Chapitre 9

Inégalités fonctionnelles et modèle gaussien

Dans ce chapitre, nous commençons par présenter des inégalités fonctionnelles expliquons de quelle manière celles-ci permettent d'obtenir de la concentration. En particulier, nous présenterons les inégalités de Poincaré et de Sobolev logarithmique. Afin de mieux appréhender ces nouveaux objets, nous nous focaliserons ensuite sur la mesure gaussienne standard sur $\mathbb{R}^n$ et présenterons une application en statistique. Ce chapitre sera également l'occasion de rappeler certaines applications obtenues dans le chapitre 5 et d'évoquer les possibles généralisations à des mesures de probabilités log-concave de la forme $d\mu = e^{-V(x)}dx$ avec V un potentiel strictement convexe. Ces extensions seront facilement visibles par le prisme des démonstration par semi-groupe.

9.1 Inégalité fonctionelles

Bien qu'il soit possible de considérer le cadre d'espace métrique mesuré, nous nous restreignons dans ce qui suit à l'espace euclidien $(\mathbb{R}^n, |\cdot|)$ (équipé de sa tribu des boréliens $\mathcal{B}(\mathbb{R}^d)$) muni d'une mesure de probabilité μ.

9.1.1 Inégalité de Poincaré

Définition 9.1.1. *Nous dirons que μ vérifie une inégalité de Poincaré de constante $C_P > 0$ pour une famille de fonction $\mathcal{A}_P$, si pour toute fonction $f \in \mathcal{A}_P$,*

$$\mathrm{Var}_\mu(f) \leq C_P \int_{\mathbb{R}^n} |\nabla f|^2 d\mu,$$

où $\mathrm{Var}_\mu(f) = \int_{\mathbb{R}^n} f^2 d\mu - \left(\int_{\mathbb{R}^n} f d\mu\right)^2$. Dans l'équation précédente, la quantité $|\nabla f|$ désigne la norme euclidienne du gradient.

Remarque. Dans le cadre d'un espace métrique mesuré (E, d, μ), il est possible de généraliser la notion de longueur de gradient d'une $f : E \to \mathbb{R}$ (en un point $x \in E$) de la manière suivante :

$$|\nabla f|(x) = \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x,y)}. \tag{9.1.1}$$

Sur $\mathbb{R}^n$, muni de la distance euclidienne, cette définition coïncide alors avec la norme euclidienne d'un gradient.

Une propriété remarquable des inégalités de Poincaré est qu'elles se tensorisent indépendamment de la dimension, plus précisément nous avons le résultat suivant [77].

Proposition 9.1.1. *Soit* (μ_i), $\quad i = 1, \dots, n$, *une famille de mesures de probabilité sur* $\mathcal{B}(\mathbb{R})$ *alors la mesure produit* $\mu = \mu_1 \otimes \dots \mu_n$ *satisfait l'inégalité suivante, pour* $f \in L^2(\mu)$,

$$\mathrm{Var}_{\mu^n}(f) \leq \sum_{i=1}^{n} \int_{\mathbb{R}^n} \mathrm{Var}_{\mu_i}(f) d\mu,$$

où Var_{μ_i} *signifie que l'on intègre uniquement par rapport à la* i*-ème coordonnée.*

En particulier, si les mesures μ_i *(avec* $i = 1, \dots, n$*) vérifient une inégalité de Poincaré (sur* $\mathbb{R}$*) de constante* $C_i > 0$*, alors la mesure produit* μ *satisfait une inégalité de Poincaré (sur* $\mathbb{R}^n$*) de constante* $C = \max_{i=1,\dots,n} C_i$.

Remarque. La démonstration de ce résultat repose sur l'inégalité de Jensen et a été utilisé dans la démonstration de l'inégalité de Poincaré 8.2.1.

Intégrabilité exponentielle

Comme il est exposé dans [77, 16] une inégalité de Poincaré de constante $C > 0$ entraine que les fonctions lipschitziennes sont exponentiellement intégrables. Plus précisement nous avons la proposition suivante [77].

Proposition 9.1.2. *Si* μ *vérifie une inégalité de Poincaré de constante* $C > 0$*, alors pour toute fonction* f 1*-lipschitzienne et pour tout* $s < \sqrt{\frac{4}{C}}$

$$\int_{\mathbb{R}^n} e^{sf} d\mu \leq e^{s \int_{\mathbb{R}^n} f d\mu} \prod_{l=0}^{\infty} \left(1 - \frac{Cs^2}{4^{l+1}}\right)^{-2^l} \tag{9.1.2}$$

Remarque. Il est à noter que la démonstration précédente repose sur le fait suivant : pour f fixée, il suffit d'appliquer l'inégalité de Poincaré à la fonction $g = e^{sf/2}$, $s \in \mathbb{R}$. Ceci permettant alors d'obtenir une inégalité fonctionnelle entre la fonction $Z(s) = \int_{\mathbb{R}^n} e^{sf} \mathrm{d}\mu$ et $Z(s/2)$. Cette inégalité servira à montrer que $Z(s)$ est bornée par 3 lorsque $s < \sqrt{\frac{4}{C}}$. Il suffit alors d'appliquer l'inégalité de Chernoff pour conclure (cf. [77, 16]).

Comme annoncé dans la remarque précédente, les inégalités de Poincaré sont liées à un phénomène de concentration exponentielle.

Corollaire 9.1.1. *Soit* μ *une mesure vérifiant une inégalité de Poincaré de constante* $C > 0$*. Alors, pour toute fonction* f *lipschitzienne et tout* $t \geq 0$,

$$\mu\left(f \geq \int_{\mathbb{R}^n} f d\mu + t\right) \leq 3e^{-3t/\sqrt{C}\|f\|_{Lip}}$$

Remarque. L'inégalité précédente correspond plutôt une inégalité de déviation (de la moyenne). Cependant le même argument appliqué à la fonction $-f$ fournit

$$\mu\left(\left|f - \int_{\mathbb{R}^n} f d\mu\right| \geq t\right) \leq 6e^{-3t/\sqrt{C}\|f\|_{Lip}}, \quad t \geq 0,$$

qui exprime bien une inégalité de concentration (autour de la moyenne).

9.1.2 Inégalité de Sobolev logarithmique

Comme nous allons le voir, l'inégalités de Sobolev logarithmique permet d'obtenir des résultats plus fort que ceux obtenus par l'inégalité de Poincaré.

