H)$ with covariance operator Q, then Q is trace class ($Q \in \mathcal{L}_1^+$), that is for every ONB $\{e_n, n\geq 0\}$ of H, $Tr(Q) = \sum_{n\geq 0} \langle Qe_n, e_n\rangle < \infty$

Furthermore we have the following analog of the existence of exponential moments for $\|x\|_H^2$; which is a consequence of Fernique's theorem

Let μ be a Gaussian measure on $\mathcal{B}(H)$ with covariance operator Q. Then

for $s < \frac{1}{2 Tr Q}$, $\int_H e^{s\|x\|_H^2} \mu(dx) \leq \frac{1}{\sqrt{1-2sT_r(Q)}} < \infty$

As in the case of finite-dimensional Gaussian vectors, we have the following

$\boxed{\text{Proposition 1.26}}$ Given $m \in H$ and a symmetric, non negative trace-class operator Q, there exists a Gaussian measure μ on $\mathcal{B}(H)$ with mean m and covariance operator Q.

Definition 1.27 A H-valued stochastic process $(W_t, t\geq 0)$ is a Q Wiener process if $W_0 \equiv 0$,

$W_\cdot(\omega) : [0, \infty) \longrightarrow H$ is continuous

W_t has independent increments

$W_t - W_s \sim N(0, (t-s)Q)$ for $0 \leq s < t$

Let $(e_k)_{k\geq 0}$ be an ONB of H made of eigen vectors of Q : $Qe_k = \lambda_k e_k$ for $\lambda \geq 0$. Then

$W_t = \sum_{k\geq 0} \sqrt{\lambda_k}\, e_k\, \beta_k(t)$, where $\beta_k(t) = \langle W_t, Q^{-\frac{1}{2}} e_k\rangle$ are independent 1-d Brownian motions

The set $H_0 = Q^{\frac{1}{2}} H$ is the reproducing kernel Hilbert space (RKHS) of (W_t)

Note that the operator $Q^{\frac{1}{2}}$ is Hilbert-Schmidt

H_0 is a Hilbert space endowed with the scalar product $\langle u, v\rangle_{H_0} = \langle Q^{-\frac{1}{2}} u, Q^{-\frac{1}{2}} v\rangle_H$

The following result extends Schilder's theorem to any H-valued Q Brownian motion (W_t) on $[0,T]$

Theorem 1.28 (Schilder's theorem for Hilbert-valued Brownian motion)

Let H be a Hilbert space, Q be a trace class operator on H and $H_0 = Q^{\frac{1}{2}} H$ be the RKHS of an H-valued, Q Brownian motion $(W_t, t\geq 0)$.

Then the family $\mathcal{J}_\varepsilon$ of distributions of $\sqrt{\varepsilon} W : \Omega \times [0,T] \longrightarrow H$ satisfies on $C([0,T]; H)$ a LDP with the good rate function

$$\Lambda^*(f) = \begin{cases} \inf\left\{ \frac{1}{2}\int_0^T \|\dot{f}(s)\|^2_{H_0}\, ds \; ; \; f(\cdot) = \int_0^{\cdot} \|\dot{f}(s)\|^2_{H_0}\, ds, \; \dot{f} \in L^2(0,T; H_0)\right\} \\ +\infty \quad \text{otherwise} \end{cases}$$

We finally state a version of Schilder's theorem for the Brownian sheet, which is the driving noise of some SPDEs.

Theorem 1.29 (Schilder's theorem for the Brownian sheet)

Let $\{Z(t,x),\ t\in[0,T],\ x\in[0,L]\}$ be the Brownian sheet, that is a Gaussian process on $[0,T]\times[0,L]$ with zero mean, and covariance

$$E\left(Z(s,x)\, Z(t,y)\right) = (s\wedge t)\,(x\wedge y).$$

Then the family of distributions of $\sqrt{\varepsilon}\, Z$ satisfy on $\mathcal{C}([0,T]\times[0,L]; \mathbb{R})$ a LDP with good rate function

$$I(f) = \begin{cases} \frac{1}{2}\int_{[0,T]\times[0,L]} \dot{f}(s,y)^2\, ds\, dy, \; f(t,x) = \int_0^t\int_0^x \dot{f}(s,y)\, ds\, dy \; ; \; \dot{f} \in L^2 \\ +\infty \quad \text{otherwise} \end{cases}$$

Schilder's theorem can be extended to arbitrary Gaussian measures on Banach spaces. Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a separable Banach space and μ be a probability on $\mathcal{B}(\mathcal{X})$ such that $\int_{\mathcal{X}} e^{i\langle\lambda,x\rangle}\mu(dx) = \exp(-\frac{1}{2}Q_\mu(\lambda,\lambda))$ for $\lambda\in\mathcal{X}^*$, where $Q_\mu : \mathcal{X}^*\times\mathcal{X}^*\to[0,+\infty)$ is bilinear; Q_μ is the covariance operator of μ.

An example of such a measure is the Wiener measure W on $\Theta = \{f\in\mathcal{C}([0,\infty);\mathbb{R}) : f(0)=0, \lim_{t\to\infty}\frac{|f(t)|}{t}=0\}$. Then $Q_W(\lambda,\lambda) = \int_0^\infty \lambda(t,\infty)^2\,dt$

In general, there exists $\alpha>0$ such that $\int_{\mathcal{X}}\exp(\alpha\|x\|^2_{\mathcal{X}})\mu(dx)<\infty$, and for $\lambda\in\mathcal{X}^*$, $\Lambda(\lambda)=\frac{1}{2}Q_\mu(\lambda,\lambda)=\frac{1}{2}\int_{\mathcal{X}}\langle\lambda,x\rangle^2\mu(dx)\le B\|\lambda\|^2_{\mathcal{X}^*}$, where $B=\int_{\mathcal{X}}\|x\|^2_{\mathcal{X}}\mu(dx)\in(0,\infty)$.

Let $\Lambda^*_\mu(x) = \sup\{\langle\lambda,x\rangle - \Lambda_\mu(\lambda) : \lambda\in\mathcal{X}^*\}$. Then Λ^*_μ is a good rate function such that $\Lambda^*_\mu(ax) = a^2\Lambda^*_\mu(x)$ for $x\in\mathcal{X}$ and $a\in\mathbb{R}$.

For $\varepsilon>0$, let μ_ε be the pushforward of μ by the map $x\in\mathcal{X}\mapsto\sqrt{\varepsilon}x\in\mathcal{X}$.

Theorem 1.30 The family μ_ε satisfies a LDP with the good rate function Λ^*

Furthermore, if μ is a centered Gaussian measure on a separable Banach space $\mathcal{X}$, there exists a separable Hilbert space H and a linear continuous map $S : H\to\mathcal{X}$, which is compact. Then

$$\Lambda^*_\mu(x) = \begin{cases} \frac{1}{2}\|S^{-1}x\|_H & \text{if } x\in S(H) \\ +\infty & \text{otherwise}\end{cases}$$

2 Contraction principle.- Freidlin Wentzell approach

The aim of this section is to provide tools which enable to deduce LDP from known ones.

2.1 Contraction principle, Transfer principle

We at first transfer the LDP using a continuous mapping.

Theorem 2.1 (Contraction principle) Let X, Y be Hausdorff topological spaces, $f : X \longrightarrow Y$ be a continuous map, $I : X \longrightarrow [0,+\infty]$ be a "good" rate function (as defined in Section 1.4). Then

(i) For $y \in Y$, let $I'(y) = \inf\{I(x) : x \in X,\ y = f(x)\}$ (with the convention $\inf \phi = +\infty$). Then I' is a good rate function on Y.

(ii) Let (μ_ε) be a family of probabilities on $\mathcal{B}(X)$ which satisfies a LDP with the good rate function I. Then the family of pushforward probabilities $(\mu_\varepsilon \circ f^{-1})$ satisfies a LDP with the good rate function I'.

Proof (i) The map I' is clearly defined from Y to $[0,+\infty]$. Hence we only have to check that for any $a \in [0,+\infty)$, $\{I' \le a\}$ is a compact subset of Y.

Let $y \in Y$ be such that $I'(y) < \infty$, and let (X_k) be a sequence of elements of X such that $y = f(X_k)$ and $I(X_k) \le I'(y) + \frac{1}{k} \le I'(y)+1$ for all $k \ge 1$.

Since $\{I \le I'(y)+1\}$ is a compact subset of X, there exists a subsequence $X_{k(n)}$ such that as $n \to \infty$, $X_{k(n)} \to x$. Since f is continuous, $f(X_{k(n)}) = y$ converges to $f(x) = y$, and since I is l.s.c., $I(x) \le \underline{\lim}\ I(X_{k(n)}) = I'(y)$.

Hence for $a \in [0,\infty)$, $\{I' \le a\} = f(\{I \le a\})$ is a compact subset of Y.

(iv) Let $A \in \mathcal{B}(Y)$; then

$$\inf\{I'(y) : y \in A\} = \inf\{I(x) : x \in f^{-1}(A)\}$$

Since f is continuous, if A is closed (resp. open), then $f^{-1}(A)$ is closed (resp. open). If G is an open set of Y,

$$\underline{\lim}\, \varepsilon \ln (\mu^\varepsilon \circ f^{-1})(G) = \underline{\lim}\, \varepsilon \ln \mu_\varepsilon (f^{-1}(G)) \geq -\inf\{I(x) : x \in f^{-1}(G)\}$$
$$= -\inf\{I'(y) : y \in G\}$$

and a similar argument proves upper estimates for closed sets. □

One would like to use the contraction principle to deduce LDP for diffusion processes $X^\varepsilon_t = x + \int_0^t \sigma(X^\varepsilon_s)\sqrt{\varepsilon}\, dW_s + \int_0^t b(X^\varepsilon_s)\, ds$ given (W_t) a r-dimensional Brownian motion, $x \in \mathbb{R}^d$, $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^r$ Lipschitz coefficients. The k-dimensional Brownian motion $(\sqrt{\varepsilon} W)$ satisfies a LDP on $C_0([0,T]; \mathbb{R}^k)$ with the good rate function $I(f) = \frac{1}{2}\int_0^T \|\dot{f}_s\|^2_{\mathbb{R}^r}\, ds$ if $f(t) = \int_0^t \dot{f}_s\, ds$, $\dot{f} \in L^2(0,T; \mathbb{R}^k)$, $I(f) = +\infty$ otherwise. However in general, the map $\sqrt{\varepsilon} W \mapsto X^\varepsilon$ is NOT continuous from $C_0([0,T]; \mathbb{R}^k)$ to $C_x([0,T]; \mathbb{R}^d)$. The lack of continuity comes from the Lévy area $\int_0^t W^1_s\, dW^2_s$, where W^1 and W^2 are independent Brownian motions. One way to overcome this problem is to use rough paths, a theory initiated by T. Lyons. It requires "strong" assumptions on the diffusion coefficient σ.

The continuity of the above map is true if the noise is additive that is $k=d$ and $\sigma = \mathrm{Id}$. Then if $Y^\varepsilon_t = X^\varepsilon_t - \sqrt{\varepsilon} W_t$, $Y^\varepsilon_t = x + \int_0^t b(Y^\varepsilon_s + \sqrt{\varepsilon} W_s)\, ds$. If f is continuous, $b(s,x) = b(f_s + x)$ is uniformly continuous in x for fixed t.

Hence $\dot{F}_s = b(F_s + f_s)$, $F_0 = x$, has a unique solution, and

$$F : C_0([0,T];\mathbb{R}^d) \longrightarrow C_x([0,T];\mathbb{R}^d) \quad \text{is continuous}$$
$$f \longmapsto F(f)_t = x + \int_0^t b\,(F(f)_s + f_s)\,ds$$

Indeed $|F(f_1)_t - F(f_2)_t| \leq \int_0^t c\,|F(f_1)_s - F(f_2)_s|\,ds + c\int_0^t |f_2(s) - f_1(s)|\,ds$

and Gronwall's lemma implies $\sup_{s\leq t} |F(f_1)_s - F(f_2)_s| \leq c e^c \|f_1 - f_2\|_\infty$

Therefore, (Y^ε) satisfies on $C_x([0,T];\mathbb{R}^d)$ a LDP with the good rate function

$$\tilde{I}(f) = \begin{cases} \frac{1}{2}\int_0^T |\dot{f}_s - b(f(s))|^2\,ds & \text{if } f \in x + \mathcal{H} \\ +\infty & \text{otherwise} \end{cases}$$

The proof is left as an exercise.

The aim of the following result is to weaken the continuity assumption.

Theorem 2.2 (transfer principle) Let (E_1, d_1) and (E_2, d_2) be metric spaces.

Let $X_1^\varepsilon : \Omega \to E_1$ and $X_2^\varepsilon : \Omega \to E_2$ be families of random variables indexed by $\varepsilon > 0$. Suppose that

(1) (X_1^ε) satisfies a LDP on E_1 with the good rate function $I_1 : E_1 \to [0,+\infty]$

(2) There exist a map $S : \{I_1 < \infty\} \longrightarrow E_2$ such that

(i) For any $a \in [0,+\infty)$, $S|_{\{I_1 \leq a\}}$ is continuous

(ii) $\forall h \in E_1$, $I_1(h) < \infty$, $\forall \eta > 0$, $\forall R > 0$, $\exists \delta > 0$

$$\overline{\lim}\, \varepsilon \ln P\big(d_2(X_2^\varepsilon, S(h)) \geq \eta,\ d_1(X_1^\varepsilon, h) \leq \delta\big) \leq -R$$

Then the function $I_2 : E_2 \to [0,+\infty)$ defined by

$$I_2(f) = \inf\{ I_1(h) : S(h) = f\} \quad (\text{with the convention } \inf \emptyset = +\infty)$$

is a good rate function, and (X_2^ε) satisfies a LDP with rate function I_2.

Proof (1) Let $a \in [0,+\infty)$ and $f \in E_2$ be such that $I_2(f) \le a < \infty$. Let (h_n) be a sequence of elements of E_1 such that $S(h_n) = f$ and $I_2(f) \le I_1(h_n) \le I_2(f) + \frac{1}{n}$. The sequence (h_n) belongs to $\{I_1 \le a+1\}$ which is compact. Hence there exists a subsequence (h_{n_k}) such that $h_{n_k} \to h \in \{I_1 \le a+1\}$. Since $S|_{\{I_1 \le a+1\}}$ is continuous, $f = S(h_{n_k}) \xrightarrow[k]{} S(h) = f$. Since I_1 is l.s.c.,

$I_1(h) \le \underline{\lim}_k I_1(h_{n_k}) = I_2(f)$. Hence $\{I_2 \le a\} \subset S(\{I_1 \le a\})$. The converse inclusion is obvious. Since E_2 is a Hausdorff space, $\{I_2 \le a\} = S(\{I_1 \le a\})$ is compact.

(2) Let us check the lower estimates for X_2^ε. Let G be an open set of E_2, $f \in G$ be such that $a = I_2(f) < \infty$, and let $\eta > 0$ be such that $B_{d_2}(f, \eta) \subset G$. We will prove that $\underline{\lim}\, \varepsilon \ln P(d_2(X_2^\varepsilon, f) < \eta) \ge -I_2(f)$.

From the proof of (1), there exist $h \in E_1$ such that $f = S(h)$ and $I_2(f) = I_1(h)$. Fix $R > a$ and apply the condition (2-ii). There exists $\delta > 0$ such that

$$\overline{\lim}\, \varepsilon \ln P(d_2(X_2^\varepsilon, f) \ge \eta,\ d_1(X_1^\varepsilon, h) < \delta) \le -R.$$

$$\text{Then } P(d_1(X_1^\varepsilon, h) < \delta) = P(d_2(X_2^\varepsilon, f) < \eta,\ d_1(X_1^\varepsilon, h) < \delta) + P(d_2(X_2^\varepsilon, f) \ge \eta,\ d_1(X_1^\varepsilon, h) < \delta).$$

We next prove the following

Lemma 2.3 Let (a_ε) and (b_ε) be families of strictly positive real numbers. Then $\underline{\lim}\, \varepsilon \ln(a_\varepsilon + b_\varepsilon) \le \max(\underline{\lim}\, \varepsilon \ln a_\varepsilon,\ \overline{\lim}\, \varepsilon \ln b_\varepsilon)$

Proof of Lemma 2.3 **Case 1** $\overline{\lim}\, \varepsilon \ln(b_\varepsilon) \le \gamma \le \underline{\lim}\, \varepsilon \ln(a_\varepsilon)$

For any $\alpha > 0$, one can find $\varepsilon_0 > 0$ such that for $\varepsilon \in (0, \varepsilon_0]$,

$\varepsilon \ln(b_\varepsilon) \le \gamma+\alpha$, $\varepsilon \ln(a_\varepsilon) \ge \gamma-\alpha$, that is $b_\varepsilon \le \exp\left(\frac{\gamma+\alpha}{\varepsilon}\right)$, $a_\varepsilon \ge \exp\left(\frac{\gamma-\alpha}{\varepsilon}\right)$

Then $a_\varepsilon + b_\varepsilon \le a_\varepsilon + a_\varepsilon e^{\frac{2\alpha}{\varepsilon}} \le 2a_\varepsilon e^{\frac{2\alpha}{\varepsilon}}$

and $\underline{\lim}\, \varepsilon \ln(a_\varepsilon + b_\varepsilon) \le \underline{\lim}\, \varepsilon \ln(a_\varepsilon) + 2\alpha$

The proof is complete as $\alpha \to 0$

<u>Case 2</u> $\underline{\lim}\, \varepsilon \ln(a_\varepsilon) \le \gamma = \overline{\lim}\, \varepsilon \ln(b_\varepsilon)$

Let $\varepsilon_n \to 0$ be such that $\lim_n \varepsilon_n \ln a_{\varepsilon_n} = \underline{\lim}\, \varepsilon \ln(a_\varepsilon)$, then $\overline{\lim}\, \varepsilon_n \ln(b_{\varepsilon_n}) \le \gamma$. Therefore given any $\alpha > 0$, for n large enough, $\varepsilon_n \ln(b_{\varepsilon_n}) \le \gamma + \alpha$ and $\varepsilon_n \ln(a_{\varepsilon_n}) \le \gamma + \alpha$, which imply

$$\varepsilon_n \ln(a_{\varepsilon_n} + b_{\varepsilon_n}) \le \varepsilon_n \ln\left(2e^{\frac{\gamma+\alpha}{\varepsilon_n}}\right) \le \varepsilon_n \ln 2 + \gamma + \alpha$$

Thus, $\underline{\lim}\, \varepsilon \ln(a_\varepsilon + b_\varepsilon) \le \underline{\lim}\, \varepsilon_n \ln(a_{\varepsilon_n} + b_{\varepsilon_n}) \le \gamma + \alpha$; this concludes the proof as $\alpha \to 0$ $\square$

Using Lemma 2.3, we deduce for $a_\varepsilon = P(d_2(X_2^\varepsilon, f) < \eta)$ and $b_\varepsilon = P(d_1(X_1^\varepsilon, h) < \delta,\ d_2(X_2^\varepsilon, f) \ge \eta)$,

$$-a = -I_2(f) \le \max\left(\underline{\lim}\, \varepsilon \ln P(d_2(X_2^\varepsilon, f) < \eta),\ -R\right)$$

Since $R > a$, $-R < -a$ and the maximum in the RHS is $-a$

$\underline{\lim}\, \varepsilon \ln P(d_2(X_2^\varepsilon, f) < \eta) \le -a$. Hence $\underline{\lim}\, \varepsilon \ln P(d_2(X_2^\varepsilon, f) < \eta) \ge -I_2(f)$.