Définition et conséquences

Définition 9.1.2. *On dit qu'une mesure μ statisfait une inégalité de Sobolev logarithmique pour une certaine classe de fonctions $\mathcal{A}_{LS}$, lorsqu'il existe une constante $C_{LS} > 0$ telle que, pour toute fonction $f \in \mathcal{A}_{LS}$, on ait :*

$$\mathrm{Ent}_\mu(f^2) \leq C_{LS} \int_{\mathbb{R}^n} |\nabla f|^2 d\mu \tag{9.1.3}$$

où $\mathrm{Ent}_\mu(f) = \int_{\mathbb{R}^n} (f \log f) d\mu - \left(\int_{\mathbb{R}^n} f d\mu\right) \log \left(\int_{\mathbb{R}^n} f d\mu\right)$ *désigne l'entropie d'une fonction mesurable positive f.*

Remarque. il est important de noter qu'une inégalité de Sobolev logarithmique entraine une inégalité de Poincaré et que nous avons la relation suivante entre les constantes C_{LS} et C_P. :

$$C_P \leq \frac{1}{2} C_{LS}.$$

En effet, soit $f \in \mathcal{A}_{LS}$. Un développement limité de la fonction

$$h \mapsto (1+h)\log(1+h)$$

en 0 nous donne :

$$\mathrm{Ent}_\mu\big((1+\epsilon f)^2\big) = \frac{\epsilon^2}{2}\mathrm{Var}_\mu(f) + o(\epsilon^2).$$

D'autre part, nous avons

$$\int_{\mathbb{R}^n} [(1+\epsilon f)']^2 d\mu = \epsilon^2 \int_{\mathbb{R}^n} f'^2 d\mu$$

En faisant tendre ϵ vers 0^+, nous obtenons une inégalité de Poincaré, de constante $\frac{1}{2}C_{LS}$.

Notons que l'ensemble $\mathcal{A}_{LS}$ contient strictement $\mathcal{A}_P$, une inégalité de Sobolev logarithmique permet donc de considérer une classe beaucoup plus grande de fonctions.

Par exemple, dans le cas de la mesure gaussienne γ_1, considérons la fonction $f(x) = e^{c\frac{x^2}{2}}$ avec $c \in]0,1[$. Il n'est pas difficile de montrer que $\mathrm{Var}_{\gamma_1}(f) = +\infty$ lorsque $c \in [\frac{1}{2}, 1[$ tandis que $\mathrm{Ent}_{\gamma_1}(f) < +\infty$ pour tout $c \in]0,1[$.

Tout comme les inégalités de Poincaré, les inégalités de Sobolev logarithmique jouissent d'une propriété de tensorisation indépendante de la dimension. Plus précisément, nous avons le résultat suivant (cf. [77]).

Proposition 9.1.3. *Soit $(\mu_i)_{i=1,\dots,n}$ une famille de mesures de probabilité sur $\mathcal{B}(\mathbb{R})$ alors, pour toute fonction f suffisamment régulière, la mesure produit $\mu = \mu_1 \otimes \ldots \otimes \mu_n$ vérifie l'inégalité suivante*

$$\mathrm{Ent}_{\mu^n}(f) \leq \sum_{i=1}^{n} \int_{\mathbb{R}^n} \mathrm{Ent}_{\mu_i}(f) d\mu.$$

En particulier, si les mesures μ_i (pour $i = 1, \dots, n$) vérifient une inégalité de Sobolev logarithmique (sur $\mathbb{R}$) de constante $C_i > 0$ alors μ en vérifie également une, de constante $C = \max_{i=1,\dots,n} C_i$.

Démonstration. Pour démontrer ceci, il faut utiliser la formulation variationnelle de l'entropie :

$$\mathrm{Ent}_\mu(f) = \sup \left\{ \mathbb{E}_\mu[fg] \quad \text{avec} \quad \mathbb{E}_\mu[e^g] = 1 \right\}$$

qui se démontre facilement à partir de l"inégalité $uv \leq u \log u - u + e^v$ pour $u \geq 0$ et $v \in \mathbb{R}$. Considérons ensuite une fonction $g : \mathbb{R}^n \to \mathbb{R}$ telle que $\mathbb{E}_\mu[e^g] = 1$ et exprimons celle-ci de la manière suivante :

$$g = \sum_{i=1}^{n} g_i$$

avec

$$g_1 = g - \log \int_{\mathbb{R}} e^g d\mu_1 \quad \text{et} \quad g_i = \log \frac{\int_{\mathbb{R}^{i-1}} e^g d\mu_1 \dots d\mu_{i-1}}{\int_{\mathbb{R}^{i-1}} d\mu_i \dots d\mu_i} \quad \text{pour tout} \quad i = 2, \dots, n.$$

Il est ensuite facile de vérifier que $\mathbb{E}_{\mu_i}[e^{g_i}] = 1$ pour tout $i = 1, \dots, n$. En conséquence, grâce à la formulation variationnelle de l'entropie sous μ_i, nous avons

$$\sum_{i=1}^{n} \mathbb{E}_{\mu_i}[fg_i] \leq \sum_{i=1}^{n} \mathrm{Ent}_{\mu_i}[f].$$

Par suite, nous en déduisons que

$$\mathbb{E}_\mu[fg] = \sum_{i=1}^{n} \mathbb{E}_\mu[fg_i] = \sum_{i=1}^{n} \mathbb{E}_\mu\left[\mathbb{E}_{\mu_i}[fg_i]\right] \leq \sum_{i=1}^{n} \mathbb{E}_\mu\left[\mathrm{Ent}_{\mu_i}[f]\right],$$

la proposition est donc démontrée. □

Nous allons voir dans la section suivante les conséquences d'une telle inégalité vis à vis de la concentration de la mesure.

Intégrabilité exponentielle

L'inégalité de Sobolev logarithmique étant plus forte que celle de Poincaré, il est naturel de s'attendre à obtenir des propriétés de concentration plus fortes. Ce fut l'observation de Herbst qui a remarqué qu'une telle inégalité entrainait un contrôle beaucoup plus précis de la transformée de Laplace. Plus précisement nous avons le résultat suivant :

Proposition 9.1.4 (Herbst). *Si μ satisfait une inégalité de Sobolev logarithmique de constante $C = C_{LS}$, alors pour toute fonction f 1-lipschitzienne et tout $\sigma^2 < \frac{1}{C}$ nous avons*

$$\int_{\mathbb{R}^n} e^{\sigma^2 f^2/2} d\mu < \infty.$$

Plus précisément, toute fonction f 1-lipschitzienne est intégrable et pour tout $\lambda \in \mathbb{R}$

$$\int_{\mathbb{R}^n} e^{\lambda f} \mathrm{d}\mu \leq e^{s \int_{\mathbb{R}^n} f d\mu + C\lambda^2/2}.$$

Remarque. Cet argument avait déjà été mis en oeuvre dans le chapitre 5 et sera étudié plus en détail dans un chapitre ultérieur.

Comme dans le cas de l'inégalité de Poincaré, ceci se retranscrit via l'argument de Chernoff en inégalité de déviation autour de la moyenne.

Proposition 9.1.5. *Si μ satisfait une inégalité de Sobolev logarithmique de constante $C = C_{LS}$, alors pour toutes fonctions f lipschitziennes et tout $t \geq 0$*

$$\mu\left(f \geq \int_{\mathbb{R}^n} f d\mu + t\right) \leq e^{-t^2/2C\|f\|_{Lip}^2}.$$

Une inégalité de concentration autour de la moyenne est facilement obtenue en remplaçant f par $-f$ dans ce qui précède :

$$\mu\left(\left|f - \int_{\mathbb{R}^n} f d\mu\right| \geq t\right) \leq 2e^{-t^2/2C\|f\|_{Lip}^2}, \quad t \geq 0$$

Démonstration. Il suffit de combiner l'argument de Herbst avec l'inégalité de Chernoff pour enfin optimiser en $\lambda \geq 0$. □

9.2 Modèle gaussien

Considérons $\mathbb{R}$ muni de sa métrique euclidienne ainsi que la mesure gaussienne standard sur $\mathbb{R}$. Rappelons que celle-ci admet la densité de probabilité suivante par rapport à la mesure de Lebesgue

$$d\gamma_1(x) = e^{-x^2/2} \frac{dx}{\sqrt{2\pi}} \quad x \in \mathbb{R}.$$

Le modèle gaussien est un exemple parfait pour illustrer les inégalités fonctionnelles que nous venons d'exposer dans ce chapitre.

Proposition 9.2.1. *La mesure γ_1 vérifie une inégalité de Poincaré avec une constante optimale $C_P = 1$ et $\mathcal{A}_P$ consiste en l'ensemble des fonctions $f : \mathbb{R} \to \mathbb{R}$ de l'espace de Sobolev $H^1(\mathbb{R}, \gamma_1)$.*

Cette proposition entraine que la mesure (produit) gaussienne γ_n sur $\mathbb{R}^n$ vérifie une égalité de Poincaré de constante $C_P = 1$, c'est à dire, pour toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulière,

$$\mathrm{Var}_{\gamma_n}(f) \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n \tag{9.2.1}$$

Remarque. Une démonstration par semi-groupe de ce résultat sera produite ultérieurement.

De l'inégalité de Poincaré pour la mesure gaussienne standard, nous en déduisons une inégalité de Poincaré pour la mesure induite par un vecteur gaussien dans $\mathbb{R}^n$ de matrice de covariance Γ.

Proposition 9.2.2. *Soit $X = (X_1, \dots, X_n)$ un vecteur gaussien de matrice de covariance Γ. L'inégalité suivante est toujours vérifiée :*

$$\mathrm{Var}\big(f(X)\big) \leq \mathbb{E}\big[\langle M\nabla f, M\nabla f\rangle\big] = \mathbb{E}\big[\langle \Gamma\nabla f, \nabla f\rangle\big],$$

avec ${}^tMM = \Gamma$.

Démonstration. Nous pouvons supposer, sans perdre de généralité, que Γ, qui est toujours positive, soit également définie positive. Ainsi, il existe $M \in M_n(\mathbb{R})$ telle que $\Gamma = {}^tMM$. Notons alors que, si Z suit une loi $\mathcal{N}(0, I_d)$ alors, MZ suit une loi $\mathcal{N}(0, \Gamma)$. En appliquant l'inégalité de Poincaré, vérifiée par la mesure gaussienne de $\mathbb{R}^n$, à la fonction $g(x) := f(Mx)$ nous obtenons que :

$$\mathrm{Var}\big(f(X)\big) \leq \mathbb{E}\big[\langle M\nabla f, M\nabla f\rangle\big] = \mathbb{E}\big[\langle \Gamma\nabla f, \nabla f\rangle\big].$$

□

Nous en déduisons le résultat suivant (cf. [65, **?**],

Corollaire 9.2.1. *Soit $X = (X_1, \dots, X_n)$ un vecteur gaussien de matrice de covariance Γ. L'inégalité suivante est toujours vérifiée :*

$$\mathrm{Var}\big(\max_{i=1,\dots,n} X_i\big) \leq \max_{i=1,\dots,n} \mathrm{Var}(X_i).$$

Démonstration. Posons $M_n = \max_{i=1,\dots,n} X_i$ le maximum des coordonnées d'un vecteur gaussien standard. On remarque alors que $f(X_1, \dots, X_n) = M_n = \sum_{i=1}^n X_i 1_{A_i}$, avec, pour tout i=1,. . . ,n, $A_i = \{X_i = \max_{j=1,\dots,n} X_j\}$. Ainsi, $\partial_i f = 1_{A_i}$ pour tout $i = 1, \dots, n$. Puisque $(A_i)_{i=1,\dots,n}$ est une partition de $\mathbb{R}^n$, l'inégalité est immédiate en appliquant l'inégalité de Poincaré, satisfaite par la loi de $(X_1, \dots, X_n)$, à la fonction $f(x) = \max_{i=1,\dots,n} x_i$. Cependant, pour être rigoureux, il aurait fallut faire la démonstration avec une fonction plus régulière que $\max_{i=1,\dots,n}$. On aurait pu utiliser la fonction suivante

$$F_\beta(x) = \frac{1}{\beta} \log\big(\sum_{i=1}^n e^{\beta x_i}\big) \quad \text{avec} \quad \beta > 0$$

où $x = (x_1, \dots, x_n)$, qui converge vers $\max_{i=1,\dots,n} x_i$ lorsque $\beta \to \infty$ (il s'agit d'une approximation de Gibbs) et conclure par convergence dominée. □

Remarque. Par convergence monotone, ce résultat s'étend à des supremum de processus gaussiens. Plus précisément, si (T, d) est un espace métrique et $(X_t)_{t\in T}$ un processus gaussien, nous avons

$$\mathrm{Var}\big(\sup_{t\in T} X_t\big) \leq \sup_{t\in T} \mathrm{Var}(X_t)$$

Application

Exemple 9.2.1. Soit $X = (X_{ij})_{1\leq i\leq j\leq n}$ une matrice du GUE (*Gaussian Unitary Ensemble*). Autrement dit, nous considérons une matrice (de taille $n \times n$) auto-adjointe gaussienne centrée de variance σ^2. Ceci revient à considérer une matrice hermitienne X de taille $n \times n$ dont les entrées, au dessus de la diagonale, sont des variables aléatoires gaussiennes complexes (respectivement réelles sur la diagonale) centrées et de variance σ^2 (autrement dit, la partie réelle et imaginaire sont des variables aléatoires gaussiennes centrées de variance $\sigma^2/2$). Si λ_n désigne la plus grande valeur propre de cette matrice, l'inégalité précédente entraine que

$$\mathrm{Var}\big(\lambda_n\big) \leq \sigma^2$$

Remarque. Comme il est présenté dans [78, 37], la borne obtenue est sous-optimale.

Passons maintenant à l'inégalité de Sobolev logarithmique.

Proposition 9.2.3. *La mesure gaussienne γ_1 sur $\mathbb{R}$ vérifie une inégalité de Sobolev logarithmique avec une constante $C_{LS} = 2$, autrement dit*

$$\mathrm{Ent}_{\gamma_1}(f^2) \leq 2 \int_{\mathbb{R}} f'^2 d\gamma_1.$$

Remarque. Une démonstration de ceci, par semi-groupe, sera donnée ultérieurement.

Via la propriété de tensorisation, Ce résultat entraine que la mesure gaussienne standard γ_n sur $\mathbb{R}^n$ vérifie une inégalité de Sobolev logarithmique, c'est à dire, pour f suffisamment régulière,

$$\mathrm{Ent}_{\gamma_n}(f^2) \leq 2 \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n. \tag{9.2.2}$$

Comme nous le verrons dans le chapitre suivant cette inégalité fournit une nouvelle manière permettant de justifier le contrôle de la transformée de Laplace d'une fonction lipschitzienne (cf. proposition (5.2.1) et ses conséquences donné dans le chapitre 5) : il s'agit de l'argument de Herbst.