We next prove <u>upper estimate for closed sets</u>. Let F be closed. If $\inf\{I_2(f) : f \in F\} = 0$, the upper estimate is trivial. Hence, we assume that $\inf\{I_2(f) : f \in F\} > a > 0$ and prove that $\overline{\lim}\, \varepsilon \ln P(X_2^\varepsilon \in F) \le -a$. Letting a decrease to $\inf\{I_2(f) : f \in F\}$ will conclude the proof.

Let $K_a = \{I_2 \leq a\} = S(\{I_1 \leq a\})$ and $R > a$. Let $h \in E_1$ be such that $f = S(h) \notin F$ and $I_1(h) \leq a$. Let $\rho_h > 0$ be such that $B_{d_2}(S(h), \rho_h) \cap F = \emptyset$.

Using the assumption (2-ii), let $\delta_h > 0$ be such that

$$\overline{\lim_{\varepsilon}}\ \varepsilon \ln P(d_2(X_2^\varepsilon, S(h)) > \rho_h,\ d_1(X_1^\varepsilon, h) < \delta_h) \leq -R.$$

Then $\{I_1 \leq a\} \subset \bigcup_{I(h) \leq a} B_{d_1}(h, \delta_h)$.

Since $\{I_1 \leq a\}$ is compact there exists a finite sub-covering of $\{I_1 \leq a\}$ by balls $B_{d_1}(h_i, \delta_i)$, $1 \leq i \leq n$, where $\delta_i = \delta_{h_i}$.

Let $U = \bigcup_{i=1}^n B_{d_1}(h_i, \delta_i) \supset \{I_1 \leq a\}$. Then $F \subset \bigcap_{i=1}^n \left(B_{d_2}(S(h_i), \rho_i)\right)^c$

and

$$P(X_2^\varepsilon \in F) \leq P(X_1^\varepsilon \notin U) + P(X_1^\varepsilon \in U, X_2^\varepsilon \in F)$$

$$\leq P(X_1^\varepsilon \notin U) + \sum_{i=1}^n P(d_1(X_1^\varepsilon, h_i) < \delta_i,\ d_2(X_2^\varepsilon, S(h_i)) \geq \rho_i)$$

Lemma 1.10 (i) and upper estimate for X_1^ε and U^c imply

$$\overline{\lim_{\varepsilon}}\ \varepsilon \ln P(X_2^\varepsilon \in F) \leq \max\left(\overline{\lim}\ \varepsilon \ln P(X_1^\varepsilon \notin U),\ \max_{1 \leq i \leq n} \overline{\lim}\ \varepsilon \ln P(d_1(X_1^\varepsilon, h_i) < \delta_i, d_2(X_2^\varepsilon, S(h_i)) \geq \rho_i)\right)$$

$$\leq \max\left(-\inf\{I_1(h) : h \notin U\}, -R\right)$$

The inclusion $\{I_1 \leq a\} \subset U$ implies $U^c \subset \{I_1 > a\}$, so that $\inf\{I_1(h), h \notin U\} \geq a$. Since $R > a$, we conclude $\overline{\lim}\ \varepsilon \ln P(X_2^\varepsilon \in F) \leq -a$ $\square$

The transfer principle will provide a proof of LDP for diffusion processes.

Exercise Let $X = (X_1, \dots, X_d)$ be a Gaussian vector with mean 0 and covariance matrix Γ. Prove that the family of distributions of $\sqrt{\varepsilon} X$ satisfies a LDP with a rate function to be determined.

Give a more precise expression of this rate function if Γ is invertible.

2.2 The Freidlin-Wentzell theorem

Let $(W_t, t\geq 0)$ be a r - dimensional Brownian motion, $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^r$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ be Lipschitz functions, $x \in \mathbb{R}^d$. Then for any $\varepsilon > 0$, the SDE

$$X_t^\varepsilon = x + \sqrt{\varepsilon}\int_0^t \sigma(X_s^\varepsilon)\,dW_s + \int_0^t b(X_s^\varepsilon)\,ds$$

has a unique solution. Let $\mathcal{H} = \{h_t = \int_0^t \dot{h}_s\,ds : \dot{h} \in L^2(0,T;\mathbb{R}^r)\}$ and for $h \in \mathcal{H}$, let $S(h)_t = x + \int_0^t \sigma(S(h)_s)\dot{h}_s\,ds + \int_0^t b(S(h)_s)\,ds$ denote the controlled equation driven by h. Set $\|h\|_{\mathcal{H}}^2 = \int_0^T |\dot{h}_s|_{\mathbb{R}^r}^2\,ds$.

Then the solution of the SDE exists, is unique and a.s. belongs to $C([0,T];\mathbb{R}^d)$

Furthermore, given any $p \in [2,\infty)$, $\sup_{\varepsilon \in (0,1]} E\left(\sup_{t\in[0,T]} |X_t^\varepsilon|^p\right) < \infty$.

Using the Picard iteration method and Gronwall's lemma, it is easy to prove that the controlled equation $S(h)$ has a unique solution in $C([0,T];\mathbb{R}^d)$.

Theorem 2.4 (Freidlin-Wentzell) The family of diffusion processes (X^ε) satisfies on $C([0,T];\mathbb{R}^d)$ a LDP with the good rate function

$$\tilde{I}(f) = \inf\left\{\tfrac{1}{2}\|h\|_{\mathcal{H}}^2 : h \in \mathcal{H},\ S(h) = f\right\} \quad (\text{with the convention } \inf\emptyset = +\infty)$$

Proof We will apply the transfer principle for $X_1^\varepsilon = \sqrt{\varepsilon}\,W$ and $X_2^\varepsilon = X^\varepsilon$

Lemma 2.5 For any $a \in [0,\infty)$, the map $h \mapsto S(h)$ from $\{\|h\|_{\mathcal{H}} \leq a\}$ endowed with the topology of $C([0,T];\mathbb{R}^r)$ to $C([0,T];\mathbb{R}^d)$ is continuous.

Proof We approximate S by a sequence S^n which are continuous and converge to S uniformly on $\{\|\cdot\|_{\mathcal{H}} \leq a\}$. Since this set is compact, this will conclude the proof.

The idea is to "freeze" the coefficients at the left end point of a dyadic interval. To ease notations suppose $T=1$ and for $s \in [0,1]$, let $\underline{s}_n = 2^{-n} \lfloor 2^n s \rfloor$, that is $\underline{s}_n = k2^{-n}$ if $s \in [k2^{-n}, (k+1)2^{-n})$, $k=0,..,2^n$.

We prove that

(i) $S^n|_{\{\|h\|_{\mathcal{H}} \le a\}}$ is continuous for all $n \ge 1$,

(ii) $\lim_n \sup_{\|h\|_{\mathcal{H}} \le a} \| S^n(h) - S(h) \|_\infty = 0.$

Set $\tilde{s}_m = s$ or $\tilde{s}_m = \underline{s}_n$; let $T(h)$ be defined by

$$T(h)_t = x + \int_0^t \sigma(T(h)_{\tilde{s}_m}) \dot{h}_s ds + \int_0^t b(T(h)_{\tilde{s}_m}) ds$$

Using the Picard iteration scheme for $T(h)$, one proves that if $\|h\|_H \le a$ $\sup_{\|h\|_{\mathcal{H}} \le a} \sup_{s \le T} |T(h)_s|^2 \le C_1 \exp(C_2)$, where the constants C_1 and C_2 only depend on x, a, b and σ.

Let $g, h \in \mathcal{H}$ be such that $\|g\|_{\mathcal{H}} \vee \|h\|_{\mathcal{H}} \le a$. Set $\Delta^n(t) = |S^n(g)_t - S^n(h)_t|$

Then for $k=0,.., 2^n-1$,

$$\sup_{k2^{-n} \le t \le (k+1)2^{-n}} \Delta^n(t) \le \Delta^n(k2^{-n}) + |\sigma(S^n(g)_{k2^{-n}}) - \sigma(S^n(h)_{k2^{-n}})| \left| \int_{k2^{-n}}^t \dot{g}_u du \right|$$

$$+ |\sigma(S^n(h)_{k2^{-n}})| \left| \int_{k2^{-n}}^t (\dot{g}_u - \dot{h}_u) du \right| + 2^{-n} |b(S^n(g)_{k2^{-n}}) - b(S^n(g)_{k2^{-n}})|$$

$$\le \Delta^n(k2^{-n}) + c\sqrt{a}\, \Delta^n(k2^{-n}) + c(a) \|f-g\|_\infty + c\, \Delta^n(k2^{-n})$$

$$\le C(a) \Delta^n(k2^{-n}) + c(a) \|f-g\|_\infty$$

Hence for every n, $\sup_{t \in [0,T]} \Delta^n(t) \le 2^n C(a) \|f-g\|_\infty$ and $S^n|_{\{\|\cdot\|_{\mathcal{H}} \le a\}}$ is continuous.

Let $h \in \mathcal{H}$ be such that $\|h\|_{\mathcal{H}} \le a$. Then $\|S(h)\|_\infty \le C(a) < \infty$, and

$$\sup_{t \in [0,T]} |S(h)_t - S(h)_{\underline{t}_n}| \le \sup_{t \in [0,T]} \int_{\underline{t}_n}^t \left[|\sigma(S(h)_s)| |\dot{h}_s| + |b(S(h)_s)| \right] ds$$

$$\le C(a) \sup \left\{ \int_I (1 + |\dot{h}_s|) ds : I \text{ interval} \ni \lambda(I) \le 2^{-n} \right\}$$

$$\le C(a)\, 2^{-n/2}.$$

Hence for every $t \in [0,T]$,

$$|S(h)_t - S^n(h)_t|^2 \le 2\Big(\int_0^t |\sigma(S(h)_s) - \sigma(S^n(h)_{\underline{s}_n})|\,|\dot h_s|\,ds\Big)^2$$
$$+ 2\Big(\int_0^t |b(S(h)_s) - b(S^n(h)_{\underline{s}_n})|\,ds\Big)^2$$
$$\le C \sup_t |S(h)_t - S(h)_{\underline{t}_n}|^2 + C\Big(\int_0^t |S(h)_{\underline{s}_n} - S^n(h)_{\underline{s}_n}|\,(1+|\dot h_s|)\,ds\Big)^2$$
$$\le C(a)\,2^{-n} + C\int_0^t (1+|\dot h_s|^2) \sup_{u\le s} |S(h)_u - S^n(h)_u|^2\,ds$$

Gronwall's lemma concludes the proof of the convergence of $\sup_{t\in[0,1]} |S(h)_t - S^n(h)_t|$ to 0 uniformly on $\{\|\cdot\|_{\mathcal{H}} \le a\}$; this completes the proof of the lemma □

Proposition 2.6 (The Freidlin-Wentzell inequality)

Let X^ε and $S(h)$ be as in the statement of Theorem 2.4. Then given $a>0$, $\eta>0$ and $R>0$, there exists $\delta>0$ such that for any $h\in\mathcal{H}$ such that $\|h\|_{\mathcal{H}} \le a$, $\overline{\lim}\, \varepsilon \ln P(\|X^\varepsilon - S(h)\|_\infty \ge \eta,\ \|\sqrt{\varepsilon} W - h\|_\infty \le \delta) \le -R$

Proof The proof contains four steps

(1) Using Girsanov's theorem, reduce the proof to the case $h=0$

(2) Using a localization, reduce to the case σ and b are bounded

(3) Using Gronwall's lemma, reduce the proof to a control of the stochastic integral

(4) Use an exponential inequality on stochastic integrals and time discretization

Step 1 (reduce to $h=0$). Let $h\in\mathcal{H}$ such that $\|h\|_{\mathcal{H}} \le a$; for $\varepsilon>0$ let $\tilde{P}^\varepsilon$ be defined by $\frac{d\tilde P^\varepsilon}{dP} = \exp\Big(\frac{1}{\sqrt\varepsilon}\int_0^T \dot h_s\,dW_s - \frac{1}{2\varepsilon}\int_0^T |\dot h_s|^2\,ds\Big)$

Then $\tilde W^\varepsilon_t = W_t - \frac{h_t}{\sqrt\varepsilon}$ is a $\tilde P^\varepsilon$ Brownian motion, and under $\tilde P^\varepsilon$, X^ε is solution of $\tilde X^\varepsilon_t = x + \sqrt\varepsilon\int_0^t \sigma(\tilde X^\varepsilon_s)\,d\tilde W^\varepsilon_s + \int_0^t [\, b(\tilde X^\varepsilon_s) + \sigma(\tilde X^\varepsilon_s)\,\dot h_s\,]\,ds$

Let $\tilde{b}(s,x) = b(x) + \sigma(x)\dot{h}_s$, and for $g \in \mathcal{H}$ let $\tilde{S}(g)$ be the corresponding controlled equation where $b(x)$ is replaced by $\tilde{b}(s,x)$. Then $\tilde{S}(0)_t = S(h)_t$. Suppose that the Freidlin-Wentzell inequality holds at 0 with $\tilde{b}$ instead of b. For $R > 2a^2$, and $\eta > 0$, let $\delta > 0$ be such that $\overline{\lim}\ \varepsilon \ln \tilde{P}^\varepsilon(A^\varepsilon) \le -4R$, where

$$A^\varepsilon = \{ \|X^\varepsilon - S(h)\|_\infty \ge \eta,\ \|\sqrt{\varepsilon} W - h\|_\infty \le \delta \} = \{ \|\tilde{X}^\varepsilon - \tilde{S}(0)\|_\infty \ge \eta,\ \|\sqrt{\varepsilon}\tilde{W}\|_\infty \le \delta \}.$$

Then $P(A^\varepsilon) = \int 1_{A^\varepsilon} \frac{dP}{d\tilde{P}^\varepsilon} d\tilde{P}^\varepsilon \le (\tilde{P}^\varepsilon(A^\varepsilon))^{\frac12} \left(\int \left(\frac{dP}{d\tilde{P}^\varepsilon}\right)^2 d\tilde{P}^\varepsilon \right)^{\frac12}$

We check that $\int \left(\frac{dP}{d\tilde{P}^\varepsilon}\right)^2 d\tilde{P}^\varepsilon \le \exp\left(\frac{a^2}{\varepsilon}\right)$.

Since $W_t = \tilde{W}^\varepsilon_t + \frac{h_t}{\sqrt{\varepsilon}}$, $\frac{dP}{d\tilde{P}^\varepsilon} = \exp\left(-\frac{1}{\sqrt{\varepsilon}} \int_0^T \dot{h}_s\, d\tilde{W}^\varepsilon_s - \frac{1}{2\varepsilon}\int_0^T |\dot{h}_s|^2 ds\right)$, and

$$\int \left(\frac{dP}{d\tilde{P}^\varepsilon}\right)^2 d\tilde{P}^\varepsilon = \int \exp\left[-\frac{2}{\sqrt{\varepsilon}}\int_0^T \dot{h}^\varepsilon_s\, d\tilde{W}^\varepsilon_s - \frac{1}{\varepsilon}\int_0^T |\dot{h}_s|^2 ds\right] d\tilde{P}^\varepsilon$$
$$= \int \exp\left[-\frac{2}{\sqrt{\varepsilon}}\int_0^T \dot{h}^\varepsilon_s\, d\tilde{W}^\varepsilon_s - \frac{4}{2\varepsilon}\int_0^T |\dot{h}_s|^2 ds\right] d\tilde{P}^\varepsilon \ \exp\left(\frac{1}{\varepsilon}\int_0^1 |\dot{h}_s|^2 ds\right)$$
$$\le \exp\left(\frac{a^2}{\varepsilon}\right).$$

This implies, using the inequality $R > 2a^2$

$$\overline{\lim}\ \varepsilon \ln P(A^\varepsilon) \le \frac12 \left[\overline{\lim}\ \varepsilon \ln \tilde{P}^\varepsilon(A^\varepsilon) + \varepsilon \frac{a^2}{\varepsilon}\right] \le \frac12 [-R + a^2] \le -\frac{R}{4}.$$

We now replace b by $\tilde{b}(s,x)$, work under $\tilde{P}^\varepsilon$ (denoted P) and suppose $h = 0$ ($\tilde{S}(0) = S(h)$).

<u>Step 2</u> Reduce the proof to the case σ and b bounded. (denote W the B.M.)

We have proved $\sup_{\|h\|_{\mathcal{H}} \le a} \sup_{t \in [0,T]} |S(h)_t| = \sup_{\|h\|_{\mathcal{H}} \le a} \sup_{t\in[0,T]} |\tilde{S}(0)_t| \le M < \infty$.

Let $\tau^\varepsilon = \inf\{t \ge 0 : |X^\varepsilon_t - S(h)_t| \ge \eta\} \wedge T$; the continuity of X^ε and $S(h)$ imply $P(\sup_{t\in[0,T]} |X^\varepsilon_t - S(h)_t| \ge \eta,\ \|\sqrt{\varepsilon}W\|_\infty < \delta) = P(\sup_{t \le \tau} |X^\varepsilon_t - S(h)_t| \ge \eta,\ \|\sqrt{\varepsilon}W\|_\infty < \delta)$

For $t \le \tau^\varepsilon(\omega)$, we have $|X^\varepsilon_t| \le M + \eta$. Let $\bar{\sigma}$ (resp. $\bar{b}$) denote the coefficient σ (resp. b) truncated at $c(1 + M + \eta)$, so that $\bar{\sigma}(X^\varepsilon_t) = \sigma(X^\varepsilon_t)$ and $\bar{b}(X^\varepsilon_t) = b(X^\varepsilon_t)$ for $t \le \tau^\varepsilon(\omega)$; Let $\bar{X}^\varepsilon$ (resp. $\bar{S}(0)$) denote the solution of $\quad \bar{X}^{\varepsilon}_t = x + \int_0^t \bar{\sigma}(\bar{X}^{\varepsilon}_s)dW_s + \int_0^t [\bar{b}(\bar{X}^{\varepsilon}_s) + \bar{\sigma}(\bar{X}^{\varepsilon}_s)\dot{h}_s]ds$,

resp $\quad \bar{S}(0)_t = x + \int_0^t [\bar{b}(\bar{S}(0)_s) + \bar{\sigma}(\bar{S}(0)_s)\,\dot{h}_s]ds$,

then $\bar{X}^{\varepsilon}_t = X^{\varepsilon}_t$ and $\bar{S}(0)_t = \tilde{S}(0)_t = S(h)_t$ for $t \leq \tau^{\varepsilon}(\omega)$. Hence it is enough

to check that $\overline{\lim}\, \varepsilon \ln P(\|\bar{X}^{\varepsilon} - \bar{S}(0)\|_\infty \geq \eta,\ \|\sqrt{\varepsilon}w\|_\infty \leq \delta) \leq -R$

We may and do assume that the coefficients σ and b are bounded. (and keep the notations X^{ε}, $\tilde{S}(0)$, σ and b.