Notamment, l'inégalité de Sobolev logarithmique satisfaite par la mesure gaussienne entraine l'inégalité de concentration suivante :

$$\mathbb{P}\bigg(\big|\sup_{t\in T} X_t - \mathbb{E}[\sup_{t\in T} X_t]\big| \geq r\bigg) \leq 2e^{-\frac{r^2}{2\sigma^2}}, \quad r \geq 0 \tag{9.2.3}$$

avec $(X)_{t\in T}$ un processus, centré, continu (p.p.) indexé par un ensemble pré-compact T (i.e. pour tout $\epsilon > 0$, T peut-être recouvert d'un nombre fini de boule de rayon ϵ) et $\sigma^2 = \sup_{t\in T} \mathbb{E}[X_t^2]$.

Il existe de nombreux résultats concernant les inégalités de Sobolev logarithmiques que nous ne présenterons pas dans ce cours. Notamment la perturbation d'une inégalité de Sobolev logarithmique ou encore la méthode de Lyapounov permettant d'établir l'existence d'une inégalité de Sobolev logarithmique ; nous renvoyons le lecteur vers l'ouvrage [16, 10] pour plus de détails à ce sujet.

9.3 Application en statistiques

Dans cette section nous allons introduire quelques éléments de statistiques permettant de faire de la sélection de modèle. En particulier, nous allons présenter la méthode du LASSO. Avant toutes choses, nous devons préciser le cadre dans lequel nous allons travailler.

9.3.1 Modèles linéaire gaussien généralisé

Définition 9.3.1. *Soit $\mathbb{H}$ un espace de Hilbert (séparable) et $\big(W(t)\big)_{t\in\mathbb{H}}$ un processus gaussien centré. Ce processus est dit isonormal si*

$$\mathbb{E}[W(t)W(s)] = \langle s,t\rangle_{\mathbb{H}}$$

où $\langle\cdot,\cdot,\rangle_{\mathbb{H}}$ désigne le produit scalaire de $\mathbb{H}$.

Remarque. La norme de $\mathbb{H}$ sera désigné par $\|\cdot\|_{\mathbb{H}}$.

Le problème qui va nous concerner est le suivant : pour tout $t\in\mathbb{H}$, nous observons une version perturbée de $\langle s,t\rangle_{\mathbb{H}}$. Autrement dit, nous avons à disposition les éléments suivants :

$$Y(t) = \langle s,t\rangle_{\mathbb{H}} + \epsilon W(t) \tag{9.3.1}$$

avec $\epsilon > 0$ fixé et nous souhaitons retrouver (ou, tout du moins, approcher) le vecteur $s\in\mathbb{H}$.

L'avantage de cette formulation du problème est qu'elle englobe de nombreux cas de figures.

Exemple 9.3.1. Soit $X = (X_1,\dots,X_n)$ un vecteur gaussien standard et supposons que nous observions les données suivantes

$$Y_j = s_j + \sigma X_j \quad \text{pour tout} \quad j = 1,\dots,n$$

avec $\sigma > 0$. Le but est alors d'approcher le vecteur $s = (s_1,\dots,s_n)$. Cet exemple classique est bien exprimable dans le cadre décrit ci-dessus. En effet, il suffit de choisir

$$\mathbb{H} = \mathbb{R}^n \quad \text{muni de} \quad \langle u,v\rangle_{\mathbb{H}} = \frac{1}{n}\sum_{i=1}^n u_i v_i$$

et le processus isonormal

$$W(t) = \sqrt{n}\langle X,t\rangle_{\mathbb{H}}.$$

Ainsi, en choisissant $\epsilon = \frac{\sigma}{\sqrt{n}}$ nous avons bien égalité entre $Y(t) = \langle Y,t\rangle_{\mathbb{H}}$ et $\langle s,t\rangle_{\mathbb{H}} + \epsilon W(t)$.

Voici un autre exemple.

Exemple 9.3.2. Supposons que nous ayons à disposition une réalisation d'un processus stochastique $\zeta(x)$ pour $x \in [0,1]$ et que celui-ci soit solution de l'équation différentielle stochastique suivante

$$d\zeta(x) = s(x)dx + \epsilon dB(x) \quad \text{avec} \quad \zeta(0) = 0$$

où B est un mouvement brownien, s une fonction de carré intégrable et $\epsilon > 0$. Choisissons ensuite $\mathbb{H} = L^2\big([0,1]\big)$ muni du produit scalaire usuel $\langle s,t\rangle_{\mathbb{H}} = \int_0^1 s(x)t(x)dx$. Il n'est pas difficile de vérifier que

$$W(t) = \int_0^1 t(x)dB(x) \quad \text{pour tout} \quad t \in L^2\big([0,1]\big)$$

est un processus isonormal sur $\mathbb{H}$. Il suffit ensuite de choisir $Y(y) = \int_0^1 t(x)d\zeta(x)$ pour que l'identité (9.3.1) soit vérifiée dans notre contexte.

9.3.2 Choix de modèle

Définition 9.3.2. *Un modèle est un sous ensemble convexe et fermé $S \subset \mathbb{H}$*

Nous allons donc chercher à approcher l'élément $s \in \mathbb{H}$ par l'élément de S qui minimise

$$\inf_{t\in S} \|t-s\|^2_{\mathbb{H}}.$$

De manière équivalente, cela revient à déterminer l'infimum suivant :

$$\inf_{t\in S} -2\langle s,t\rangle_{\mathbb{H}} + \|t\|^2_{\mathbb{H}}$$

Puisque s est inconnu, nous allons le remplacer par son analogue bruité, et chercher à minimiser (sur S) la fonction suivante

$$\gamma(t) = -2Y(t) + \|t\|^2_{\mathbb{H}} \quad \text{avec} \quad t \in S.$$

Notons au passage qu'il n'est pas certain qu'un tel minimiseur existe, nous allons cependant faire cette hypothèse et noterons cet estimateur par $\hat{s}$. La qualité de l'estimateur $\hat{s}$ sera évaluée en terme de risque quadratique : i.e. en fonction de

$$\mathbb{E}\big[\|\hat{s}-s\|^2_{\mathbb{H}}\big].$$

Le problème de sélection de modèle consiste alors à choisir un modèle (parmi une famille de modèles) dont l'estimateur $\hat{s}$ a le risque quadratique le plus petit possible. De manière plus précise, nous avons à disposition une famille de modèles $(S_m)_{m\in\mathcal{M}}$ (avec $\mathcal{M}$ un ensemble au plus dénombrable) où, pour tout $m \in \mathcal{M}$, $S_m \subset \mathbb{H}$ est un ensemble convexe et fermé. Pour tout $m \in \mathcal{M}$, nous désignons par $\hat{s}_m \in S_m$ l'estimateur des moindres carrés associé à l'ensemble S_m.

Une sélection de modèle est une procédure utilisant les données observées afin de sélectionner une valeur $\hat{m} \in \mathcal{M}$ et de choisir $\hat{s}_m$ comme estimateur final. Idéalement, le risque associé $\mathbb{E}[\|\hat{s}_m - s\|^2_{\mathbb{H}}]$ est le plus proche possible du risque minimal $\inf_{m\in\mathcal{M}} \mathbb{E}[\|\hat{s}_m - s\|^2_{\mathbb{H}}]$.