<u>Step 3</u> Reduce the proof to the stochastic integral.

Lemma 2.7 (<u>generalized Gronwall lemma</u>)

Let $f \geq 0$, $\varphi \in L_1^+(dt)$ be such that $f(t) \leq a + \int_0^t \varphi(s)f(s)ds$ for any $t > 0$.

Then if f is bounded, $f(t) \leq a\exp(\int_0^t \varphi(s)ds)$.

<u>Proof of lemma 2.7</u> Let f be bounded by M. Then for every t,

$$f(t) \leq a + \int_0^t [a + \int_0^{s_1} f(s_2)\varphi(s_2)ds_2]\,\varphi(s_1)ds_1$$

$$\leq a + a\int_0^t \varphi(s_1)ds_1 + \cdots + a\int_0^t \varphi(s_1)\int_0^{s_1}\varphi(s_2)\cdots\left(\int_0^{s_{n-1}}\varphi(s_n)ds_n\right)\cdots ds_2\,ds_1$$

$$+ \int_0^t \varphi(s_1)\int_0^{s_1}\varphi(s_2)\cdots\int_0^{s_n} f(s_{n+1})\varphi(s_{n+1})\,ds_{n+1}\,ds_n\cdots ds_2\,ds_1$$

Set $\Phi(t) = \int_0^t \varphi(s)ds$; then $\dot{\Phi}(t) = \varphi(t)$ for almost every t, and for every integer $k \geq 1$

$$\int_0^{t_1}\varphi(s_1)\int_0^{s_1}\varphi(s_2)\cdots\int_0^{s_{k-1}}\varphi(s_k)\,ds_k\,ds_{k-1}\cdots ds_2\,ds_1 = \frac{\Phi(t)^k}{k!}$$

Hence $f(t) \leq a\left[1 + \Phi(t) + \frac{\Phi^2(t)}{2!} + \cdots + \frac{\Phi^n(t)}{n!}\right] + M\,\frac{\Phi(t)^{n+1}}{(n+1)!}$

This concludes the proof as $n \to \infty$

□

For every $t \in [0,1]$, using the Lipschitz property of σ and b, we deduce

$$|X^\varepsilon_t - S(h)_t| \le |\sqrt{\varepsilon}\int_0^t \sigma(X^\varepsilon_s)dW_s| + c\int_0^t [1+|\dot{h}_s|]\, |X^\varepsilon_s - S(h)_s| ds$$

For almost every ω, $X^\varepsilon_\cdot - S(h)_\cdot$ is continuous on $[0,1]$, hence bounded.

Therefore, $f(s) = \sup\{|X^\varepsilon_u - S(h)_u| : u \le s\}$ is bounded on $[0,1]$.

Furthermore, $\int_0^T [1+|\dot{h}_s|]ds \le T + \|\dot{h}\|_{L^2} \le T + \sqrt{2a}$. Therefore Lemma 2.7 implies $|X^\varepsilon_t - S(h)_t| \le \sqrt{\varepsilon}\, |\int_0^t \sigma(X^\varepsilon_s)dW_s| \exp(c(a))$, and the proof is reduced to checking $\overline{\lim}\, \varepsilon \ln P(\|\sqrt{\varepsilon}\int_0^\cdot \sigma(X^\varepsilon_s)dW_s\|_\infty \ge \eta,\ \|\sqrt{\varepsilon}W\|_\infty \le \delta) \le -R$.

<u>Step 4</u> Exponential inequalities of stochastic integrals.

$\boxed{\text{Lemma 2.8}}$ (<u>Stroock</u>) Let $Y : [0,T]\times\Omega \to \mathbb{R}^d \otimes \mathbb{R}^r$, $Z : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ be progressively measurable processes, let τ be a stopping time such that

$$\sup_{\|\xi\|_{\mathbb{R}^d}=1}\ \sup_{t\le\tau(\omega)} \Big\|\sum_{i=1}^d \langle \xi_i, Y^i_\cdot(t,\omega)\rangle_{\mathbb{R}^r}\Big\| \le A, \text{ and } \sup_{t\le\tau(\omega)} \|Z(t,\cdot)\|_{\mathbb{R}^d} \le B.$$

Then for $R > BT$,

$$P\Big(\sup_{t\le T}\Big|\int_0^{t\wedge\tau} Y(s)\, dW_s + \int_0^{t\wedge\tau} Z(s)\, ds\Big|_{\mathbb{R}^d} \ge R\Big) \le 2d \exp\Big(-\frac{(R-BT)^2}{2dTA^2}\Big).$$

<u>Remark</u> In this statement, ${}^tY$ is an operator from $\mathbb{R}^d$ to $\mathbb{R}^r$.

<u>Proof of Lemma 2.8</u> For $t \le T$, $\|\int_0^t Z(s,\omega)ds\|_{\mathbb{R}^d} \le TB$, and for $TB < R$,

$$\Big\{\sup_{t\le T}\Big\|\int_0^{t\wedge\tau} Y(s)\, dW_s + \int_0^{t\wedge\tau} Z(s)ds\Big\|_{\mathbb{R}^d} \ge R\Big\} \subset \Big\{\sup_{t\le T}\Big\|\int_0^{t\wedge\tau} Y(s)dW_s\Big\|_{\mathbb{R}^d} \ge R - TB\Big\}$$

Let v be a vector of $\mathbb{R}^d$, and set

$$M_t(v) = \exp\Big(\langle v, \int_0^{t\wedge\tau} Y_s dW_s\rangle - \frac{1}{2}\int_0^{t\wedge\tau} \Big\|\sum_{i=1}^d \langle v_i, Y^i_s\rangle\Big\|^2_{\mathbb{R}^r} ds\Big)$$

Then $M_\cdot(v)$ is an exponential martingale. Doob's inequality implies that for any vector v of $\mathbb{R}^d$ with $\|v\|_{\mathbb{R}^d}=1$, and every $\lambda>0$,

$$P\Big(\sup_{0\le t\le T}\langle v,\int_0^{t\wedge\tau} Y_s\,dW_s\rangle \ge \frac{R-BT}{\sqrt d}\Big) \le P\Big(\sup_{t\in[0,T]} M_t(\lambda v)\ge \exp\big(\lambda\frac{R-BT}{\sqrt d} - \frac{\lambda^2}{2}A^2T\big)\Big)$$
$$\le \exp\Big(-\lambda\frac{R-BT}{\sqrt d} + \frac{\lambda^2}{2}A^2T\Big)$$

Indeed, for $s\le\tau(\omega)$, $\lambda\,\|\sum_{i=1}^d\langle v_i, Y^i_\cdot(s,\omega)\rangle\|_{\mathbb{R}^r}\le\lambda A^2$

The minimum over $\lambda>0$ of the above upper bound is achieved for $\lambda=\frac{R-BT}{\sqrt d\,TA^2}$

This implies

$$P\Big(\sup_{t\le T}\langle v,\int_0^{t\wedge\tau} Y_s\,dW_s\rangle\ge\frac{R-BT}{\sqrt d}\Big)\le\exp\Big(-\frac{(R-BT)^2}{2dTA^2}\Big).$$

Let $v_1,\dots,v_d$ be an ONB of $\mathbb{R}^d$. Then

$$P\Big(\sup_{t\in[0,T]}\Big\|\int_0^{t\wedge\tau} Y_s\,dW_s\Big\|\ge R-BT\Big)=P\Big(\sup_{t\in[0,T]}\sum_{i=1}^d\langle v_i,\int_0^{t\wedge\tau} Y_s\,dW_s\rangle^2\ge(R-BT)^2\Big)$$
$$\le\sum_{i=1}^d P\Big(\sup_{t\in[0,T]}\Big|\langle v_i,\int_0^{t\wedge\tau} Y_s\,dW_s\rangle\Big|\ge\frac{R-BT}{\sqrt d}\Big)\le 2d\exp\Big(-\frac{(R-BT)^2}{2dTA^2}\Big)$$

$\square$

To complete the proof of Step 4, we use some time discretization. Let $T=1$. For every integer $n\ge1$, let $X^{\varepsilon,n}_t=X^\varepsilon_{\underline{t}_n}$ (that is $X^{\varepsilon,n}_t=X^\varepsilon_{k2^{-n}}$ for $t\in[k2^{-n},(k+1)2^{-n})$). Then for $\eta>0$, $\delta>0$, $\gamma>0$ and $n\ge1$,

$\{\sup_{t\in[0,1]}|\sqrt\varepsilon\int_0^t\sigma(X^\varepsilon_s)dW_s|\ge\eta,\ \|\sqrt\varepsilon W\|_\infty<\delta\}\subset A^\varepsilon\cup B^\varepsilon\cup C^\varepsilon$, where

$$A^\varepsilon=\{\sup_{t\le1}|X^\varepsilon_t-X^{\varepsilon,n}_t|>\gamma\},$$
$$B^\varepsilon=\{\sup_{t\le1}|X^\varepsilon_t-X^{\varepsilon,n}_t|\le\gamma\}\cap\{\sup_{t\le1}|\sqrt\varepsilon\int_0^t[\sigma(X^\varepsilon_s)-\sigma(X^{\varepsilon,n}_s)]dW_s>\frac{\eta}{2}\}$$
$$C^\varepsilon=\{\sup_{t\le1}|\sqrt\varepsilon\int_0^t\sigma(X^{\varepsilon,n}_s)\,dW_s|>\frac{\eta}{2}\}\cap\{\|\sqrt\varepsilon W\|_\infty<\delta\}$$

We first choose γ such that $\overline{\lim}\,\varepsilon\ln P(B^\varepsilon)\le -R$; we then choose n such that $\overline{\lim}\,\varepsilon\ln P(A^\varepsilon)\le-R$. We finally choose δ such that $C_\varepsilon=\emptyset$.

(i) Check that given η and n, one can chose $\delta > 0$ such that $C^\varepsilon = \emptyset$.

Indeed, for $t_k = k2^{-n}$, $\varepsilon \in (0,1]$,

$$\sqrt{\varepsilon}\int_0^t \sigma(X^{\varepsilon,n}_s)dW_s = \sum_{k=0}^{2^n-1} \sqrt{\varepsilon}\,\sigma(X^\varepsilon_{t_k})[W_{t\wedge t_{k+1}} - W_{t\wedge t_k}],$$

and $\left|\sqrt{\varepsilon}\int_0^t \sigma(X^{\varepsilon,n}_s)dW_s\right| \le 2^n\, 2\delta\, C$ on $\{\|\sqrt{\varepsilon}W\|_\infty < \delta\}$ for $\|\sigma\|_\infty \le C$. On choses δ such that $2^{n+1}\delta C < \frac{\eta}{2}$; this implies $C_\varepsilon = \emptyset$.

(ii) We next choose $\gamma > 0$ such that $\overline{\lim}\, \varepsilon \ln P(B^\varepsilon) \le -R$.

Indeed, using Lemma 2.8 with $T = \tau = 1$, $A = C\sqrt{\varepsilon}\gamma$, $B = 0$, $R = \frac{\eta}{2}$, we obtain

$$P(B^\varepsilon) \le 2d\exp\left(-\frac{\frac{\eta^2}{4}}{2d\,C^2\gamma^2\varepsilon}\right)$$

Hence for $\frac{\eta^2}{8d\,C^2\gamma^2} \ge R$, $\overline{\lim}\,\varepsilon\ln P(B^\varepsilon) \le -R$

(iii) We next prove that $\overline{\lim}\,\varepsilon \ln P(A^\varepsilon) \le -R$ for large n

For any $\gamma > 0$, if $t_k = k2^{-n}$, $k = 0, 1, \dots, 2^n$.

$$P\Big(\sup_{t\in[0,1]} |X^\varepsilon_t - X^{\varepsilon,n}_t| > \gamma\Big) \le \sum_{k=0}^{2^n-1} P\Big(\sup_{t_k\le t\le t_{k+1}} \Big|\int_{t_k}^t \sqrt{\varepsilon}\,\sigma(X_s)dW_s\Big| \ge \frac{\gamma}{2}\Big)$$
$$+ \sum_{k=0}^{2^n-1} P\Big(\sup_{t_k\le t\le t_{k+1}} \Big|\int_{t_k}^t [b(X^\varepsilon_s) + \sigma(X^\varepsilon_s)\dot{h}_s]ds\Big| \ge \frac{\gamma}{2}\Big)$$

Each deterministic integral is upper bounded by $C\left|\int_{t_k}^t (1+|\dot{h}_s|)ds\right|$. Hence, the Cauchy-Schwarz inequality implies that $\left|\int_{t_k}^t (1+|\dot{h}_s|)ds\right| \le 2^{-\frac{n}{2}}\, C(a)$. Hence for n large enough, each term $P\Big(\sup_{t_k\le t\le t_{k+1}} \Big|\int_{t_k}^t [b(X^\varepsilon_s) + \sigma(X^\varepsilon_s)\dot{h}_s]ds\Big| \ge \frac{\gamma}{2}\Big)$ is null

For each $k = 0, \dots, 2^n - 1$, Lemma 2.7 used with $A = \sqrt{\varepsilon}C$, $B = 0$, $T = 2^{-n}$ and $\tau = 1$ implies we deduce $P\Big(\sup_{t_k\le t\le t_{k+1}} \Big|\int_{t_k}^t \sqrt{\varepsilon}\,\sigma(X_s)dW_s\Big| \ge \frac{\gamma}{2}\Big) \le 2d\exp\left(-\frac{\frac{\gamma^2}{4}}{2d\,2^{-n}C^2\varepsilon}\right)$.

Hence, $\sum_{k=0}^{2^{n-1}} P\left(\sup_{t_k\le s\le t_{k+1}} \left|\sqrt{\varepsilon}\int_{t_k}^{t}\sigma(X_s^\varepsilon)dW_s\right| \ge \frac{\gamma}{2}\right) \le 2^{n+1} d \exp\left(-\frac{\gamma^2}{8d\,2^{-n}c^2\varepsilon}\right)$,

and $\overline{\lim}\, \varepsilon \ln P(A^\varepsilon) \le -\frac{\gamma^2 2^n}{8dc^2} \le -R$ if $2^{n-3} \ge \frac{dc^2}{\gamma^2}R$.

The successive choice of γ, n and δ imply by Lemma 1.10 (c) that

$$\overline{\lim}\, \varepsilon \ln P\left(\left\|\sqrt{\varepsilon}\int_0^{\cdot}\sigma(X_s^\varepsilon)dW_s\right\|_\infty \ge \eta,\ \|\sqrt{\varepsilon}W\|_\infty < \delta\right) \le -R.$$

This concludes the proof. $\square$

We can write more precisely the rate function in Theorem 2.4 if $d=r$ and $\sigma(y)$ is invertible for every y. Indeed, if $f_t = x + \int_0^t \sigma(f_s)\dot{h}_s ds + \int_0^t b(f_s)ds$, f is absolutely continuous and $\dot{f}_t = \sigma(f_t)\dot{h}_t + b(f_t)$

Hence, if $\sigma(f_t)$ is invertible $\dot{h}_t = \sigma(f_t)^{-1}[\dot{f}_t - b(f_t)]$

Therefore, in this case $\tilde{I}(f) = \frac{1}{2}\int_0^T |\sigma(f_t)^{-1}\{\dot{f}_t - b(f_t)\}|_{\mathbb{R}^d}\, dt$ if f is absolutely continuous and $\tilde{I}(f) = +\infty$ otherwise

If σ is a rectangular matrix, for $y\in\mathbb{R}^d$ set $a(y) = \sigma(y)\tilde{\sigma}(y)$. This $d\times d$ matrix defines a quadratic form Q_y on $\mathbb{R}^d$ by

$$Q_y(v) = \|\tilde{\sigma}(y)v\|^2_{\mathbb{R}^d} = v\,a(y)\,\tilde{v}.$$

The adjoint quadratic form on $\mathbb{R}^r$ is defined by

$$Q_y^*(v) = \inf\{\|z\|^2_{\mathbb{R}^r} : z\in\mathbb{R}^r,\ \sigma(y)z = v\}$$

Then Q_y^* is the Legendre transform of Q_y.

The rate function, $\tilde{I}$ in Theorem 2.4 is

$$\tilde{I}(f) = \frac{1}{2}\int_0^1 Q_{f_t}^*[\dot{f}_t - b(f_t)]dt$$ if f is absolutely continuous, and $\tilde{I}(f) = +\infty$ otherwise.

2.3 Diffusion exit from a Domain

Consider the diffusion $dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\,\sigma(X_t^\varepsilon)\,dW_t$ on $\mathbb{R}^d$, with initial condition $X_0^\varepsilon = x$ in a bounded domain G of $\mathbb{R}^d$. Suppose that (W_t) is an r-dimensional standard Brownian motion and that $b: \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma: \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^k$ are Lipschitz functions.

We will make the following assumption throughout this section

$\boxed{\text{Assumption A1}}$ The unique stable equilibrium point in G of the ODE $\dot{\phi}_t = b(\phi_t)$ is $0 \in G$. Suppose furthermore that $\phi_t \in G$ for any $t > 0$ if $\phi_0 \in G$, and that $\lim_{t\to+\infty} \phi(t) = 0$ for $\phi_0 \in \overline{G}$.

When $x \in G$ and ε is small, one can guess that X_t^ε will stay inside the domain G for some time. The question is "how long"? Where does it leave G if it does?

Suppose that ∂G is smooth enough, and let τ^ε be the stopping time defined by

$$\tau^\varepsilon = \inf\{t > 0 : X_t^\varepsilon \in \partial G\}.$$

We will denote by E_x the expected value with respect to the diffusion $(X^\varepsilon_\cdot)$ with initial condition $x \in G$.

Given $y, z \in \mathbb{R}^d$, $t > 0$, let

$$V(y,z,t) := \inf\{ I_{y,t}(f) : f \in C([0,T]; \mathbb{R}^d),\ f(t) = z\},$$ where

$$I_{y,t}(f) = \inf\left\{ \frac{1}{2}\int_0^t |\varphi(s)|^2 ds : \varphi \in L^2(0,t;\mathbb{R}^d),\ f(t) = y + \int_0^t [b(f(s)) + \sigma(f(s))\varphi(s)]\,ds \right\}$$
(with $\inf \emptyset = +\infty$)

Then $V(y,z,t)$ is the minimal energy required to reach z at time t, starting from y at time 0 and using a controlled equation.

Set $\quad V(y,z) = \inf\{V(y,z,t) : t>0\}$.