Une méthode répandue pour sélectionner un modèle consiste à faire de la minimisation pénalisée. Cela revient à minimiser non pas la fonction $\gamma(t)$ mais plus plutôt une version pénalisée : nous cherchons donc à sélectioner $\hat{m} \in \mathcal{M}$ minimisant

$$\gamma(\hat{s}_m) + \mathrm{pen}(m) \quad \text{pour} \quad m \in \mathcal{M}$$

Remarque. Voici une autre motivation derrière cette minimisation pénalisée : dans de nombreux exemples, nous faisons face à un système de n équations avec p inconnues ; lorsque $p > n$ la solution du problème de minimisation des moindres carrées n'est plus unique, la contrainte de pénalisation permet alors de résoudre ce problème en réduisant le nombre d'inconnues.

Dans ce qui suit, nous présentons la méthode du *Least Absolute Shrinkage and Selection Operator.*

9.3.3 Méthode du LASSO

Dans ce qui suit, nous nous plaçons dans le contexte du modèle linéaire gaussien (généralisé) introduit plus tôt et nous allons nous concentrer sur un exemple particulier.

Considérons $F = \{\phi_1, \dots, \phi_n\} \subset \mathbb{H}$, nous cherchons alors s comme un élément de l'espace vectoriel $G = \mathrm{Vect}(F)$ engendré par la famille (qui n'est pas forcément libre) F. Autrement dit, s est de la forme

$$s = \sum_{i=1}^{n} \lambda_i \phi_i \quad \text{avec, pour tout}\, i = 1, \dots, n, \quad \lambda_i \in \mathbb{R}.$$

L'espace vectoriel G est équipé de la norme l^1 :

$$\|t\|_1 = \inf \left\{ \sum_{i=1}^{n} |\beta_i| \,:\, \beta \in \mathbb{R}^n \quad \text{tel que} \quad \sum_{i=1}^{n} \beta_i \phi_i = t \right\}.$$

La méthode de sélection que nous allons étudié porte le nom de LASSO, dans celle-ci nous allons chercher à minimiser la fonction suivante

$$\gamma(t) + r\|t\|_1 \quad \text{pour tout} \quad t \in G \tag{9.3.2}$$

avec un certain $r > 0$. Le résultat principal est le suivant :

Théorème 9.3.1. *Si $\tilde{s}$ est un minimiseur de (9.3.2) et si le paramètre r est choisit tel que*

$$r \geq 4\epsilon(1 + \sqrt{\log n})$$

alors il existe $C \geq 1$ telle que

$$\mathbb{E}\big[\|\tilde{s} - s\|_{\mathbb{H}}^2\big] \leq C \bigg(\inf_{t \in G} \big(\|s - t\|_{\mathbb{H}}^2 + r\|t\|_1 + r\epsilon \big) \bigg)$$

pour tout $\epsilon > 0$.

Remarque. Le théorème précédent nous assure que l'erreur quadratique de l'estimateur LASSO bruité (par le processus $W(t)$) n'est jamais pire que celle de la version déterministe (sans aléa perturbateur).

La démonstration que nous allons proposer repose beaucoup sur l'observation suivante

$$\gamma(\tilde{s}) + r\|\tilde{s}\|_1 = \inf_{R\geq 0} \inf_{\|t\|_1 \leq R} \big(\gamma(t) + rR\big)$$

avec $\tilde{s}$ l'estimateur LASSO. Afin d'obtenir une famille dénombrable de modèles, nous allons « discrétiser » la famille de boules l_1 en définissant

$$S_m = \big\{t \in G,\ \|t\|_1 \leq m\epsilon\big\} \quad \text{avec} \quad m \geq 1.$$

A présent, si $\hat{m}$ désigne le plus petit entier tel que $\tilde{s} \in S_{\hat{m}}$ nous avons

$$\gamma(\tilde{s}) + r\hat{m}\epsilon \leq \inf_{m\geq 1} \inf_{t\in S_m} \big(\gamma(t) + rm\epsilon\big) + r\epsilon.$$

Autrement dit, $\tilde{s}$ est équivalent à une version approchée d'un estimateur des moindres carrés (pénalisé) sur l'ensemble des modèles $(S_m)_{m\geq 1}$.

Démonstration. Choisissons des poids de la forme $x_m = \delta m$ avec $\delta > 0$ une constante numérique dont la valeur sera choisie ultérieurement. Notons dès maintenant que, pour de tels poids,

$$\sum_{m\geq 1} e^{-x_m} = \frac{e^\theta}{e^\theta - 1} := \Sigma_\theta.$$

- **Construction du processus gaussien adéquat :**

  Pour tout $m \geq 1$, nous désignerons par s_m la projection de s sur S_m, autrement dit

  $$\|s - s_m\| = \inf_{t\in S_m} \|s - t\|$$

  Alors, par définition de $\tilde{s}$, nous avons

  $$\gamma(\tilde{s}) + r\hat{m}\epsilon \leq \gamma(s_m) + rm\epsilon + r\epsilon \quad \text{pour tout} \quad m \geq 1. \tag{9.3.3}$$

  Observons également que, par définition de $Y(t)$, l'identité suivante est vérifiée

  $$\|s\|^2 + \gamma(t) = \|t - s\|^2 - 2\epsilon W(t).$$

  En particulier, si $t = \tilde{s}$ nous avons

  $$\begin{array}{rcl} \|\tilde{s} - s\|^2 & = & \|s\|^2 + \gamma(\tilde{s}) + 2\epsilon W(\tilde{s}) \\ & \leq & \|s\|^2 + \gamma(s_m) + rm\epsilon + r\epsilon - r\hat{m}\epsilon + 2\epsilon W(\tilde{s}) \quad \text{d'après l'inégalité (9.3.3)} \\ & = & \|s - s_m\|^2 + 2\epsilon\big[W(\tilde{s}) - W(s_m)\big] - r\hat{m}\epsilon + rm\epsilon + r\epsilon \end{array}$$

  car $\gamma(s_m) = -2\langle s, s_m\rangle - 2\epsilon W(s_m) + \|s_m\|^2$.

  Pour tout $l \geq 1$, considérons $y_l > 0$ un nombre positif dont la valeur sera choisie plus tard et posons, pour tout $t \in S_l$,

  $$2w_l(t) = \big[\|s - s_m\| + \|s - t\|\big]^2 + y_l^2.$$

Enfin, définissons V_l par

$$V_l = \sup_{t \in S_l} \left[\frac{W(t) - W(s_m)}{w_l(t)} \right]$$

En prenant en compte ces nouvelles définitions, l'inégalité obtenue précédemment prend la forme suivante :

$$\|\tilde{s} - s\|^2 \leq \|s - s_m\|^2 + 2\epsilon w_{\hat{m}}(\tilde{s}) V_{\hat{m}} - r\hat{m}\epsilon + rm\epsilon + r\epsilon. \tag{9.3.4}$$

- **Utilisation de l'inégalité de concentration** (9.2.3) **:**

Le reste de la démonstration consiste à étudier les fluctuations du supremum gaussien V_l à l'aide de l'inégalité de concentration (9.2.3). Celle-ci, appliquée à $(V_l)_{l \geq 1}$ nous assure que

$$\mathbb{P}\big(V_l \leq \mathbb{E}[V_l] + r\big) \leq e^{-\frac{r^2}{2v_l}} \quad r \geq 0 \tag{9.3.5}$$

avec

$$v_l = \sup_{t \in V_l} \mathrm{Var}\left(\frac{W(t) - W(s_m)}{w_l(t)} \right) = \sup_{t \in S_l} \frac{\|t - s_m\|^2}{w_l^2(t)}$$

puisque le processus $W(t)$ est isonormal. En outre, si nous procédons au changement de variable $r = \sqrt{2v_l(x_l + z)}$, l'équation (9.3.5) devient

$$\mathbb{P}\big(V_l \geq \mathbb{E}[V_l] + \sqrt{2v_l(x_l + z)}\big) \leq e^{-x_l} e^{-z} \quad \text{avec} \quad z > 0 \text{ et } l \geq 1. \tag{9.3.6}$$

En outre, puisque $w_l(t) \geq \big(\|s - s_m\| + \|s - t\|\big) y_l \geq \|t - s_m\| y_l$, nous pouvons contrôler le supremum des variances :

$$i.e. \quad v_l \leq y_l^{-2}.$$

Ainsi, en additionnant, pour $l \geq 1$, les inégalités (9.3.6) nous obtenons (pour tout $z > 0$) un évènement Ω_z tel que

$$\mathbb{P}(\Omega_z) \geq 1 - \Sigma e^{-z}.$$

Ainsi, pour tout $l \geq 1$, sur Ω_z nous avons

$$V_l \leq \mathbb{E}[V_l] + y_l^{-1} \sqrt{2(x_l + z)} \tag{9.3.7}$$

- **Contrôle de l'espérance** $\mathbb{E}[V_l]$ **et de** V_l :

Soit $l \geq 1$, pour contrôler $\mathbb{E}[V_l]$ nous allons utiliser l'inégalité suivante

$$\mathbb{E}[V_l] \leq \mathbb{E}\left[\sup_{t \in S_l} \frac{W(t) - W(s_l)}{\inf_{t \in S_l} w_l(t)} \right] + \mathbb{E}\left[\frac{\big(W_{s_l} - W(s_m)\big)_+}{\inf_{t \in S_l} w_l(t)} \right]$$

Il est alors possible de montrer (cf. [33] pour les détails techniques) que, pour tout $l \geq 1$, nous ayons (sur l'évènement Ω_z)

$$\epsilon V_l \leq K^{-1/2}$$

avec $K = \frac{4\sqrt{2}}{5}$. Cela suppose, bien entendu, de choisir la valeur de δ puis de y_l judicieusement. En particulier, ce qui précède implique donc la majoration suivante (valable sur Ω_z)

$$\epsilon V_{\hat{m}} \leq K^{-1/2}.$$

- **Conclusion :**

En combinant ce qui précède avec (9.3.4), il est alors possible de montrer (après un certain nombre de calculs et en choisissant $\theta = \frac{1-\sqrt{\log 2}}{K}$) que sur l'évènement Ω_z nous avons

$$\begin{aligned} \|\tilde{s}-s\|^2 &\leq A'\|s-s_m\|^2 + K^{-1/4}\|s-\tilde{s}\|^2 \\ &+ 4m\epsilon^2(\sqrt{\log N}+1) - r\hat{m}\epsilon \\ &+ rm\epsilon + r\epsilon + \frac{4K\epsilon^2}{\sqrt{K}-1}\left(\frac{1}{2\pi}+2z\right) \end{aligned}$$

avec $A' > 0$ une constante dépendant de K. Il faut ensuite utiliser l'hypothèse

$$r \geq 4\epsilon(1+\sqrt{\log N})$$

afin d'obtenir (après d'autres calculs et une intégration en z) l'inégalité suivante

$$\begin{aligned} \mathbb{E}\big[\|\tilde{s}-s\|^2\big] &\leq C(K)\left[\inf_{m\geq 1}\left(\inf_{\|t\|_1\leq m\epsilon}\|s-t\|^2 + rm\epsilon + r\epsilon\right) + (1+\Sigma_\theta)\epsilon^2\right] \\ &\leq C(K)\left[\inf_{t\in G}\big(\|s-t\|^2 + r\|t\|_1\big) + 2r\epsilon + (1+\Sigma_\theta)\epsilon^2\right]. \end{aligned}$$

ce qui conclut la preuve.

$\square$

Remarque. A un moment donné, dans la démonstration, il est nécessaire de contrôler $\mathbb{E}[\sup_{t\in S_l} W(t)]$. Or, dans le contexte précédent

$$\sup_{t\in S_l} W(t) \leq m\epsilon \max_{i=1,\ldots,N}|W(\phi)|$$

Puisque les variables $W(\phi)$, $i = 1,\ldots,N$ sont des variables aléatoires gaussiennes standards nous avons $\mathbb{E}\big[\max_{i=1,\ldots,n}|W(\phi)|\big] \leq \sqrt{2\log 2N}$. L'argument repose sur le fait suivant : si $g_1,\ldots,g_n$ désigne des variables aléatoires gaussiennes (non nécessairement indépendantes mais de même variance) alors

$$\begin{aligned} \mathbb{E}[\max_{i=1,\ldots,n}|g_i|] &= \frac{1}{\beta}\mathbb{E}\big[\log e^{\beta\max_{i=1,\ldots,n}|g_i|}\big] \\ &\leq \frac{1}{\beta}\mathbb{E}\big[\log\sum_{i=1}^n e^{\beta|g_i|}\big] \\ &\leq \frac{1}{\beta}\log\sum_{i=1}^n\mathbb{E}\big[e^{\beta g_i}+e^{-\beta g_i}\big] \quad \text{par l'inégalité de Jensen} \\ &= \frac{\beta}{2} + \frac{\log 2n}{\beta} \end{aligned}$$

Il ne reste plus qu'à optimiser en choisissant $\beta = \sqrt{2\log 2n}$. Ceci expliquant la présence du terme logarithmique dans la démonstration précédente.

Dans la section suivante nous allons développer des méthodes qui permettent d'établir simplement les inégalités fonctionnelles que nous venons de rencontrer.

9.4 Démonstrations par semi-groupes et méthodes d'interpolation

Cette section a pour objet de proposer des démonstrations élémentaires des inégalités fonctionnelles présentées précédemment. Nous nous inspirons de la présentation faite dans [76, 16], à laquelle nous renvoyons le lecteur pour plus de détails. Dans un premier temps, nous introduirons quelques notations provenant de la théorie de Bakry-Émery afin de mettre en place les méthodes d'interpolation via l'utilisation de semi-groupes.

9.4.1 Notations

Considérons un espace mesurable $(E, \mathcal{F})$ equipé d'une mesure σ-finie μ. Lorsque μ est finie, nous la renormaliserons afin d'avoir une mesure de probabilité. Notons par $L^p = L^p(\mu)$, $1 \leq p \leq \infty$, les espaces de Lebesgue par rapport à la mesure μ, et nous poserons $\|\cdot\|_p$ pour désigner la norme L^p.