For every $z \in \mathbb{R}^d$, $V(0,z)$ is the quasi-potential

We will need additional assumptions

Assumption A2 $\quad \overline{V} = \inf\{V(0,z) : z \in \partial G\} < \infty$

Assumption A3 $\quad$ There exists $M<\infty$ such that for any $\delta>0$ small enough and all x, y such that $|x-z|+|z-y| \le \delta$ for some $z \in \partial G \cup \{0\}$, there exists a function u such that $\|u\|_\infty \le M$, $\phi_{T(\rho)} = y$ with $T(\rho) \to 0$ as $\rho \to 0$, and $\quad \phi_t = x + \int_0^t b(\phi_s)\,ds + \int_0^t \sigma(\phi_s)\, u_s\, ds.$

Note that if the matrices $\sigma(x)\tilde{\sigma}(x)$ are positive definite for $x=0$ and uniformly positive definite for $x \in \partial G$, then Assumption A3 is satisfied.

The following result that exit time and location are governed by energy minimization.

Theorem 2.9 $\quad$ Assume that Assumptions (A1)–(A3) are satisfied. Then

(i) For all $x \in G$ and $\delta > 0$,

$$\lim_{\varepsilon \to 0} P_x\left(e^{\frac{\overline{V}-\delta}{\varepsilon}} < \tau^\varepsilon < e^{\frac{\overline{V}+\delta}{\varepsilon}}\right) = 1,$$

and $\quad \lim_{\varepsilon\to 0} \varepsilon \ln E_x(\tau^\varepsilon) = \overline{V}$ for all $x \in G$.

(ii) Let $N \subset \partial G$ be a closed set such that $\inf\{V(0,z) : z \in N\} > \overline{V}$. Then for any $x \in G$, $\lim_{\varepsilon\to 0} P_x(X^\varepsilon_{\tau^\varepsilon} \in N) = 0$

Therefore, if there exists $z^* \in \partial G$ such that $V(0,z^*) < V(0,z)$ for any $z \in \partial G$, $z \neq z^*$, then for any $\delta > 0$ and $x \in G$,

$$\lim_{\varepsilon \to 0} P_x(|X^\varepsilon_{\tau^\varepsilon} - z^*| < \delta) = 1.$$

The proof is quite technical and will be omitted. Let us give an heuristic description of the proof.

One step consists in refining Theorem 2.4, proving LDP uniformly on the initial condition in a compact set.

Theorem 2.10 Assume that the assumptions of Theorem 2.4 are satisfied. Let $(X^{\varepsilon,y}_t)$ denote the solution to $X^{\varepsilon,y}_t = y + \int_0^t \sigma(X^{\varepsilon,y}_s)\,dW_s + \int_0^t b(X^{\varepsilon,y}_s)\,ds$, and for $\phi \in \mathcal{C}([0,T];\mathbb{R}^d)$, $I_y(\phi) = \inf\{\frac{1}{2}\|h\|^2_{\mathcal{H}} : S_y(h) = \phi\}$, where $S_y(h)_t = y + \int_0^t \sigma(S_y(h)_s)\dot{h}_s\,ds + \int_0^t b(S_y(h)_s)\,ds$.

Then for every compact set K, every closed (resp. open) set F (resp. G),

$$\overline{\lim_{\varepsilon}}\ \varepsilon \ln\left[\sup_{y\in K} P(X^{\varepsilon,y} \in F)\right] \leq -\inf\{I_y(\phi) : \phi \in F, y \in K\},$$

$$\underline{\lim_{\varepsilon}}\ \varepsilon \ln P\left[\inf_{y \in K} P(X^{\varepsilon,y} \in G)\right] \geq -\sup_{y\in K} \inf\{I_y(\phi) : \phi \in G\}.$$

One then proves that the process X^ε_t is attracted towards a small ball containing 0 without hitting ∂G on its way with a "high probability": for $B_\rho = \{x : \|x\| \leq \rho\} \subset G$, $\sigma_\rho = \inf\{t \geq 0 : X^\varepsilon_t \in B_\rho \cup \partial G\}$, $\lim_{\varepsilon} P_x(X^\varepsilon_{\sigma_\rho} \in B_\rho) = 1$.

$S_{2\rho} = \{x : \|x\| = 2\rho\} \subset G$ be a sphere around 0. One defines recursively stopping times by $\theta_0 = 0$, $\tau_m = \inf\{t \geq \theta_m : X^\varepsilon_t \in B_\rho \cup \partial G\}$, $\theta_{m+1} = \inf\{t > \tau_m : X^\varepsilon_t \in S_\rho\}$, with the convention $\theta_{m+1} = +\infty$ if $X_{\tau_m} \in \partial G$.

The process $Z_k = X^\varepsilon_{\tau_k}$ is a Markov chain.

One proves $\overline{\lim_\varepsilon}\ \varepsilon \ln P(\sup_{x \in S_\rho} X^\varepsilon_{\sigma_\rho} \in \partial G) \leq -\overline{V} + \frac{\delta}{2}$ for $\rho > 0$

small enough, and upper-estimates $P_x(\tau^\varepsilon \leq k T_0)$ for some $T_0 > 0$.

2.4 Uniform large deviations for parabolic SPDEs

We extend the results of sections 2.2 and 2.3 to solutions to the solution to a 1-d parabolic PDE subject to a random perturbation driven a space-time white noise.

Let $(W(t,x),\ t\in[0,T],\ x\in[0,1])$ be a space-time white noise and

$$\mathcal{H} = \left\{ h \in \mathcal{C}([0,T]\times[0,1]) : h(t,x) = \int_0^t\int_0^x \dot{h}(u,z)\,du\,dz \ :\ \dot{h}\in L^2((0,T)\times(0,1)) \right\}$$

Let $\sigma : [0,T]\times[0,1]\times\mathbb{R} \longrightarrow \mathbb{R}$ and $b : [0,T]\times[0,1]\times\mathbb{R}\longrightarrow\mathbb{R}$ be such that there exist constants $\bar{b}$, $\bar{\sigma}$ and B such that

$$\sup\{|\sigma(t,x,z)| + |b(t,x,z)| : (t,x)\in[0,T]\times[0,1]\} \leq B(1+|z|),$$

$$|b(t,x,y)-b(t,x,z)| \leq \bar{b}\,|y-z|,\quad |\sigma(t,x,y)-\sigma(t,x,z)|\leq\bar{\sigma}\,|y-z|,\quad \forall (t,x)\in[0,T]\times[0,1],\ \forall y,z\in\mathbb{R}$$

Consider the stochastic one-dimensional heat equation

$$\begin{cases} \partial_t X^\varepsilon_\varphi(t,x) = D^2_{xx}X^\varepsilon_\varphi(t,x) + b(t,x,X^\varepsilon_\varphi(t,x)) + \sqrt{\varepsilon}\,\sigma(t,x,X^\varepsilon_\varphi(t,x))\,\ddot{W}_{t,x} \\ X^\varepsilon_\varphi(0,x) = \varphi(x) \end{cases}$$

for $\varphi\in\mathcal{C}^{2\alpha}([0,1])$, $0\leq\alpha<\frac{1}{4}$ with Dirichlet boundary conditions

$X^\varepsilon_\varphi(t,0) = X^\varepsilon_\varphi(t,1) = 0$, or Neumann boundary conditions

$\partial_x X^\varepsilon_\varphi(t,0) = \partial_x X^\varepsilon_\varphi(t,1) = 0$. Here $\ddot{W}$ denotes the formal derivative of the Brownian sheet W.

Let G denote the Green function associated with $\partial_t u = \Delta u$ on $[0,1]$, either the Dirichlet or Neumann boundary conditions.

Using martingale measures, J. B. Walsh gave a formal meaning to the previous SPDE as follows

$$X_\varphi^\varepsilon(t,x) = G_t(x,\varphi) + \int_0^t\int_0^1 G_{t-u}(x,z)\, b(u,z,X_\varphi^\varepsilon(u,x))\,du\,dz$$
$$+\sqrt{\varepsilon}\int_0^t\int_0^1 G_{t-u}(x,z)\,\sigma(u,z,X_\varphi^\varepsilon(u,z))\,W(du,dz),$$

where $G_t(x,\varphi) = \int_0^1 G_t(x,z)\,\varphi(z)\,dz$.

Then J.B. Walsh proved that there exists a unique strong solution in the probabilistic sense (that is we keep the given Brownian sheet W, and do not only deal with distributions), but weak in the analytical sense (that is inner products). This solution belongs a.s. to $\mathcal{C}^{\alpha,2\alpha}([0,T]\times[0,1])$, that is, is α-Hölder continuous in time and 2α-Hölder continuous in space.

For $\alpha\in[0,\frac{1}{4})$, $f:[0,T]\times[0,1]\to\mathbb{R}$, set

$$\|f\|_\alpha = \sup\{|f(t,x)| : (t,x)\in[0,T]\times[0,1]\} + \sup\left\{\frac{|f(t,x)-f(s,y)|}{(|t-s|^2+|x-y|^4)^{\frac{\alpha}{2}}} : \begin{array}{l}(t,x),(s,y)\in[0,T]\times[0,1]\\ (t,x)\neq(s,y)\end{array}\right\}$$

and let $\mathcal{C}^{\alpha,0}([0,T]\times[0,1])$ denote the set of functions $f:[0,T]\times[0,1]\to\mathbb{R}$ such that

$$\|f\|_\alpha<\infty \quad\text{and}\quad \lim_{|t-s|+|x-y|\to 0}\frac{|f(t,x)-f(s,y)|}{|t-s|^\alpha+|x-y|^{2\alpha}} = 0.$$

Let $\mathcal{C}^{\alpha,0}([0,1])$ be defined in a similar way.

Using a version of the transfer principle uniform on the initial condition in a compact subset of $\mathcal{C}^{\alpha,0}([0,1])$ and arguments similar to those used to prove Theorem 2.4 and Theorem 2.10, we deduce the following

Theorem 2.11 Let $\alpha\in(0,\frac{1}{4})$, K be a compact subset of $\mathcal{C}^{2\alpha,0}([0,1])$. Then for every closed (resp. open) set F (resp. $\mathcal{O}$) of $\mathcal{C}^{\alpha,0}([0,T]\times[0,1])$

$$\overline{\lim}\,\varepsilon\ln\left[\sup_{\varphi\in K} P(X_\varphi^\varepsilon\in F)\right] \le -\inf\left\{\tfrac{1}{2}\|h\|_{\mathcal{H}}^2 : S_\varphi(h)=f,\ \varphi\in K,\ f\in F\right\},$$

$$\underline{\lim}\,\varepsilon\ln\left[\inf_{\varphi\in K} P(X_\varphi^\varepsilon\in\mathcal{O})\right] \ge -\sup_{\varphi\in K}\inf\left\{\tfrac{1}{2}\|h\|_{\mathcal{H}}^2 : S_\varphi(h)=f,\ f\in\mathcal{O}\right\},$$

where $S_\varphi(h)(t,x) = G_t(x,\varphi) + \int_0^t \int_0^1 G_{t-u}(x,z)\{\sigma(u,z,S_\varphi(h)(u,z))\dot{h}(u,z)$
$+ b(u,z,S_\varphi(h)(u,z))\}\,du\,dz$

The proof depends on upper estimates of the Green function, its space and time derivatives.

Suppose now that b and σ do not depend on t, and that the solution satisfies homogeneous Dirichlet boundary conditions. For an initial condition φ in a bounded open domain A of $\mathcal{C}^{2\alpha,0}([0,1])$ such that $\varphi(0)=\varphi(1)=0$, set

$$\tau^\varepsilon = \inf\{t \geq 0 : X^\varepsilon_\varphi(t,\cdot) \in A^c\} = \inf\{t \geq 0 : X^\varepsilon_\varphi(t,\cdot) \in \partial A\}$$

Since closed balls of $\mathcal{C}^{2\alpha}([0,1])$ are not compact, we have to introduce $\mathcal{C}^{2\beta,0}([0,1])$ for $\beta \in (0,\alpha)$. Let $\tilde{A}$ denote the closure of A in $\mathcal{C}^{2\beta,0}([0,1])$. Then $\tilde{A}$ is a compact subset of $\mathcal{C}^{2\beta,0}([0,1])$. In order to prove estimates on the exit time τ^ε, we make assumptions similar to that in Section 2.3.

(H1) Given an initial condition $\psi(0,\cdot) \in \tilde{A}$, the trajectory ψ of the solution of the PDE $\partial_t\psi(t,x) = \Delta\psi(t,x) + b(x,\psi(t,x))$ satisfy
$\lim_{t\to+\infty} \|\psi(t,\cdot)\|_\beta = 0$.
Furthermore, if $\psi(0,x) \in A$, then $\psi(t,\cdot) \in A$ for all $t>0$

The statements of (H2) and (H3), similar to (A2) and (A3), is omitted.

We then have the following for $\hat{V} = \inf\{V(0,\varphi) : \varphi \in \partial A\}$ and $V(\phi,\psi) = \inf_{t>0} V(\phi,\psi,t)$

Theorem 2.12 Let σ, b be as above, and suppose that (H1)-(H3) hold. Then

(i) For every $\varphi \in A$, $\delta > 0$, $\lim_{\varepsilon\to 0} P_\varphi\left(e^{\frac{\hat{V}-\delta}{\varepsilon}} < \tau^\varepsilon < e^{\frac{\hat{V}+\delta}{\varepsilon}}\right) = 1$, $\lim_{\varepsilon\to 0} \varepsilon \ln E_\varphi(\tau^\varepsilon) = \hat{V}$.

(ii) For $N \subset \partial A$ such that $\hat{V} < \inf\{V(0,\psi), \psi \in N\}$, $\lim_{\varepsilon} P_\varphi(X^\varepsilon_\varphi(\tau^\varepsilon,\cdot) \in N) = 0$ for $\varphi \in A$

3 Varadhan's lemma - Weak convergence approach

3.1 Varadhan's lemma

In this section we suppose that X_ε is a family of random variables taking values in a regular topological space $\mathcal{X}$ (for every $x \in \mathcal{X}$, F closed subset of $\mathcal{X}$ such that $x \notin F$, there exist disjoint open sets G_1 and G_2 such that $x \in G_1$ and $F \subset G_2$).

$\boxed{\text{Theorem 3.1}}$ (Varadhan's lemma) Suppose that the family (μ_ε) of distributions of X_ε satisfies a LDP with the good rate function $I : \mathcal{X} \to [0,+\infty]$. Let $\phi : \mathcal{X} \to \mathbb{R}$ be a continuous function such that either

(i) $$\lim_{M\to\infty} \overline{\lim_{\varepsilon}}\ \varepsilon \ln E\left[e^{\frac{\phi(X_\varepsilon)}{\varepsilon}} 1_{\{\phi(X_\varepsilon)\geq M\}}\right] = -\infty$$

or

(ii) there exists $\gamma > 1$ such that $$\overline{\lim_{\varepsilon}}\ \ln E\left[e^{\frac{\gamma\phi(X_\varepsilon)}{\varepsilon}}\right] < \infty$$

Then $$\lim_{\varepsilon\to 0} \varepsilon \ln E\left[e^{\frac{\phi(X_\varepsilon)}{\varepsilon}}\right] = \sup\{\phi(x) - I(x) : x \in \mathcal{X}\}$$

<u>Remark.</u> This is a natural extension of the Laplace method to infinite dimensional spaces. Indeed let $\mathcal{X} = \mathbb{R}$ and suppose that μ_ε has a density $\frac{d\mu_\varepsilon}{dx} \approx e^{-\frac{I(x)}{\varepsilon}}$ with respect to the Lebesgue measure. Then

$$\int_{\mathbb{R}} e^{\frac{\phi(x)}{\varepsilon}} \mu_\varepsilon(dx) \simeq \int_{\mathbb{R}} e^{\frac{1}{\varepsilon}[\phi(x) - I(x)]} dx$$

If ϕ and I are twice differentiable and $\phi - I$ is concave with a unique minimum x^* then
$$\phi(x) - I(x) = \phi(x^*) - I(x^*) + \frac{(x-x^*)^2}{2}[\phi - I]''(\xi), \quad \xi \in [x^*, x].$$

Hence $$\int_{\mathbb{R}} e^{\frac{1}{\varepsilon}\phi(x)} d\mu_\varepsilon(x) \simeq e^{\frac{1}{\varepsilon}[\phi(x^*) - I(x^*)]} \int_{\mathbb{R}} e^{-\frac{1}{3}(x-x^*)^2 B(x)} dx$$

for some function $B \geq 0$. Then the Laplace method implies that
$$\overline{\lim_{\varepsilon}}\ \varepsilon \ln \int_{\mathbb{R}} e^{-\frac{1}{2}B(x)(x-x^*)^2} dx = 0.$$

The proof of Theorem 3.1 is a consequence of three lemmas. The first one relates the assumptions (i) and (ii)

Lemma 3.2 In the statement of Theorem 3.1, assumption (ii) implies (i).

<u>Proof</u> Set $Z_\varepsilon = \exp\left[\frac{1}{\varepsilon}(\Phi(X_\varepsilon) - M)\right]$ and let $\gamma > 1$. Then

$$e^{-\frac{M}{\varepsilon}} E\left[e^{\frac{1}{\varepsilon}\Phi(X_\varepsilon)} 1_{\{\Phi(X_\varepsilon) \geq M\}}\right] = E\left(Z_\varepsilon 1_{\{Z_\varepsilon \geq 1\}}\right) \leq E\left[(Z_\varepsilon)^\gamma\right]$$

$$\leq e^{-\frac{\gamma M}{\varepsilon}} E\left[e^{\gamma \frac{\Phi(X_\varepsilon)}{\varepsilon}}\right]$$

This implies $\overline{\lim_\varepsilon} \, \varepsilon \ln E\left[e^{\frac{\Phi(X_\varepsilon)}{\varepsilon}} 1_{\{\Phi(X_\varepsilon) \geq M\}}\right] \leq -(\gamma - 1)M + \overline{\lim_\varepsilon} \, \varepsilon \ln E\left[e^{\frac{\gamma\Phi(X_\varepsilon)}{\varepsilon}}\right]$

Letting $M \to \infty$, we conclude the proof of (i). □

Lemma 3.3 Let $\Phi : \mathcal{X} \to \mathbb{R}$ be l.s.c. and let (X_ε) satisfy the LDP lower bounds with $I : \mathcal{X} \to [0, +\infty]$ Then

$$\underline{\lim} \, \varepsilon \ln E\left[e^{\frac{\Phi(X_\varepsilon)}{\varepsilon}}\right] \geq \sup\{\Phi(x) - I(x) : x \in \mathcal{X}\}$$

<u>Proof</u> Fix $x \in \mathcal{X}$ and $\delta > 0$. Since Φ is l.s.c., there exists a neighborhood G of x such that $\inf_{y \in G} \Phi(y) \geq \Phi(x) - \delta$. Hence

$$\underline{\lim_\varepsilon} \, \varepsilon \ln E\left[e^{\frac{\Phi(X_\varepsilon)}{\varepsilon}}\right] \geq \underline{\lim_\varepsilon} \, \varepsilon \ln E\left[e^{\frac{\Phi(X_\varepsilon)}{\varepsilon}} 1_{\{X_\varepsilon \in G\}}\right]$$

$$\geq \inf_{y \in G} \Phi(y) + \underline{\lim_\varepsilon} \, \varepsilon \ln P(X_\varepsilon \in G)$$

$$\geq \inf_{y \in G} \Phi(y) - \inf_{y \in G} I(y) \geq \Phi(x) - \delta - I(x)$$

The proof is complete, as $\delta \to 0$ and $x \in \mathcal{X}$ is arbitrary □

Lemma 3.4 Let $\Phi : \mathcal{X} \to \mathbb{R}$ be u.s.c., let condition (i) in the statement of Theorem 3.1 be satisfied, and let the LDP upper estimates hold with the good rate function I. Then $\overline{\lim_\varepsilon} \, \varepsilon \ln E\left[e^{\frac{\Phi(X_\varepsilon)}{\varepsilon}}\right] \leq \sup\{\Phi(x) - I(x) : x \in \mathcal{X}\}$

<u>Proof:</u> Suppose at first that Φ is bounded from above, and let $\sup_{x\in\mathcal{X}} \Phi(x) \leq M < \infty$.

Then condition (i) holds trivially for Φ. Fix $\alpha < \infty$, and set

$K(\alpha) = \{x \in \mathcal{X} : I(x) \leq \alpha\}$. Then $K(\alpha)$ is compact.

Let $\delta > 0$; since I is l.s.c, Φ is u.s.c. and $\mathcal{X}$ is regular, for every $x \in K(\alpha)$

there exists a neighborhood A_x of x such that

$$\inf_{y\in \bar{A}_x} I(y) \geq I(x) - \delta \quad \text{and} \quad \sup_{x\in \bar{A}_x} \Phi(y) \leq \Phi(x) + \delta .$$

Extract finitely many sets A_{x_i}, $i = 1, \ldots, N$ from the covering

$K(\alpha) \subset \bigcup_{x\in K(\alpha)} A_x$. Then

$$E[e^{\frac{\Phi(X_\varepsilon)}{\varepsilon}}] \leq \sum_{i=1}^{N} E[e^{\frac{\Phi(X_\varepsilon)}{\varepsilon}} 1_{\{X_\varepsilon \in A_{x_i}\}}] + e^{\frac{M}{\varepsilon}} P(X_\varepsilon \in (\bigcup_{i=1}^{N} A_{x_i})^c)$$

$$\leq \sum_{i=1}^{N} e^{\frac{1}{\varepsilon}[\Phi(x_i)+\delta]} P(X_\varepsilon \in \bar{A}_{x_i}) + e^{\frac{M}{\varepsilon}} P(X_\varepsilon \in (\bigcup_{i=1}^{N} A_{x_i})^c).$$

Using the LDP upper bounds for $\bar{A}_{x_i}$, $i = 1, \ldots, N$ and $(\bigcup_{i=1}^{N} A_{x_i})^c$, as well as

Lemma 1.10 (i), we deduce

$$\overline{\lim_{\varepsilon}}\, \varepsilon \ln E[e^{\frac{\Phi(X_\varepsilon)}{\varepsilon}}] \leq \max\Big[\max_{i=1,\ldots,N} \big(\Phi(x_i) + \delta - \inf\{I(y) : y \in \bar{A}_{x_i}\}\big),\ M - \inf\{I(y) : y \in (\bigcup_{i=1}^{N} A_{x_i})^c\}\Big]$$

$$\leq \max\Big[\max_{i=1,\ldots,N} \big(\Phi(x_i) - I(x_i) + 2\delta\big),\ M - \alpha\Big)$$

$$\leq \max\big(\sup\{\Phi(x) - I(x) : x \in \mathcal{X}\},\ M - \alpha\big) + 2\delta .$$

Letting $\delta \to 0$ and $\alpha \to +\infty$, we conclude the proof.

We next suppose that Φ satisfies condition (i) in the statement of Theorem 3.1, and set

$\Phi_M(x) = \inf(\Phi(x), M) \leq \Phi(x)$. Since $\Phi_M \leq M$, splitting Ω into $\{\Phi(X_\varepsilon) > M\}$ and

its complement, using the above result for bounded Φ and Lemma 1.10 (i), we deduce

$$\overline{\lim_{\varepsilon}}\, \varepsilon \ln E[e^{\frac{\Phi(X_\varepsilon)}{\varepsilon}}] \leq \sup_{x\in\mathcal{X}} (\Phi(x) - I(x)) \vee \overline{\lim_{\varepsilon}}\, \varepsilon \ln E[e^{\frac{\Phi(X_\varepsilon)}{\varepsilon}} 1_{\{\Phi(X_\varepsilon) \geq M\}}].$$

Using condition (i), we conclude the proof as $M \to \infty$. □

3.2 Bryc's lemma

We next prove that if $\mathcal{X}$ is a completely regular topological space (that is, $\mathcal{X}$ is Hausdorff and given a closed set F and $x \in \mathcal{X}$ such that $x \notin F$, there exists a continuous map $f : \mathcal{X} \to [0,1]$ such that $f(x) = 1$ and $f(y) = 0$ for $y \in F$), then an inverse of Varadhan's lemma holds.

Let $\mathcal{C}_b(\mathcal{X})$ denote the set of continuous and bounded functions on $\mathcal{X}$.

Theorem 3.5 (Bryc) Let (μ_ε) be an exponentially tight family of probabilities on $\mathcal{B}(\mathcal{X})$, where $\mathcal{X}$ is a totally regular topological space. Suppose that for every $f \in \mathcal{C}_b(\mathcal{X})$,
$$\Lambda(f) := \lim_{\varepsilon \to 0} \varepsilon \ln \int_{\mathcal{X}} e^{\frac{f(x)}{\varepsilon}} \mu_\varepsilon(dx)$$
exists. Then (μ_ε) satisfies a LDP with the good rate function $I(x) = \sup\{ f(x) - \Lambda(f) : f \in \mathcal{C}_b(\mathcal{X})\}$.

Furthermore, for every $f \in \mathcal{C}_b(\mathcal{X})$, $\Lambda(f) = \sup\{ f(x) - I(x) : x \in \mathcal{X}\}$.

Remark In general, I is not the Fenchel-Legendre transform of Λ. Indeed, I is not necessarily convex.

Proof. Since $\Lambda(0) = 0$, $I \geq 0$. Since I is the supremum of continuous functions, it is l.s.c. Proposition 1.11 shows that the proof reduces to prove upper (resp lower) estimates for compact (resp open) sets.

Lemma 3.6 (Upper bounds for compact sets)

If $\Lambda(f)$ exists for every $f \in \mathcal{C}_b(\mathcal{X})$, then $\overline{\lim}\, \varepsilon \ln \mu_\varepsilon(K) \leq -\inf\{I(x) : x \in K\}$ for any compact subset K of $\mathcal{X}$.

Proof The argument is similar to that used in the proof of the Ellis-Gärtner Theorem 1.12.

Let K be a compact subset of $\mathcal{X}$, $\delta>0$, and set $I^\delta(x) = \inf\{I(x)-\delta, \frac{1}{\delta}\}$.

For $x\in K$, let $f_x \in \mathcal{C}_b(\mathcal{X})$ be such that $f_x(x) - \Lambda(f_x) \geq I^\delta(x)$.

Since f_x is continuous there exits a neighborhood $\mathcal{V}_x$ of x such that

$f_x(y) \geq f_x(x) - \delta$ for every $y \in \mathcal{V}_x$.

Then
$$\mu_\varepsilon(\mathcal{V}_x) \leq \int 1_{\mathcal{V}_x}(y)\, e^{\frac{1}{\varepsilon}[f_x(y) - f_x(x) + \delta]}\, d\mu_\varepsilon(y)$$
$$\leq e^{\frac{1}{\varepsilon}[-f_x(x)+\delta]} \int_{\mathcal{X}} e^{\frac{1}{\varepsilon} f_x(y)}\, d\mu_\varepsilon(y)$$

Hence $\overline{\lim_\varepsilon}\ \varepsilon \ln \mu_\varepsilon(\mathcal{V}_x) \leq \Lambda(f_x) - f_x(x) + \delta$

We extract a finite subcovering $\bigcup_{i=1}^N \mathcal{V}_{x_i} \supset K$ of $\bigcup_{x\in K} \mathcal{V}_x \supset K$. Using Lemma 1.10 (i), we deduce

$$\overline{\lim_\varepsilon}\ \varepsilon \ln \mu_\varepsilon(K) \leq \max_{1\leq i\leq N} \{\Lambda(f_{x_i}) - f_{x_i}(x_i)\} + \delta$$
$$\leq -\min_{x\in K}\{f_x(x) - \Lambda(f_x)\} + \delta \leq -\inf\{I^\delta(x) : x\in K\} + \delta$$

The proof is complete by letting $\delta \to 0$. $\square$

$\boxed{\text{Lemma 3.7}}$ (<u>lower bounds</u>) Let $\Lambda(f)$ exist for every $f\in\mathcal{C}_b(\mathcal{X})$. Then for any open subset G, $\underline{\lim}\ \varepsilon\, \mu_\varepsilon(G) \geq -\inf\{I(x) : x\in G\}$.

<u>Proof</u>. Let $G\neq\emptyset$ be an open subset of $\mathcal{X}$, and $x\in G$. The proof reduces to checking $\underline{\lim}\ \varepsilon \ln \mu_\varepsilon(G) \geq -I(x)$.

Since $\mathcal{X}$ is completely regular, there exists a continuous function $f : \mathcal{X}\to[0,1]$ such that $f(x)=1$ and $f(y)=0$ for $y\in G^c$. For $M>0$ set

$f_M(y) = M[f(y)-1]$. Then

$$\int_{\mathcal{X}} e^{\frac{f_M(y)}{\varepsilon}} \mu_\varepsilon(dy) \leq e^{-\frac{M}{\varepsilon}} \mu_\varepsilon(G^c) + \mu_\varepsilon(G) \leq e^{-\frac{M}{\varepsilon}} + \mu_\varepsilon(G)$$

We have $f_M \in \mathcal{C}_b(\mathcal{X})$ and $f_M(x) = 0$. Hence Lemma 2.3 implies

$$\max\left(\underline{\lim}_{\varepsilon} \ \varepsilon \ln \mu_\varepsilon(G), -M \right) \geq \underline{\lim}_{\varepsilon} \ \varepsilon \ln \int_{\mathcal{X}} e^{\frac{f_M(y)}{\varepsilon}} \mu_\varepsilon(dy) = \Lambda(f_M)$$

$$\geq -\left[f_M(x) - \Lambda(f_M) \right] \geq -\sup\{ f(x) - \Lambda(f) : f \in \mathcal{C}_b(\mathcal{X}) \} = -I(x)$$

For $M > I(x)$, the max in the left hand side is $\underline{\lim}\ \varepsilon \ln \mu_\varepsilon(G) \geq -I(x)$. $\square$

<u>Remark</u> We deduce from the Varadhan and Bryc lemmas the following characterization of a LDP for a family (μ_ε) of probabilities on a totally regular topological space $\mathcal{X}$, such as a Polish [complete, separable, metric] space.

$\boxed{\text{Theorem 3.8}}$ Let (X^ε) be a family of random variables taking values in a Polish space $\mathcal{X}$ and let $I : \mathcal{X} \to [0, +\infty]$ be a good rate function. Then the family of distributions of (X^ε) satisfies a LDP with the good rate function I if and only if for every continuous and bounded function $f : \mathcal{X} \to \mathbb{R}$,

$$\lim_{\varepsilon \to 0} -\varepsilon \ln E\left[\exp\left(-\frac{1}{\varepsilon} f(X^\varepsilon)\right)\right] = \inf\{ f(x) + I(x) : x \in \mathcal{X} \}$$

3.3 The weak convergence approach of LDP

Let $(\mathcal{X},\mathcal{F})$ be a measurable space, $\phi:\mathcal{X}\to\mathbb{R}$ be a bounded measurable function.

Then for any probability μ on $\mathcal{F}$, we have the following variational formula

$$\ln\int_{\mathcal{X}} e^{\phi}\,d\mu = \sup\left\{\int_{\mathcal{X}}\phi f\,d\mu - \int_{\mathcal{X}} f\ln f\,d\mu : f\in L^1(\mu),\ f\geq 0\ \mu \text{ a.e. and } \int f\,d\mu=1\right\}$$

Furthermore, the supremum is uniquely achieved for $f(x)=\dfrac{e^{\phi(x)}}{\int_{\mathcal{X}} e^{\phi(x)}\mu(dx)}$

(The proof can be read in [DZ], (6.2.15) page 264.)

Let $H(\nu|\mu)=\int f\ln f\,d\mu$ if $\frac{d\nu}{d\mu}=f$, $H(\nu|\mu)=+\infty$ if ν is not AC with respect to μ.

This variational can be rephrased: For $\phi:\mathcal{X}\to\mathbb{R}$ bounded and measurable

$$-\ln\int_{\mathcal{X}} e^{-\phi(x)}\mu(dx) = \inf\left\{H(\nu|\mu)+\int\phi(x)\nu(dx) : \nu \text{ probability on } \mathcal{F}\right\}.$$

We then deduce a representation theorem for a standard r-dimensional Brownian motion W. Let $\Omega=\mathcal{C}([0,T];\mathbb{R}^r)$ be endowed with its Borel σ-field and let μ be the r-dimensional Wiener measure.

Definition 3.9 The set $\mathcal{A}$ of predictable controls is the set of predictable processes $v:[0,T]\times\Omega\to\mathbb{R}^r$ such that $\int_0^T\|v_t\|_r^2\,dt<\infty$ μ a.s.

$\boxed{\textbf{Theorem 3.10}}$ (Boué-Dupuis) Let $f:[0,T]\times\mathbb{R}^d\to\mathbb{R}$ be a bounded Borel map; then

$$-\ln E\left[e^{-f(W)}\right]=\inf\left\{E\left[\frac{1}{2}\int_0^T\|v_s\|_{\mathbb{R}^r}^2\,ds + f\left(W_\cdot+\int_0^\cdot v_s\,ds\right)\right] : v\in\mathcal{A}\right\}$$

__Proof:__ Let v be a bounded element of $\mathcal{A}$. Then

$M_t=\exp\left(\sum_{i=1}^d\int_0^t v_s^{(i)}\,dW_s^i-\frac{1}{2}\int_0^t\|v_s\|_d^2\,ds\right)$, $t\in[0,T]$ is a martingale

and Girsanov's theorem implies that $\tilde{W}_t=W_t-\int_0^t v_s\,ds$ is a d dimensional Brownian motion under the probability $\tilde{P}$ such that $\frac{d\tilde{P}}{dP}=M_T$ on $\mathcal{F}_T$.

Then $H(\tilde{P}|P)=\int_\Omega\ln\left(\frac{d\tilde{P}}{dP}\right)\frac{d\tilde{P}}{dP}\,dP=\int_\Omega\ln\left(\frac{d\tilde{P}}{dP}\right)d\tilde{P}$

$$= \int_\Omega \Big(\sum_{i=1}^{d} V_s^{(i)} dW_s^i - \frac{1}{2} \int \|V_s\|_d^2 ds \Big) d\tilde{P}$$

$$= \int_\Omega \Big(\sum_{i=1}^{d} \int_0^T V_s^{(i)} d\tilde{W}_s^i + \sum_{i=1}^{d} \int_0^T V_s^{(i)} V_s^{(i)} ds - \frac{1}{2} \int_0^T \|V_s\|_d^2 ds \Big) d\tilde{P}$$

$$= \frac{1}{2} \int_\Omega \int_0^T \|V_s\|_d^2 ds \, d\tilde{P}$$

since the stochastic integral is centered.

Therefore $H(\tilde{P}|P) + \int_\Omega f(\omega) d\tilde{P}(\omega) = \int_\Omega \Big[\frac{1}{2} \int_0^T \|V_s\|_d^2 ds + f\big(\tilde{W} + \int_0^{\cdot} V_s ds\big) \Big] d\tilde{P}$

The variational formula implies

$$-\ln E[e^{-f(\omega)}] \leq \inf\Big\{ \int_\Omega \Big[\frac{1}{2} \int_0^T \|V_s\|_d^2 ds + f\big(\tilde{W} + \int_0^{\cdot} V_s ds\big) \Big] d\tilde{P} \ : \ V \in \mathcal{A}_b \Big\}$$

We next have to extend this inequality from $V \in \mathcal{A}_b$ to $V \in \mathcal{A}$, and replace $\tilde{W}$ (resp $\tilde{P}$) by W (resp. P).

Step 1 Suppose that V is a simple process, that is

$V = \xi_0 1_{\{0\}} + \sum_{i=0}^{j-1} \xi_i 1_{(t_i, t_{i+1}]}$, where ξ_i are $\mathcal{F}_{t_i}$-measurable and $t_0 = 0 < t_1 < \dots < t_j \leq T$.

Let $T_V(\omega)$ be defined by $(T_V(\omega))_t = \omega_t - \int_0^t V_s(\omega) ds$.

Then $\tilde{V} = T_V$ is again a simple process defined by

$\tilde{V}_t(\omega) = \xi_0(\omega) 1_{\{0\}}(t) + \sum_{i=0}^{j-1} \tilde{\xi}_i(\omega) 1_{(t_i, t_{i+1}]}(t)$, and $\tilde{\xi}_0 = \xi_0$, while for $i = 1, \dots, j-1$,

$\tilde{\xi}_i$ are defined by induction: $\tilde{\xi}_i(\omega) = \xi_i(\varphi_i(\omega))$, $\varphi_i(\omega) = W_{t_i} - \sum_{k=0}^{i-1} \tilde{\xi}_k (t_{k+1} - t_k)$.

Then $\tilde{V}$ is a $\tilde{P}$ Brownian motion, and the distribution of $(\tilde{W}, \tilde{v})$ under $\tilde{P}$ coincides with that of (W, v) under P. Therefore, since $\tilde{v} \in \mathcal{A}_b$,

$$-\ln E[e^{-f(W)}] \leq \int_\Omega \Big[\frac{1}{2} \int_0^T \|\tilde{V}_s\|^2 ds + f\big(\tilde{W} + \int_0^{\cdot} \tilde{V}_s ds\big) \Big] d\tilde{P}$$

$$\leq \int_\Omega \Big[\frac{1}{2} \int_0^T \|V_s\|^2 ds + f\big(W + \int_0^{\cdot} V_s ds\big) \Big] dP$$

Hence $-\ln E[e^{-f(W)}] \leq +\inf\Big\{ \int_\Omega \Big[\frac{1}{2} \int_0^T \|V_s\|^2 ds + f\big(W_\cdot + \int_0^{\cdot} V_s ds\big) \Big] dP \ ; \ V \in \mathcal{A}_b \Big\}$.

Step 2 Suppose that $V \in \mathcal{A}$ is bounded, and P a.s., $\|V_s(\omega)\| \leq M$.