L'objet d'intérêt fondamental est la famille $(P_t)_{t\geq 0}$ d'opérateurs positifs agissant sur l'ensemble des fonctions bornées f sur E par

$$P_t f(x) = \int_E f(y) p_t(x, dy), \quad x \in E,$$

et satisfaisant la propriété dite de semi-groupe :

$$P_s \circ P_t = P_{s+t},\ s, t \geq 0,\ P_o = Id.$$

Les noyaux positifs $p_t(x, dy)$ sont appelés noyaux de transition. Nous supposerons toujours que les opérateurs P_t sont bornés et continus dans $L^2(\mu)$ au sens où, pour tout f dans $L^2(\mu)$, $\|P_t f\|_2 \leq C(t)\|f\|_2$ pour tout $t \geq 0$ et $\|P_t f - f\|_2 \to 0$ lorsque $t \to 0$.

Nous dirons que $(P_t)_{t\geq 0}$ est markovien si $P_t 1 = 1$ pour tout $t \geq 0$. Ces semi-groupes markoviens sont naturellement associés à des processus de Markov $(X_t)_{t\geq 0}$, à valeurs dans E, par la relation

$$P_t f(x) = \mathbb{E}[f(X_t)|X_0 = x], \quad x \in E$$

Le mouvement brownien $(B_t)_{t\geq 0}$ à valeurs dans $\mathbb{R}^n$, issu de l'origine avec pour noyau (de la chaleur)

$$p_t(x, dy) = \frac{1}{(2\pi t)^{n/2}} e^{-|x-y|^2/2t} dy, \quad t \geq 0, \quad x \in \mathbb{R}^n,$$

est un exemple classique de tels semi-groupes (il s'agit notamment de celui que nous avons employé au chapitre 5).

Néanmoins, nous travaillerons la plupart du temps (pour une raison qui deviendra évidente par la suite) avec le processus d'Ornstein-Uhlenbeck $(X_t)_{t\geq 0}$ qui peut être représenté comme le processus gaussien suivant

$$X_t = \sqrt{2}e^{-t}\int_0^t e^s dB_s,\ t\geq 0.$$

Le semi-groupe associé peut être représenté par la formule, dite de Mehler, suivante :

$$P_tf(x) = \int_{\mathbb{R}^n} f\big(e^{-t}x + (1-e^{-2t})^{1/2}y\big)d\gamma_n(y), \quad x\in\mathbb{R}^n, \quad t\geq 0. \tag{9.4.1}$$

Un objet essentiel lié aux semi-groupes est leur générateur infinitésimal L. Notons par $\mathcal{D}(L)$ le domaine dans $L^2(\mu)$ du générateur infinitésimal du semi-groupe $(P_t)_{t\geq 0}$, celui-ci étant défini comme l'ensemble des fonctions f de $L^2(\mu)$ pour lesquelles la limite suivante existe

$$Lf = \lim_{t\to 0}\frac{1}{t}(P_tf - f).$$

Le semi-groupe $(P_t)_{t\geq 0}$ laisse le domaine $\mathcal{D}(L)$ stable, et pour toute fonction $f\in\mathcal{D}(L)$, P_tf satisfait l'équation (dite de la chaleur) associée à L suivante,

$$\partial_t P_tf = P_tLf = LP_tf. \tag{9.4.2}$$

Réciproquement, le générateur L et son domaine $\mathcal{D}(L)$ détermine complètement $(P_t)_{t\geq 0}$: il existe un unique semi-groupe $(P_t)_{t\geq 0}$ d'opérateurs bornés sur $L^2(\mu)$ satisfaisant (9.4.2) pour toutes les fonctions du domaine $\mathcal{D}(L)$ [123].

Par exemple, le semi-groupe du mouvement brownien a pour générateur un demi du laplacien Δ sur $\mathbb{R}^n$ et dans ce cas particulier (9.4.2) est l'équation de la chaleur classique (que nous avons utilisé dans le chapitre 5). Dans notre contexte, avec le semi-groupe d'Ornstein-Uhlenbeck, nous avons le générateur suivant $L = \Delta - x\cdot\nabla$ qui peut être vu comme le laplacien de l'espace gaussien $(\mathbb{R}^n,\gamma)$ (on parle dans ce cas de Witten-Laplacien). L'équation de la chaleur associée à $L = \Delta - x\cdot\nabla$ correspond à une formulation probabiliste de l'équation, plus usuelle en théorie des dérivées partielles, de Fokker-Planck linéaire (cf. [122, 121]).

Certaines mesures μ auront un rôle particulier vis-à-vis du semi-groupe $(P_t)_{t\geq 0}$ lorsque celles-ci vérifient des propriétés d'invariance et de reversibilité. Plus précisément une mesure μ est dite réversible (en temps) par rapport à $(P_t)_{t\geq 0}$ si pour tout $f,g\in L^2(\mu)$,

$$\int_E fP_tg d\mu = \int_E gP_tf d\mu.$$

μ est invariante par rapport à $(P_t)_{t\geq 0}$ si pour toute fonction $f\in L^1(\mu)$,

$$\int_E P_tf d\mu = \int_E f d\mu.$$

Lorsque μ est finie, on peut choisir $g=1$ pour constater que les mesures reversibles sont des mesures invariantes. Les mesures reversibles, respectivement invariantes, μ sont décrites de manière équivalentes comme celles pour lesquelles $\int fLg d\mu = \int gLf d\mu$, respectivement $\int Lf d\mu = 0$, pour tout $f, g \in \mathcal{D}(L)$.

Concernant le semi-groupe d'Ornstein-Uhlenbeck, sa mesure invariante et réversible est la mesure gaussienne standard γ_n sur $\mathbb{R}^n$ (il s'agit de la raison principale qui permet de considérer ce semi-groupe dans un cadre gaussien). Etant donné un générateur infinitésimal L, nous pouvons définir $\mathcal{E}$, l'énergie de Dirichlet du semi-groupe de Markov P_t, par

$$\mathcal{E}_\mu(f,g) = \int_E f(-Lg)d\mu.$$

Comme nous pouvons le vérifier dans le cas du semi-groupe d'Ornstein-Uhlenbeck, après une intégration par partie en espace, nous obtenons l'expression suivante :

$$\mathcal{E}_{\gamma_n}(f,g) = \int_{\mathbb{R}^n} \nabla f \cdot \nabla g d\gamma_n.$$

Lorsque μ est une mesure de probabilité, nous dirons que $(P_t)_{t\geq 0}$ est ergodique si $P_t f \to \int_E f d\mu$ μ-presque sûrement lorsque $t \to \infty$.

9.4.2 Inégalités fonctionnelles par interpolation

Une des forces de la théorie de Bakry-Émery et des méthodes d'interpolation, est qu'elle permet de démontrer de manière élémentaire et synthétique les inégalités de Poincaré et de Sobolev logarithmique. Nous décrirons plus loin comment les résultats obtenus pour la mesure gaussienne peut s'étendre facilement à un cadre plus général ; certains aspects de la théorie seront éludés et nous invitons le lecteur à se référer aux livres suivants [16, 18, 76]. Nous énonçons ci-dessous une proposition, regroupant les propriétés fondamentales satisfaites par le semi-groupe d'Ornstein-Uhlenbeck, celles-ci sont essentielles pour les démonstrations. Nous démontrerons ensuite l'inégalité de Poincaré, puis celle de Sobolev logarithmique pour la mesure gaussienne γ_n sur $\mathbb{R}^n$.