Then there exist a sequence (V_n) of simple processes in $\mathcal{A}_b$, such that $\|V^n_s\| \leq M$ a.s. and $E\int_0^T \|V^n_s - V_s\|^2 ds \to 0$ as $n \to \infty$. Then $(W, \int_0^\cdot V^n_s ds)$ converges in distribution to $(W, \int_0^\cdot V_s ds)$, and for every n,

$$-\ln E[e^{-f(W)}] \leq E\left[\frac{1}{2}\int_0^T \|V^n_s\|^2 ds + f(W_\cdot + \int_0^\cdot V^n_s ds)\right]$$

Note that $\sup_n E\left[\frac{1}{2}\int_0^T \|V^n_s\|^2 ds\right] \leq \frac{M^2 T}{2}$ and $\lim_n E\left[f(W + \int_0^\cdot V^n_s ds)\right] = E\left[f(W + \int_0^\cdot V_s ds)\right]$

Furthermore, $E\left[\frac{1}{2}\int_0^T \|V^n_s\|^2 ds\right] = H(\tilde{P}_n | P)$, where $\frac{d\tilde{P}_n}{dP} = \exp\left(\int_0^T (V^n_s, dW_s) - \frac{1}{2}\int_0^T \|V^n_s\|^2 ds\right)$.

Since $H(\cdot | P)$ is l.s.c, we deduce $H(\tilde{P} | P) \leq E\left[\frac{1}{2}\int_0^T \|V_s\|^2 ds\right]$ for $V \in \mathcal{A}_b$.

<u>Step 3</u> Suppose $V \in \mathcal{A}$ and let $V^n_s = V_s 1_{\{\|V_s\| \leq n\}}$ for $n \geq 1$, $s \in [0,T]$.

Then $\sup_n H(P_n | P) \leq \sup_n E\left(\frac{1}{2}\int_0^T \|V^n_s\|^2 ds\right) \leq \frac{1}{2} E\left(\int_0^T \|V_s\|^2 ds\right) < \infty$.

The dominated convergence theorem concludes the proof of the upper bound

$$-\ln E[e^{-f(W)}] \leq \inf\left\{E\left[\frac{1}{2}\int_0^T \|V_s\|^2 ds + f(W_\cdot + \int_0^\cdot V_s ds)\right] : V \in \mathcal{A}\right\}$$

<u>We next prove the lower bound.</u> Let γ_0 be the probability which minimizes the variational formula, that is $\frac{d\gamma_0}{dP}(w) = e^{-f(w)} \times \frac{1}{E[e^{-f(W)}]}$. Then γ_0 and P are equivalent. Let $R_t = E\left[\frac{d\gamma_0}{dP} | \mathcal{F}_t\right]$. Then $(R_t, \mathcal{F}_t, t \in [0,T])$ is a martingale which is bounded from below by $e^{-2\|f\|_\infty}$, and from above by $e^{2\|f\|_\infty}$.

This martingale can be expressed as a stochastic integral: $R_t = 1 + \int_0^t u_s dW_s$
$= 1 + \int_0^t \tilde{V}_s R_s dW_s$, and $E\int_0^T \|\tilde{V}_s\|^2 R_s^2 ds < \infty$. Since $R_t \geq e^{-2\|f\|_\infty}$, we deduce that $E\int_0^T \|\tilde{V}_s\|^2 ds < \infty$, and also $\int_\Omega \int_0^T \|\tilde{V}_s\|^2 ds\, d\gamma_0 < \infty$

Then $R_t = \exp\left(\int_0^T \tilde{V}_s dW_s - \frac{1}{2}\int_0^T \|\tilde{V}_s\|^2 ds\right)$ and $\tilde{W}_t = W_t - \int_0^t \tilde{V}_s ds$ is a γ_0 Brownian motion. Hence, since $H(\gamma_0 | P) = E\left(\frac{1}{2}\int_0^T \|\tilde{V}_s\|^2 ds\right)$, we deduce

$$-\ln E[e^{-f(W)}] = \int_\Omega \Big[\frac{1}{2}\int_0^T \|\tilde v_s\|^2 ds + f(W)\Big] d\tilde P = \int_\Omega \Big[\frac{1}{2}\int_0^T \|\tilde v_s\|^2 ds + f\big(\tilde W + \int_0^{\cdot} \tilde v_s ds\big)\Big] d\tilde P$$

(1) Assume that f is continuous

Then one can construct a sequence $(\tilde v^n)$ of bounded simple processes in $\mathcal{A}$ such that $\int_\Omega \int_0^T \|\tilde v^n_s - \tilde v_s\|^2 ds\, d\tilde P \to 0$ as $n \to \infty$. Thus given $\varepsilon > 0$, we may choose n and C such that $\|\tilde v^n_s\| \le C$ a.s for $s \in [0,T]$,

and $\int_\Omega \Big[\frac{1}{2}\int_0^T \|\tilde v_s\|^2 ds + f\big(\tilde W + \int_0^{\cdot} \tilde v_s ds\big)\Big] d\tilde P \geq \int_\Omega \Big[\frac{1}{2}\int_0^T \|\tilde v^n_s\|^2 ds + f\big(\tilde W + \int_0^{\cdot} \tilde v^n_s ds\big)\Big] d\tilde P - \varepsilon$

Then $\tilde v^n_t = \xi_0 1_{\{0\}}(t) + \sum_{i=0}^{l-1} \xi_i 1_{(t_i, t_{i+1}]}(t)$, where $t_0 = 0 < t_1 < \dots < t_l = T$,

and ξ_i is $\mathcal{F}_{t_i}$-measurable. For $i = 0, \dots, l-1$, set

$$\tilde Z^{(i)}_s = \tilde W_{t_{i-1}+s} - \tilde W_{t_{i-1}} \quad \text{for } s \in [0, t_i - t_{i-1}]$$

The continuity of f implies the existence of a continuous function

$F: (\mathbb{R}^d)^l \times \prod_{j=1}^{l} \mathcal{C}([0, t_j - t_{j-1}]; \mathbb{R}^d)$ such that for $\xi = (\xi_0, \dots, \xi_{l-1})$, $\tilde Z^i = (\tilde Z^{(1)}, \dots, \tilde Z^{(i)})$,

$$f\big(\tilde W + \int_0^{\cdot} \tilde v^n_s ds\big) = F(\xi, \tilde Z^l).$$

Furthermore there exists a function $G: (\mathbb{R}^d)^l \to \mathbb{R}$ such that $\frac{1}{2}\int_0^T \|\tilde v^n_s\|^2 ds = G(\xi_0, \dots, \xi_{l-1})$.

Hence $\int_\Omega \Big[\frac{1}{2}\int_0^T \|\tilde v_s\|^2 ds + f\big(\tilde W + \int_0^{\cdot} \tilde v_s ds\big)\Big] d\tilde P \geq \int_\Omega \big[G(\xi) + F(\xi, \tilde Z^l)\big] d\tilde P - \varepsilon$.

Each random variable ξ_i can be written as a bounded measurable function of $\tilde Z^{(j)}$, $j < i$. Using an approximation of this function by continuous ones we can construct a simple process

$\bar v_t = \Gamma_0 1_{[0,t_1]}(t) + \sum_{i=1}^{l-1} \Gamma_i(Z^i) 1_{(t_i, t_{i+1}]}(t)$, $0 \le t \le T$, where $Z^i = (Z^{(1)}, \dots, Z^{(i)})$,

$Z^{(i)}_s = W_{t_{i-1}+s} - W_{t_{i-1}}$, $s \in [0, t_i - t_{i-1}]$, $\Gamma_0 \in \mathbb{R}^d$, $\Gamma_i : \prod_{j=1}^{i} \mathcal{C}([0, t_j - t_{j-1}]; \mathbb{R}^d) \to \mathbb{R}^d$ is continuous, $\|\Gamma_j\| \le C$, and $\int_\Omega \big[G(\xi) + F(\xi, \tilde Z^l)\big] d\tilde P \geq \int_\Omega \big[G(\Gamma(\tilde Z^l)) + F(\Gamma(\tilde Z^l), \tilde Z^l)\big] dP - \varepsilon$.

Therefore, $-\ln E[e^{-f(W)}] = \int_\Omega \left[\frac{1}{2}\int_0^T \|\tilde{v}_s\|^2 ds + f(\tilde{W} + \int_0^{\cdot} \tilde{v}_s ds)\right] d\tilde{P}$

$$\geq \int_\Omega \left[\frac{1}{2}\int_0^T \|\bar{v}_s\|^2 ds + f(W + \int_0^{\cdot} \bar{v}_s ds)\right] dP - 2\varepsilon$$

As $\varepsilon \to 0$, this proves the lower bounds for a continuous function f.

(2) <u>If f is not continuous</u>, let (f_n) be a sequence of continuous bounded functions on $\Omega = \mathcal{C}([0,T]; \mathbb{R}^d)$, such that $f_n \to f$ P a.s. and $\sup_n \|f_n\|_\infty \leq \|f\|_\infty$.

The result above applied to each function f_n implies for every $n \geq 1$

$$-\ln E(e^{-f_n(W)}) \geq E\left[\frac{1}{2}\int_0^T \|\bar{v}^n_s\|^2 ds + f_n(W + \int_0^{\cdot} \bar{v}^n_s ds)\right]$$

Furthermore, $\sup_n H(\tilde{P}_n | P) = \sup_n \frac{1}{2} E(\int_0^T \|\bar{v}^n_s\|^2 ds) \leq 2\|f_n\|_\infty \leq 2\|f\|_\infty$

This implies that the pair $(W, \int_0^{\cdot} \bar{v}^n_s ds)$ is tight. Hence there exists a subsequence $(n_j, j \geq 0)$ such that $(W, \int_0^{\cdot} \bar{v}^{n_j}_s ds)$ converges in distribution to $(W, \int_0^{\cdot} \bar{v}_s ds)$.

We next prove that $E[e^{-f_n(W)}] \to E[e^{-f(W)}]$ and

$$\left|E[f_n(W + \int_0^{\cdot} \bar{v}^n_s ds)] - E[f(W + \int_0^{\cdot} \bar{v}^n_s ds)\right| \to 0 \text{ as } n \to \infty$$

Since $\|f_n\|_\infty \leq \|f\|_\infty$ and $f_n \to f$ P a.s, the dominated convergence theorem proves the first convergence.

To prove the second one, let $\tilde{P}_n$ be the distribution of $W + \int_0^{\cdot} \bar{v}^n_s ds$.

We prove that $\lim_{j\to\infty} \sup_{n \in \mathbb{N}} \int \|f_j - f\| d\tilde{P}_n = 0$. This will imply

$$\left|\int f_j \, d\tilde{P}_j - \int f \, d\tilde{P}_j\right| \leq \int \|f_j - f\| \, d\tilde{P}_j \leq \alpha \text{ for large } j.$$

Given $\varepsilon > 0$,

$$\int \|f_j - f\| d\tilde{P}_n \leq \int_{\{\|f_j - f\|_\infty > \varepsilon\}} \|f_j - f\|_\infty \, d\tilde{P}_n + \varepsilon$$

$\leq 2\|f\|_\infty \, \tilde{P}_n(\|f_j - f\|_\infty \geq \varepsilon) + \varepsilon$

Thus, it is enough to prove that for any $\varepsilon > 0$ $\lim_j \sup_{n\in\mathbb{N}} \tilde{P}_n(\|f_j - f\|_\infty \geq \varepsilon) = 0$

For any $c \in (1,\infty)$, every $n \in \mathbb{N}$ and $j \in \mathbb{N}$,

$$\tilde{P}_n(\|f_j - f\|_\infty > \varepsilon) = \int 1_{\{\|f_j - f\|_\infty \geq \varepsilon\}} \frac{d\tilde{P}_n}{dP}\, dP$$

$$\leq \int 1_{\{\|f_j - f\|_\infty > \varepsilon\} \cap \{\frac{d\tilde{P}_n}{dP} > c\}} \frac{d\tilde{P}_n}{dP}\, dP + c\, P(\|f_j - f\|_\infty > \varepsilon).$$

For fixed $\varepsilon > 0$, $c > 1$, one can choose j_0 such that $c\, P(\|f_j - f\|_\infty > \varepsilon) < \varepsilon$ for $j \geq j_0$.

Furthermore,

$$\int 1_{\{\|f_j - f\|_\infty > \varepsilon\} \cap \{\frac{d\tilde{P}_n}{dP} > c\}} \frac{d\tilde{P}_n}{dP}\, dP \leq \frac{1}{\ln(c)} \int 1_{\{\|f_j - f\|_\infty > \varepsilon\} \cap \{\frac{d\tilde{P}_n}{dP} > c\}} \ln\left(\frac{d\tilde{P}_n}{dP}\right) \frac{d\tilde{P}_n}{dP}\, dP$$

$$\leq \frac{1}{\ln(c)} \int \left[e^{-1} + \frac{d\tilde{P}_n}{dP} \ln\left(\frac{d\tilde{P}_n}{dP}\right)\right] \leq \frac{1}{\ln(c)} \left[e^{-1} + H(\tilde{P}_n|P)\right] \leq \frac{\frac{1}{e} + 2\|f\|_\infty}{\ln(c)}$$

Thus, given $\varepsilon > 0$ we may choose $c > 1$ large enough to have $\dfrac{\frac{1}{e} + 2\|f\|_\infty}{\ln(c)} \leq \varepsilon$.

Since $\bar{v}^n \in \mathcal{A}$, we have for n large enough

$$-E(e^{-f(W)}) \geq E\left[\frac{1}{2}\int_0^T \|\bar{v}^n_s\|^2 ds + f\left(W + \int_0^\cdot \bar{v}^n_s ds\right)\right] - \varepsilon$$

$$\geq \inf\left\{E\left[\frac{1}{2}\int_0^T \|v_s\|^2 ds + f\left(W + \int_0^\cdot v_s ds\right)\right] : v \in \mathcal{A}\right\} - \varepsilon.$$

This completes the proof as $\varepsilon \to 0$. $\square$

The above result can be extended to an H-valued Wiener process, where H is a Hilbert space. Recall that, as stated in Section 1.7, Definition 1.27 we define a Q-Wiener process on H, where Q is trace class, and set $H_0 = Q^{\frac{1}{2}}H$ with the scalar product $\langle h, k\rangle_{H_0} = (Q^{-\frac{1}{2}}h, Q^{-\frac{1}{2}}k)_H$.

The following result (Budhiraja-Dupuis) extends Theorem 3.10 to infinite dimensional settings.

Theorem 3.11 Let f be a Borel bounded function from $\mathcal{C}([0,T];H)$ to $\mathbb{R}$. Then

$$\ln E[e^{-f(W)}] = -\inf\left\{E\left[\frac{1}{2}\int_0^T \|v(s)\|_0^2\,ds + f\left(W + \int_0^\cdot v(s)\,ds\right)\right] : v \in \mathcal{A}\right\},$$

where $\mathcal{A}$ is the class of H_0-valued predictable processes v such that $\int_0^T \|v(s)\|_0^2\,ds < \infty$ a.s.

We next give sufficient condition to obtain LDP for the family of distributions of the random variables $X^\varepsilon = g^\varepsilon(W)$ as $\varepsilon \to 0$. It is based on the representation Theorem 3.11 and on the equivalence stated in Theorem 3.8

For every $M > 0$, set

$S_M = \{v \in L^2(0,T;H_0) : \int_0^T \|v(s)\|_0^2\,ds \le M\}$ endowed with the weak topology,

$\mathcal{A}_M = \{v \in \mathcal{A} : v(\omega) \in S_M \text{ a.s.}\}$.

Assumption 3.12 For $\varepsilon > 0$, let $g^\varepsilon : \mathcal{C}([0,T];H) \to \mathcal{E}$ be a measurable map, where $\mathcal{E}$ is a Polish space. Let $g^0 : C([0,T];H) \to \mathcal{E}$ be a measurable map.

Suppose that

(i) For $M > 0$ and any family $\{v^\varepsilon\} \subset \mathcal{A}_M$, if v^ε converges to $v \in \mathcal{A}_M$ in distribution (as random elements of S_M), then $g^\varepsilon\left(W + \frac{1}{\sqrt{\varepsilon}}\int_0^\cdot v^\varepsilon(s)\,ds\right)$ converges in distribution to $g^0\left(\int_0^\cdot v(s)\,ds\right)$.

(ii) For every $M > 0$, $\Gamma_M = \{g^0(\int_0^\cdot v(s)\,ds) : v \in S_M\}$ is a compact subset of $\mathcal{E}$.

Given $u \in \mathcal{E}$, set $I(u) = \inf\left\{\frac{1}{2}\int_0^T \|v(s)\|_0^2\,ds : v \in L^2(0,T;H_0),\ g^0\left(\int_0^\cdot v(s)\,ds\right) = u\right\}$.

Theorem 3.13 Let $X^\varepsilon = g^\varepsilon(W)$, $\varepsilon > 0$ where g^ε satisfies Assumption 3.12. Then the family of distributions of X^ε satisfies on $\mathcal{E}$ a LDP with the good rate function I. (Budhiraja - Dupuis).

<u>Proof</u> To use Theorem 3.8, we prove $\lim_{\varepsilon\to 0} \varepsilon \ln E[e^{-\frac{1}{\varepsilon}h(X^\varepsilon)}] = -\inf\{h(x)+I(x) : x\in\mathcal{E}\}$

for any bounded continuous function $h : \mathcal{E}\to\mathbb{R}$.

<u>Proof of the lower bound</u> Using Theorem 3.11, we have

$$-\varepsilon \ln E[e^{-\frac{1}{\varepsilon}h(X^\varepsilon)}] = \inf\Big\{ E\Big[\frac{\varepsilon}{2}\int_0^T \|v(s)\|_0^2 ds + h\circ g^\varepsilon\Big(W+\int_0^\cdot v(s)ds\Big)\Big] : v\in\mathcal{A}\Big\}$$

$$= \inf\Big\{ E\Big[\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + h\circ g^\varepsilon\Big(W+\frac{1}{\sqrt{\varepsilon}}\int_0^\cdot v(s)ds\Big)\Big] : v\in\mathcal{A}\Big\}.$$

Fix $\delta>0$. For every $\varepsilon>0$ there exists $v^\varepsilon\in\mathcal{A}$ such that

$$\inf\Big\{E\Big[\frac{1}{2}\int_0^T\|v(s)\|_0^2 + h\circ g^\varepsilon\Big(W+\frac{1}{\sqrt\varepsilon}\int_0^\cdot v(s)ds\Big)\Big] : v\in\mathcal{A}\Big\} \geq E\Big[\frac{1}{2}\int_0^T\|v^\varepsilon(s)\|_0^2 ds + h\circ g^\varepsilon\Big(W+\frac{1}{\sqrt\varepsilon}\int_0^\cdot v^\varepsilon(s)ds\Big)\Big]-\delta.$$

We prove that $\varliminf_{\varepsilon\to 0} E\Big[\frac{1}{2}\int_0^T \|v^\varepsilon(s)\|_0^2 ds + h\circ g^\varepsilon\Big(W+\int_0^\cdot \frac{v^\varepsilon(s)}{\sqrt\varepsilon}ds\Big)\Big] \geq \inf\{h(x)+I(x) : x\in\mathcal{E}\}$.

Let $K=\|h\|_\infty$; then $\sup_{\varepsilon>0} E\Big[\frac{1}{2}\int_0^T\|v^\varepsilon(s)\|_0^2 ds\Big]\leq 2(K+\delta)$.

Given $M>0$, set $\tau_M^\varepsilon = \inf\{t\geq 0 : \int_0^t \|v^\varepsilon(s)\|_0^2 ds \geq M\}\wedge T$. Then τ_M^ε is a stopping time; set $v^{\varepsilon,M}(s) = v^\varepsilon(s)\, 1_{[0,\tau_M^\varepsilon]}(s)$. The processes $v^{\varepsilon,M}$ belong to $\mathcal{A}_M$, and $P(v^\varepsilon\neq v^{\varepsilon,M}) = P\Big(\int_0^T \|v^\varepsilon(s)\|_0^2 ds \geq M\Big)\leq \frac{4(K+\delta)}{M}$. Hence,

$$E\Big[\frac{1}{2}\int_0^T\|v^\varepsilon(s)\|_0^2 ds + h\circ g^\varepsilon\Big(W+\frac{1}{\sqrt\varepsilon}\int_0^\cdot v^\varepsilon(s)ds\Big)\Big] \geq E\Big[\frac{1}{2}\int_0^T\|v^{\varepsilon,M}(s)\|_0^2 ds + h\circ g^\varepsilon\Big(W+\frac{1}{\sqrt\varepsilon}\int_0^\cdot v^{\varepsilon,M}(s)\Big)1_{\{v^{\varepsilon,M}=v^\varepsilon\}}\Big]$$
$$-\|h\|_\infty P(v^\varepsilon\neq v^{\varepsilon,M}).$$

Hence

$$\inf\Big\{E\Big[\frac{1}{2}\int_0^T\|v(s)\|_0^2 ds + h\circ g^\varepsilon\Big(W+\frac{1}{\sqrt\varepsilon}\int_0^\cdot v(s)ds\Big)\Big] : h\in\mathcal{A}\Big\} \geq E\Big[\frac{1}{2}\int_0^T\|v^{\varepsilon,M}(s)\|_0^2 + h\circ g^\varepsilon\Big(W+\frac{1}{\sqrt\varepsilon}\int_0^\cdot v^{\varepsilon,M}(s)ds\Big)\Big]$$
$$-\frac{4K(K+\delta)}{M}.$$

For fixed M, since $L^2(0,T;H_0)$ is endowed with the weak topology, a.s. $v^{\varepsilon,M}$ belongs to a compact subset of $L^2(0,T;H_0)$. Let $(v^{\tilde\varepsilon,M}, \tilde\varepsilon>0)$ denote a subfamily converging in distribution to $\tilde v\in\mathcal{A}_M$ as $\tilde\varepsilon\to 0$. Skorokhod's theorem yields the existence of $\bar v^{\varepsilon,M}$ and $\bar v$ on $(\bar\Omega,\bar{\mathcal{F}},\bar P)$ with the same distribution as $v^{\tilde\varepsilon,M}$ and $\tilde v$ under P, such that $\bar v^{\varepsilon,M}\to\bar v$ $\bar P$ a.s. as $\varepsilon\to 0$.

Fatou's lemma implies that

$$\underline{\lim}_{\varepsilon} E\left(\int_0^T \|v_s^{\tilde{\varepsilon},M}\|_{H_0}^2 ds\right) = \underline{\lim}_{\varepsilon} \bar{E}\left(\int_0^T \|\bar{v}_s^{\varepsilon,M}\|_0^2 ds\right) \geq \bar{E}\left(\int_0^T \|\bar{v}_s\|_0^2 ds\right) = E\left(\int_0^T \|\tilde{v}_s\|_0^2 ds\right)$$

Furthermore since h is continuous, Assumption 3.12 (i) implies that

$$E\left(h\circ g^{\varepsilon}\left(W + \frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} v^{\tilde{\varepsilon},M}(s)\, ds\right)\right) \longrightarrow E\left(\int_0^T g^0\left(\int_0^{\cdot} \tilde{v}(s) ds\right)\right) \text{ as } \varepsilon \to 0.$$

Hence

$$\underline{\lim}_{\varepsilon} E\left[\frac{1}{2}\int_0^T \|v^{\tilde{\varepsilon},M}(s)\|_0^2 ds + h\circ g^{\varepsilon}\left(W + \frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} v^{\tilde{\varepsilon},M}(s) ds\right)\right] \geq E\left[\frac{1}{2}\int_0^T \|\tilde{v}_s\|_0^2 ds + h\circ g^0\left(\int_0^{\cdot} \tilde{v}_s ds\right)\right]$$

$$\geq \inf\left\{\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + h(x) : x \in \mathcal{E},\ v \in L^2(0,T;H_0),\quad x = g^0\left(\int_0^{\cdot} v_s ds\right)\right\}$$

$$\geq \inf\{ I(x) + h(x) : x \in \mathcal{E}\}$$

Letting $\delta \to 0$ and $M \to \infty$, we conclude the proof of the lower bound.

<u>Proof of the upper bound</u> Since h is bounded, $\inf\{I(x) + h(x) : x \in \mathcal{E}\} < \infty$.

Fix $\delta > 0$ and let $x_0 \in \mathcal{E}$ be such that $I(x_0) + h(x_0) \leq \inf\{I(x) + h(x) : x \in \mathcal{E}\} + \delta$.

Then choose $\tilde{v} \in L^2(0,T;H_0)$ such that $x_0 = g^0\left(\int_0^{\cdot} \tilde{v}(s) ds\right)$ and $\frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2 ds \leq I(x_0) + \delta$.

Theorem 3.10 implies that for every continuous and bounded function h,

$$\overline{\lim}_{\varepsilon} -\varepsilon \ln E\left[e^{-\frac{1}{\varepsilon}h(X^{\varepsilon})}\right] = \overline{\lim}_{\varepsilon} \inf_{v \in \mathcal{A}} E\left[\frac{\varepsilon}{2}\int_0^T \|v(s)\|_0^2 ds + \varepsilon \times \frac{1}{\varepsilon} h\circ g^{\varepsilon}\left(W + \int_0^{\cdot} v(s) ds\right)\right]$$

$$\leq \overline{\lim}_{\varepsilon} E\left[\frac{\varepsilon}{2}\int_0^T \left\|\frac{\tilde{v}(s)}{\sqrt{\varepsilon}}\right\|_0^2 ds + h\circ g^{\varepsilon}\left(W + \frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} \tilde{v}(s) ds\right)\right]$$

$$\leq \frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2 ds + h\circ g^0\left(\int_0^{\cdot} \tilde{v}(s) ds\right) = \frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2 ds + h(x_0),$$

where the last line follows from Assumption 3.12 (i). Hence

$$\overline{\lim}_{\varepsilon} -\varepsilon \ln E\left[e^{-\frac{1}{\varepsilon}h(X^{\varepsilon})}\right] \leq I(x_0) + h(x_0) + \delta \leq \inf\{I(x) + h(x) : x \in \mathcal{E}\} + 2\delta.$$

Letting $\delta \to 0$, we conclude the proof of the upper estimate.

In order to use Theorem 3.8, we still have to prove that <u>I has compact level sets</u>.

Fix $M>0$ and let $x \in \mathcal{E}$ be such that $I(x) \leq M$. For $\delta>0$, there exists $v \in L^2(0,T;H_0)$ such that $x = g^0\left(\int_0^{\cdot} v(s)\,ds\right)$ and $\frac{1}{2}\int_0^T \|v(s)\|_0^2 \leq I(x)+\delta$.

Hence $\{I \leq M\} \subset \bigcap_{\delta>0} g^0(S_{M+\delta})$. Conversely, if $x \in g^0(S_{M+\delta})$ for $\delta>0$, $I(x) \leq M+\delta$. Hence $\{I \leq M\} = \bigcap_{\delta>0} g^0(S_{M+\delta})$. Assumption 3.12 (ii) implies that each set $g^0(S_{M+\delta})$ is compact. This concludes the proof. □

For a finite-dimensional diffusion $X_t^\varepsilon = x + \int_0^t \sigma(s,X_s^\varepsilon)\sqrt{\varepsilon}\,dW_s + \int_0^t b(s,X_s^\varepsilon)\,ds$, where $x \in \mathbb{R}^d$, W is an r-dimensional standard Brownian, $\sigma : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^r$ and $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ are globally Lipschitz with linear growth, Theorem 3.13 can be used to recover the Freidlin Wentzell theorem 2.4 .

In the finite dimensional setting, $H_0 = \mathbb{R}^r$. For $h^\varepsilon \in \mathcal{A}_M$, $\tilde{X}_t^\varepsilon = X_t^\varepsilon\left(W + \frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} h_s^\varepsilon\,ds\right)$ is solution of the SDE: $\tilde{X}_t^\varepsilon = x + \int_0^t \sqrt{\varepsilon}\,\sigma(s,\tilde{X}_s^\varepsilon)\,dW_s + \int_0^t \sigma(s,\tilde{X}_s^\varepsilon)h_s^\varepsilon\,ds + \int_0^t b(s,X_s^\varepsilon)\,ds$.

Then $\tilde{X}^\varepsilon = g^\varepsilon\left(W + \frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} h_s^\varepsilon\,ds\right)$, where $X^\varepsilon = g^\varepsilon(W)$.

given $h \in S_M$, $g^0\left(\int_0^{\cdot} h_s\,ds\right)$ is solution of the ODE

$$g^0\left(\int_0^{\cdot} h_s\,ds\right)_t = x + \int_0^t \sigma\left(s, g^0\left(\int_0^{\cdot} h_u\,du\right)_s\right)h_s\,ds + \int_0^t b\left(s, g^0\left(\int_0^{\cdot} h_u\,du\right)_s\right)ds$$

One can prove that if $h^\varepsilon \to h$ in distribution (when $L^2(0,T;\mathbb{R}^r)$ is endowed with the weak topology), $\tilde{X}^\varepsilon$ converges in distribution to $g^0\left(\int_0^{\cdot} h_s\,ds\right)$, and that $g^0(S_M)$ is compact.

The proofs are left as an exercise.

3.4 LDP for hydrodynamical models (related to the Navier Stokes equations)

The weak convergence approach has been used to prove LDP for several stochastic PDEs. Let us focus on two results concerning the 2D Navier-Stokes equations subjected to a small random perturbation.

3.4.1 Navier-Stokes equation with small perturbation

The Navier-Stokes equation is a celebrated PDE describing the velocity of a fluid. In dimension 3, finding a topological space where existence and uniqueness of the solution hold is still an open problem. Therefore, we will deal with the 2 dimensional case. Let $D \subset \mathbb{R}^2$ be a bounded, open and simply connected domain with the Dirichlet non-slip boundary condition. We look for a function $u = (u^1, u^2)$ of t and $x = (x_1, x_2)$ such that

$$\partial_t u - \nu \Delta u + u.\nabla u + \nabla p = f, \quad \operatorname{div} u = 0 \text{ in } D, \quad u = 0 \text{ on } \partial D,$$

where $\Delta = \partial^2_{x_1} + \partial^2_{x_2}$ is the Laplace operator, $\operatorname{div} u = \partial_1 u^1 + \partial_2 u^2$,
$(u.\nabla u)^i = \sum_{j=1}^{2} u^j \partial_j u^i$, $i = 1, 2$, and p is the pressure.

We project this equation on divergence free fields and let

$H = \{ f \in (L^2(D))^2 : \operatorname{div} u = 0 \text{ in } D, \ f.n = 0 \text{ on } \partial D \}$, where n is the outer normal to ∂D. Let $V = [H^1(D)]^2 \cap H$; for $u_1, u_2 \in V$ set

$A(u_1, u_2) = \sum_{j=1}^{2} \int_D \nabla u_1^j \cdot \nabla u_2^j \, dx$; then A is the Stokes operator, which is bilinear, symmetric and positive. Suppose that $\nu = 1$ to ease notations.

Let $B : V \times V \longrightarrow V'$ be the bilinear map defined by

$$\langle B(u_1, u_2), u_3 \rangle = \int [u_1(x) . \nabla u_2(x)] . u_3(x) = \sum_{i,j=1}^{2} \int_D u_1^j \partial_j u_2^i u_3^i \, dx, \quad u_1, u_2, u_3 \in V.$$

Then $\langle B(u_1,u_2),u_3\rangle = -\langle B(u_1,u_3),u_2\rangle$.

Furthermore, if $\mathcal{H} = L^4(D)\cap H$, the Gagliardo Nirenberg inequality implies $V\subset\mathcal{H}\subset H$, $\|u\|_{\mathcal{H}}^2 \le q_0 |u|_H \|u\|_V$ for any $u\in V$.

For any $\eta>0$ there exists $C_\eta>0$ such that

$$|\langle B(u_1,u_2),u_3\rangle| \le \eta\|u_3\|_V^2 + C_\eta \|u_1\|_{\mathcal{H}}^2\|u_2\|_{\mathcal{H}}^2,$$

$$|\langle B(u_1,u_1)-B(u_2,u_2),u_1-u_2\rangle| \le \eta\|u_1-u_2\|_V^2 + C_\eta\|u_2\|_{\mathcal{H}}^4 |u_1-u_2|_H^2.$$

We next suppose that f is a random forcing term driven by some Q Brownian motion in H, and let $H_0 = Q^{\frac12}H$. We denote by $L_Q(H_0,H)$ the set of linear operators $S: H_0\to H$ such that $SQ^{\frac12}$ is Hilbert-Schmidt from H to H; $\|S\|_{L_Q}^2 = \mathrm{trace}(SQS^*) = \sum_{k\ge1} |(SQ^{\frac12})\psi_k|_H^2$, where (ψ_k) is any orthonormal basis of H.

The noise intensity may depend on u, and is described by $\sigma:[0,T]\times V\to L_Q$. We suppose that σ satisfies the following

$\boxed{\text{Condition (C)}}$ $\sigma\in\mathcal{C}([0,T]\times V;L_Q)$ and there exist non negative constants K_i and L_i such that

(i) $\|\sigma(t,u)\|_{L_Q}^2 \le K_0 + K_1|u|_H^2 + K_2\|u\|_V^2$, $\forall t\in[0,T]$, $\forall u\in V$

(ii) $\|\sigma(t,u)-\sigma(t,v)\|_{L_Q}^2 \le L_1|u-v|_H^2 + L_2\|u-v\|_V^2$, $\forall t\in[0,T]$, $\forall u,v\in V$.

Let $\mathcal{A}$ be the set of H_0-valued $(\mathcal{F}_t)$ predictable random fields h such that $\int_0^T \|h(s)\|_0^2\,ds<\infty$ a.s. For $M>0$ let $S_M = \{h\in L^2(0,T;H_0) : \int_0^T\|h(s)\|_0^2\,ds\le M\}$ be endowed with the weak topology and let $\mathcal{A}_M = \{h\in\mathcal{A} : h(\omega)\in S_M \text{ a.s.}\}$

We consider a stochastic control equation with another intensity $\tilde{\sigma} : [0,T]\times V \to L(H_0,H)$

Condition($\tilde{C}$) $\tilde{\sigma} \in \mathcal{C}([0,T]\times V;\ L(H_0,H))$ and there exist constants $\tilde{K}_{\mathcal{H}}$, $\tilde{K}_i$ and $\tilde{L}_i$ such that

(i) $|\tilde{\sigma}(t,u)|^2_{L(H_0,H)} \leq \tilde{K}_0 + \tilde{K}_1 |u|_H^2 + \tilde{K}_{\mathcal{H}} \|u\|_{\mathcal{H}}^2$, $\quad \forall t\in[0,T],\ \forall u\in V$,

(ii) $|\tilde{\sigma}(t,u) - \tilde{\sigma}(t,v)|^2_{L(H_0,H)} \leq \tilde{L}_1 |u-v|_H^2 + \tilde{L}_2 \|u-v\|_V^2$, $\forall t\in[0,T]$, $\forall u,v\in V$.

<u>Remark:</u> If $K_2 = L_2 = 0$, we may take $\tilde{\sigma} = \sigma$ and $\tilde{K}_{\mathcal{H}} = \tilde{L}_2 = 0$.

Consider the stochastic control equation: for $h\in \mathcal{A}_M$, $\xi\in H$, let

$$d u_h(t) + [A u_h(t) + B(u_h(t), u_h(t))]\,dt = \sigma(t, u_h(t))\,dW_t + \tilde{\sigma}(t, u_h(t))\, h(t)\,dt$$

with the initial condition $u_h(0) = \xi$.

Let $X = \mathcal{C}([0,T];H) \cap L^2(0,T;V)$ and set $\|u\|_X^2 = \sup_{t\in[0,T]} |u(t)|_H^2 + \int_0^T \|u(s)\|_V^2\, ds$.

We say that $u_h(t,\cdot)$ is a weak solution of the stochastic control equation if $u\in X$ a.s. and for $v\in \mathrm{Dom}(A)$, $t\in[0,T]$,

$$(u_h(t), v) - (\xi, v) + \int_0^t [\langle u_h(s), Av\rangle + \langle B(u_h(s), u_h(s)), v\rangle]\, ds$$
$$= \int_0^t (\sigma(s, u_h(s))\,dW_s, v) + \int_0^t (\tilde{\sigma}(s, u_h(s))\, h(s), v)\, ds.$$

Theorem 3.14 Assume that conditions (C) and ($\tilde{C}$) are satisfied. Then for every $M>0, T>0$, there exists a constant $\bar{K}_2 := \bar{K}_2(T,M)$ such that if ξ is independent of (W_t), $E|\xi|_H^4 < \infty$, $h\in \mathcal{A}_M$, $K_2\in[0,\bar{K}_2)$ and $L_2 < 2$, there exists a weak solution $u_h \in X$ of the stochastic control equation with initial data $u_h(0) = \xi$. Furthermore there exist a constant $C := C(T,M)$ such that for $h\in\mathcal{A}_M$,

$$E\Big(\sup_{t\in[0,T]} |u_h(t)|_H^4 + \int_0^T \|u_h(t)\|_V^2\, dt + \int_0^T \|u_h(t)\|_{\mathcal{H}}^4\, dt\Big) \leq C\,(1 + E(|\xi|_H^4))$$

If L_2 is small enough, u_h is the unique solution in X.

The proof is decomposed in several steps.

<u>Step 1</u> Prove a priori bounds for the Galerkin approximation $u_{m,h}$ of u_h.

<u>Step 2</u> Extract subsequences of $u_{m,h}$, $\sigma(\cdot, u_{m,h})$, $\tilde{\sigma}(\cdot, u_{m,h})$,

$-[Au_{m,h} + B(u_{m,h}, u_{m,h})]$ converging weakly to $\bar{u}_h$, S_h, $\tilde{S}_h$ and

F_h in appropriate functional spaces.

Prove that $\bar{u}_h(t) = \xi + \int_0^t \bar{S}_h(s)\, dW_s + \int_0^t \tilde{S}_h(t)\, h(t)\, dt + \int_0^t F_h(s)\, ds$

<u>Step 3</u> Prove that $\bar{S}_h(t) = \sigma(t, \bar{u}_h(t))$, $\tilde{S}_h(t) = \tilde{\sigma}(t, \bar{u}_h(t))$, $F_h(t) = -[A\bar{u}_h(t) + B(\bar{u}_h(t), \bar{u}_h(t))]$

<u>Step 4</u> Prove uniqueness of the weak solution in X

We next suppose that $K_2 = L_2 = 0$, and that σ satisfies the time regularity

condition $(\bar{C})$

Condition $(\bar{C})$ There exist constants $\gamma > 0$, $C \geq 0$ such that for $u \in V$, $t_1, t_2 \in [0,T]$,

$$\|\sigma(t_1, u) - \sigma(t_2, u)\|_{L_Q} \leq C(1 + \|u\|_V)\, |t_1 - t_2|^\gamma$$

For $\varepsilon > 0$, let $du^\varepsilon(t) + [Au^\varepsilon(t) + B(u^\varepsilon(t), u^\varepsilon(t))]dt = \sqrt{\varepsilon}\, \sigma(t, u^\varepsilon(t))\, dW(t)$, $u^\varepsilon(0) = \xi \in H$

Then $u^\varepsilon = g^\varepsilon(W)$. Given $h \in L^2(0,T; H_0)$, let

$$dS(h)_t + [AS(h)_t + B(S(h)_t, S(h)_t)]dt = \sigma(t, S(h)_t)\, h_t\, dt, \quad S(h)_0 = \xi.$$

Then $S(h)_t = g^0\left(\int_0^\cdot h(s)\, ds\right)$.

Theorem 3.15 Suppose that condition (C) holds with $K_2 = L_2 = 0$, and that the time regularity condition $(\bar{C})$ is satisfied. Then the family of solutions (u^ε) satisfies in X a LDP with the good rate function I defined by

$$I(u) = \inf\left\{ \frac{1}{2}\int_0^T |h(s)|_0^2\, ds : h \in L^2(0,T; H_0),\ u = S(h) \right\}$$

We sketch the main steps of the proof.

Given $N>0$, $h_\varepsilon \in \mathcal{A}_M$ let $G_N^\varepsilon(t) = \{\omega : (\sup_{0\le s\le t} |u_{h_\varepsilon}(s)|_H^2) \vee \int_0^t \|u_h(s)\|_V^2\, ds) \le N\}$

For every integer $m\geq 1$, let $\psi_m : [0,T] \to [0,T]$ be measurable such that for $s\in[0,T]$, $\psi_m(s) \in [s, (s+c2^{-n})\wedge T]$ for some positive constant c. Then we have a very weak time regularity property of the weak solution u_h^ε.

Lemma 3.16 Fix $\varepsilon>0$, $M>0$, $N>0$ and let $\tilde{\sigma}=\sigma$ satisfy condition (C).

Let $\xi \in L^4(\Omega; H)$; then there exists a positive constant C depending on M, N and ε_0 such that for any $\varepsilon \in [0, \varepsilon_0]$, $h_\varepsilon \in \mathcal{A}_M$,

$$I_m(h,\varepsilon) := E\Big[1_{G_N(T)} \int_0^T |u^\varepsilon_{h_\varepsilon}(s) - u^\varepsilon_{h_\varepsilon}(\psi_m(s))|_H^2\, ds\Big] \le C\, 2^{-\frac{m}{2}}.$$

Let $(h_\varepsilon, \varepsilon>0)$ be a family of elements of $\mathcal{A}_M$ which converges in distribution to $h\in\mathcal{A}_M$. For each $\varepsilon>0$, let $u_{h_\varepsilon} := g^\varepsilon(W + \frac{1}{\sqrt{\varepsilon}}\int_0^\cdot h_\varepsilon(s)ds)$ be solution to

$$du_{h_\varepsilon}(t) + [Au_{h_\varepsilon}(t) + B(u_{h_\varepsilon}(t), u_{h_\varepsilon}(t))]dt = \sqrt{\varepsilon}\,\sigma(t, u_{h_\varepsilon}(t))dW_t + \sigma(t, u_{h_\varepsilon}(t))h_\varepsilon(t)\,dt,$$

$u_{h_\varepsilon}(0) = \xi$, and let $g^0(\int_0^\cdot h(s)ds) = u_h$.

Proposition 3.17 Let the assumptions of Theorem 3.15 be satisfied. Then as $\varepsilon\to 0$, $u_{h\varepsilon} = g^\varepsilon(W + \frac{1}{\sqrt{\varepsilon}}\int_0^\cdot h_\varepsilon(s)ds)$ converges in distribution to u_h in X.

<u>Proof</u> The Skorokhod theorem implies the existence of $(\tilde{h}_\varepsilon, \tilde{h}, \tilde{W})$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that $\tilde{h}_\varepsilon \to \tilde{h}$ a.s. in the weak topology of S_M, and $(\tilde{h}_\varepsilon, \tilde{W}^\varepsilon)$ has the same distribution as $(h_\varepsilon, W^\varepsilon)$, where $W^\varepsilon = W + \int_0^\cdot h_\varepsilon(s)ds$ and $\tilde{W}^\varepsilon = \tilde{W} + \frac{1}{\sqrt{\varepsilon}}\int_0^\cdot \tilde{h}_\varepsilon(s)ds$.

To ease notations, let $(\tilde{h}_\varepsilon, \tilde{h}, \tilde{W}^\varepsilon) = (h_\varepsilon, h, W^\varepsilon)$.

For every $t\in[0,T]$, $\int_0^t h_\varepsilon(s)\,ds \longrightarrow \int_0^t h(s)\,ds$ weakly in H_0.

Set $U_\varepsilon = u_{h_\varepsilon} - u_h$; then $v_\varepsilon(0) = 0$.

Using Itô's lemma for $u_{h_\varepsilon} - u_h$, we deduce

$$|v_\varepsilon(t)|_H^2 + \int_0^t \|v_\varepsilon(s)\|_V^2 ds \le 2\int_0^t \Big[C_{\frac12} \|u_h(s)\|_{\mathcal{X}}^4 + \sqrt{L_1}\, |h_\varepsilon(s)|_0 \Big] |v_\varepsilon(s)|_H^2 ds + \sum_{i=1}^3 T_i(t,\varepsilon),$$

where

$$T_1(t,\varepsilon) = 2\sqrt{\varepsilon} \int_0^t (v_\varepsilon(s), \sigma(s, u_{h_\varepsilon}(s))\, dW(s)),$$

$$T_2(t,\varepsilon) = \varepsilon \int_0^t [K_0 + K_1 |u_{h_\varepsilon}(s)|_H^2]\, ds,$$

$$T_3(t,\varepsilon) = 2\int_0^t (\sigma(s, u_h(s))[h_\varepsilon(s) - h(s)], v_\varepsilon(s))\, ds.$$

We will prove that $\|v_\varepsilon\|_X \to 0$ in probability, which will imply the desired convergence in distribution. Set

$$G_N(t) = \{\sup_{0\le s\le t} |u_h(s)|_H^2 \le N\} \cap \{\int_0^t \|u_h(s)\|_V^2 ds \le N\}$$

$$G_{N,\varepsilon}(t) = G_N(t) \cap \{\sup_{0\le s\le t} |u_{h_\varepsilon}(s)|_H^2 \le N\} \cap \{\int_0^t \|u_{h_\varepsilon}(s)\|_V^2 ds \le N\}$$

<u>Step 1</u> For any $\varepsilon_0 > 0$, $\sup_{\varepsilon\in(0,\varepsilon_0]} \sup_{h, h_\varepsilon \in \mathcal{A}_M} P(G_{N,\varepsilon}(T)^c) \le C(1 + E|\xi|_H^4) N^{-1}$

<u>Step 2</u> For fixed $N > 0$, as $\varepsilon \to 0$

$$E\Big[1_{G_{N,\varepsilon}(T)} \Big(\sup_{0\le t\le T} |v_\varepsilon(t)|_H^2 + \int_0^T \|v_\varepsilon(s)\|_V^2 ds \Big)\Big] \longrightarrow 0.$$

Gronwall's lemma implies that on $G_{N,\varepsilon}(T)$,

$$\sup_{0\le t\le T} |v_\varepsilon(t)|_H^2 \le \Big[\sup_{0\le t\le T} \{T_1(t,\varepsilon) + T_3(t,\varepsilon)\} + C(N) T\varepsilon \Big] \exp\big(2a_0 C_{\frac12} N^2 + 2\sqrt{L_1 MT}\big).$$

Indeed $\int_0^T \|u_h(s)\|_{\mathcal{X}}^4 ds \le a_0 \sup_{0\le s\le T} |u_h(s)|_H^2 \int_0^T \|u_h(s)\|_V^2 ds \le a_0 N^2$ on $G_{N,\varepsilon}(T)$.

Since $G_{N,\varepsilon}(\cdot)$ decreases,

$$E\Big[1_{G_{N,\varepsilon}(T)} \sup_{0\le t\le T} |T_1(t,\varepsilon)|\Big) \le E(\lambda_\varepsilon), \text{ where}$$

$$\lambda_\varepsilon = 2\sqrt{\varepsilon} \sup_{0\le t\le T} \Big| \int_0^t 1_{G_{N,\varepsilon}(s)} (v_\varepsilon(s), \sigma(s, u_{h_\varepsilon}(s))\, dW(s)) \Big|$$

Using the Burkholder - Davies Gundy inequality, we deduce

$E(\Lambda_\varepsilon) \le C(T,N)\sqrt{\varepsilon}$.

For $s \in [kT2^{-n}, (k+1)T2^{-n})$, set $\psi_n(s) = \bar{s}_n = T(k+1)2^{-n}$ and $t_k = kT2^{-n}$.

Then

$$E\Big[1_{G_{N,\varepsilon}(T)} \sup_{0\le t\le T} |T_3(t,\varepsilon)|\Big) \le 2\sum_{i=1}^{4} \tilde{T}_i(N,n,\varepsilon) + 2E(\bar{T}_5(N,n,\varepsilon)),$$

where

$$\tilde{T}_1(N,n,\varepsilon) = E\Big[1_{G_{N,\varepsilon}}(T) \sup_{0\le t\le T}\Big|\int_0^t \big(\sigma(s,u_h(s))[h_\varepsilon(s)-h(s)],\, v_\varepsilon(s)-v_\varepsilon(\bar{s}_n)\big)\,ds\Big|\Big],$$

$$\tilde{T}_2(N,n,\varepsilon) = E\Big[1_{G_{N,\varepsilon}}(T) \sup_{0\le t\le T}\Big|\int_0^t \big([\sigma(s,u_h(s))-\sigma(\bar{s}_n,u_h(s))][h_\varepsilon(s)-h(s)],\, v_\varepsilon(\bar{s}_n)\big)\,ds\Big|\Big],$$

$$\tilde{T}_3(N,n,\varepsilon) = E\Big[1_{G_{N,\varepsilon}}(T) \sup_{0\le t\le T}\Big|\int_0^t \big([\sigma(\bar{s}_n,u_h(s))-\sigma(\bar{s}_n,u_h(\bar{s}_n))][h_\varepsilon(s)-h(s)],\, v_\varepsilon(\bar{s}_n)\big)\,ds\Big|\Big],$$

$$\tilde{T}_4(N,n,\varepsilon) = E\Big[1_{G_{N,\varepsilon}(T)} \sup_{0\le t\le T}\ \sup_{t_{k-1}\le t\le t_k}\Big|\Big(\sigma(t_k,u_h(t_k))\int_{t_{k-1}}^{t}(h_\varepsilon(s)-h(s))\,ds,\, v_\varepsilon(t_k)\Big)\Big|\Big],$$

$$\bar{T}_5(N,n,\varepsilon) = 1_{G_{N,\varepsilon}(T)} \sum_{k=1}^{2^n}\Big|\Big(\sigma(t_k,u_h(t_k))\int_{t_{k-1}}^{t_k}(h_\varepsilon(s)-h(s))\,ds,\, v_\varepsilon(t_k)\Big)\Big|.$$

Using the Cauchy-Schwarz inequality, condition $(\bar{C})$ and Lemma 3.16, we prove that $\tilde{T}_1(N,n,\varepsilon) \le C_1 2^{-\frac{n}{4}}$, $\tilde{T}_3(N,n,\varepsilon) \le C_3 2^{-\frac{n}{4}}$, $\tilde{T}_2(N,n,\varepsilon) \le C_3 2^{-n\gamma}$, $\tilde{T}_4(N,n,\varepsilon) \le C_4 2^{-\frac{n}{2}}$.

Note that the weak convergence of h_ε to h implies that for $a,b \in [0,T]$, $a<b$, $\int_a^b h_\varepsilon(s)\,ds \to \int_a^b h(s)\,ds$ in the weak topology of H_0.

Since $\sigma(t_k, u_h(t_k))$ is a compact operator from H_0 to H,

$$\Big|\sigma(t_k,u_h(t_k))\Big(\int_{t_{k-1}}^{t_k} h_\varepsilon(s)\,ds - \int_{t_{k-1}}^{t_k} h_s\,ds\Big)\Big|_H \to 0 \text{ as } \varepsilon\to 0.$$

Hence $\bar{T}_5(N,n,\varepsilon) \to 0$ a.s. as $\varepsilon\to 0$. Since it is upper bounded by a constant, the dominated convergence theorem implies $E[\bar{T}_5(N,n,\varepsilon)] \to 0$ as $\varepsilon \to 0$.

Hence $\overline{\lim_{\varepsilon \to 0}}\, E\left[1_{G_{N,\varepsilon}(T)} \sup_{t\in[0,T]} |T_3(t,\varepsilon)|\right] \le C(N,T,M)\; 2^{-n(\gamma \wedge \frac{1}{4})}$.

Letting $n \to \infty$, we deduce $\lim_{\varepsilon \to 0} E\left[1_{G_{N,\varepsilon}(T)} \sup_{t\in[0,T]} |T_3(t,\varepsilon)|\right] = 0$.

The Markov inequality implies for any $\delta \to 0$,

$$P(\|v_\varepsilon\|_X > \delta) \le P(G_{N,\varepsilon}(T)^c) + \frac{1}{\delta^2} E\left[1_{G_{N,\varepsilon}(T)} \|v_\varepsilon\|_X^2\right].$$

Hence $\overline{\lim_{\varepsilon \to 0}}\, P(\|v_\varepsilon\|_X > \delta) \le C(T,M)\, N^{-1}$.

We conclude the proof letting $N \to \infty$. □

To use Theorem 3.13, we still have to prove the following

Proposition 3.18 Let the assumptions of Theorem 3.15 be satisfied. Fix $M>0$, $\xi \in H$ and let $K_M = \{u_h : h \in S_M\}$. Then K_M is a compact subset of X.

<u>Proof</u> Let (u_n) be a sequence of K_M, where $u_n = u_{h_n}$, $h_n \in S_M$. Since S_M is bounded in $L^2(0,T; H_0)$, it is weakly compact. Hence there exists $h \in S_M$ and a subsequence, still denoted h_n, such that $h_n \to h$ for the weak topology of S_M. One proves that $u_n = u_{h_n}$ converges to $u = u_h$ in X. The details of the proof, which is similar to that of Proposition 3.17 (and easier) are left as an exercise. □

3.4.2 Some extensions to other hydrodynamical models

(1) We may <u>couple the 2D Navier Stokes equations</u> (describing the velocity) with similar PDEs describing the temperature or the magnetic field. This gives rise to the Boussinesq equations, the magneto-hydrodynamical models or the general Bénard problem. Large deviations for these models have been proved by Chueshov–M. and Duan–M. using arguments similar to the previous ones.

(2) One may try to work with the 3D-Navier Stokes equations on $\mathbb{R}^3$, and to have an anisotropic model, that is suppose that the viscosity is positive only in two directions (but vanishes in the third one). An extra term has to be added, and can be of Brinkman–Forchheimer type due to the 3D-setting.

As the strength of the stochastic perturbation goes to 0, one can prove a LDP result. The initial condition has to be smoother in one direction (Bessaih–M.)

(3) One can let the viscosity ν converge to 0 and supposed that the noise intensity is multiplied by $\sqrt{\nu}$, that is in dimension 2 consider

$$d u_\nu(t) - \nu \Delta u_\nu(t) + (u_\nu(t)\cdot\nabla) u_\nu(t) + \nabla p = \sqrt{\nu}\, \sigma_\nu(t, u_\nu(t))\, dW(t), \quad \nabla\cdot u_\nu = 0 \text{ in } D,$$

$$\operatorname{curl} u_\nu = \partial_1 u_\nu^2 - \partial_2 u_\nu^1 = 0 \text{ on } \partial D, \quad u_\nu(t,\cdot)\cdot n = 0 \text{ on } \partial D, \quad u_\nu(0,\cdot) = \xi \text{ in } D.$$

If $\xi \in V$, under stronger regularity conditions on σ, one can prove a LDP in a more restrictive functional set. This is due to the fact that the rate function is described in terms of the Euler equation. The weak convergence approach is used with functional arguments and no "weak time regularity" (Bessaih–M.)

References

[Be Mi 1] Bessaih, H., Millet, A., Large deviations and the zero viscosity limit for 2D Navier Stokes equations with free boundary, SIAM J. Math. Anal. 44(3), 2012, p.1861-1893.

[Be Mi 2] Bessaih, H., Millet, A., On stochastic modified 3D Navier-Stokes equations with anisotropic viscosity, J. Math. Anal. Appl., online 2017.

[Bo Du] Boué, M., Dupuis, P., A variational representation for certain functionals of Brownian motion, Ann. Probab 26(4), 1998, p. 1641-1659.

[Bu Du] Budhiraja, A., Dupuis, P., A variational representation for positive functionals of infinite dimensional Brownian motion, Probab. Math. Stat. 20, 2000, p. 39-61.

[Che Mi] Chenal, F., Millet, A., Uniform large deviations for parabolic SPDEs and Applications, Stochastic Proc. Appl. 72(2), 1997, p. 161-186.

[Chu Mi] Chueshov, I., Millet, A., Stochastic 2D hydrodynamical type systems: Well-posedness and large deviations, Appl. Math. Optim. 61(3), 2010, p. 379-420.

[De Ze] Dembo, A., Zeitouni, O., Large Deviations Techniques and Applications, Springer, New-York (2010).

[Deu St] Deuschel, J.D., Stroock, D., Large Deviations, Pure and Applied Mathematics 137, Academic Press Inc., Boston (1989).

[Du El] Dupuis, P., Ellis, R.S., Weak convergence Approach to the Theory of Large Deviations, Wiley - Interscience, New-York (1997).

[SrSu] Sritharan, S.S., Sundar, P., Large deviations for the two-dimensional Navier-Stokes equations with multiplicative noise, Stoch. Process. Appl., 116, 2006, p. 1636-1659.

[St] Stroock, D., Some applications of stochastic calculus to partial differential equations, Ecole d'Eté de Probabilités de Saint-Flour XI -1981, Lecture Notes in Math. 976, 1983, p. 267-382.