Proposition 9.4.1. *Le semi-groupe d'Ornstein-Uhlenbeck $(P_t)_{t\geq 0}$ admet pour mesure invariante la mesure gaussienne γ_n sur $\mathbb{R}^n$ et vérifie les assertions suivantes :*

1. *$(P_t)_{t\geq 0}$ est ergodique.*

2. *Les inégalités ci-dessous, sont valables pour tout $x \in \mathbb{R}^n$ que nous omettrons pour alléger les notations. $(P_t)_{t\geq 0}$ satisfait une relation de commutation exacte avec l'opérateur ∇, c'est à dire, pour toute fonction f suffisamment régulière,*

$$\nabla P_t f = e^{-t} P_t \nabla f, \quad t \geq 0. \tag{9.4.3}$$

En particulier, ceci fournit l'inégalité suivante, valable pour tout $t \geq 0$ et toute fonction f suffisamment régulière,

$$|\nabla P_t f|^2 = e^{-2t} |P_t(\nabla f)|^2 \leq e^{-2t} P_t(|\nabla f|^2). \tag{9.4.4}$$

L'inégalité de Cauchy-Schwarz implique également, pour tout $t \geq 0$ et $f : \mathbb{R}^n \to \mathbb{R}_+$ suffisamment régulière,

$$|\nabla P_t f|^2 \leq P_t\left(\frac{|\nabla f|^2}{f}\right) P_t(f) \tag{9.4.5}$$

3. *En considérant $(P_t)_{t\geq 0}$ comme une famille d'opérateurs de $L^p(\gamma_n)$, $p \geq 1$, dans lui-même. L'inégalité de Jensen nous affirme que ces opérateurs sont des contractions de L^p, autrement dit pour toute fonction $f \in L^p(\gamma_n)$ et tout $t \geq 0$,*

$$\|P_t f\|_p \leq \|f\|_p.$$

Remarque. Nous verrons plus tardivement la signification de la constante $\rho = 1$ dans l'identité $\nabla P_t f = e^{-\rho t} P_t \nabla f$ dans le cas du semi-groupe d'Ornstein-Uhlenbeck alors que nous avions $\rho = 0$ dans le cas de semi-groupe associé au mouvement brownien (cf. chapitre 5).

Expliquons maintenant de quelle manière obtenir les inégalités (9.4.5) et (9.4.4).

Inégalité de Poincaré

Nous débutons par le cas le plus simple : l'inégalité de Poincaré.

Proposition 9.4.2. *La mesure gaussienne standard γ_n sur $\mathbb{R}^n$ satisfait l'inégalité suivante, pour f suffisamment régulière,*

$$\mathrm{Var}_{\gamma_n}(f) \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n.$$

Démonstration. Par ergodicité du semi-groupe, puis avec le théorème fondamental de l'analyse nous obtenons le développement de la variance le long du semi-groupe suivant :

$$\begin{aligned}
\mathrm{Var}_{\gamma_n}(f) &= \lim_{s\to\infty} \int_{\mathbb{R}^n} (P_0 f)^2 d\gamma_n - \left(\int_{\mathbb{R}^n} P_s f d\gamma_n\right)^2 && (9.4.6)\\
&= -\int_0^\infty \frac{d}{ds} \int_{\mathbb{R}^n} (P_s f)^2 d\gamma_n ds && (9.4.7)\\
&= -2\int_0^\infty \int_{\mathbb{R}^n} P_s f L(P_s f) d\gamma_n ds && (9.4.8)\\
&= 2\int_0^\infty \int_{\mathbb{R}^n} |\nabla P_s f|^2 d\gamma_n ds && (9.4.9)\\
&= 2\int_0^\infty e^{-2s} \int_{\mathbb{R}^n} |P_s(\nabla f)|^2 d\gamma_n ds && (9.4.10)
\end{aligned}$$

La dernière ligne est obtenue après une intégration par partie en espace, puis, en utilisant la relation (9.4.3) . Pour conclure, il suffit d'utiliser (9.4.4), le fait que le semi-groupe est une contraction des espaces $L^p(\gamma_n)$ et l'invariance de γ_n par rapport à $(P_t)_{t\geq 0}$.

$$\begin{aligned}
2\int_0^\infty \int_{\mathbb{R}^n} |\nabla P_s f|^2 d\gamma_n ds &= 2\int_0^\infty e^{-2s} \int_{\mathbb{R}^n} |P_s(\nabla f)|^2 d\gamma_n ds\\
&\leq 2\int_0^\infty e^{-2s} \int_{\mathbb{R}^n} P_s(|\nabla f|^2) d\gamma_n ds\\
&= \left(2\int_0^\infty e^{-2s} ds\right)\left(\int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n\right)
\end{aligned}$$

□

Remarque. Notons que la démonstration ci-dessus, fournit une représentation dynamique de la variance, le long du semi-groupe, qui sera au centre de certains résultats de superconcentration.

$$\mathrm{Var}_{\gamma_n}(f) = 2\int_0^\infty e^{-2s}\int_{\mathbb{R}^n} |P_s(\nabla f)|^2 d\gamma_n ds. \tag{9.4.11}$$

Mentionnons également que, dans le cadre de la théorie de Bakry-Émery, l'inégalité de Poincaré peut s'exprimer, de manière équivalente par un critère de « courbure dimension intégré ». Pour la mesure gaussienne γ_n, il s'agit de l'équivalence suivante, pour $f \in \mathcal{D}(L)$,

$$\mathbb{E}_{\gamma_n}[\Gamma_2(f)] \geq \mathbb{E}_{\gamma_n}[\Gamma_1(f)] \quad \Longleftrightarrow \mathrm{Var}_{\gamma_n}(f) \leq \mathbb{E}_{\gamma_n}[\Gamma_1(f)], \tag{9.4.12}$$

avec $\Gamma_1(f) = |\nabla f|^2$ et $\Gamma_2(f) = \|\mathrm{Hess} f\|_2^2 + |\nabla f|^2$.

Dans un certain sens, la constante $C_P = 1$ de l'inégalité de Poincaré gaussienne dépend de la « pire » fonction de l'ensemble $\mathcal{D}(L)$. Il est alors naturel de se demander s'il est possible d'exhiber un ensemble plus petit $\mathcal{A} \subset \mathcal{D}(L)$ sur lequel une inégalité de Poincaré serait satisfaite avec une constante $\rho_{\mathcal{A}} < 1$. Nous verrons par la suite (dans le chapitre traitant de la superconcentration), via une inégalité de Talagrand, que si $\mathcal{A}$ est constitué des statistiques d'ordres d'un échantillon gaussien de taille $n \geq 2$, on peut choisir $\rho_{\mathcal{A}} = C/\log n$.

Inégalité de Sobolev logarithmique

Comme nous allons le voir, ce type d'argument permet, de manière analogue, de démontrer une inégalité de Sobolev logarithmique.

Proposition 9.4.3. *Rappelons que γ_n satisfait une inégalité de Sobolev logarithmique, c'est-à-dire, pour toute fonction f suffisament régulière,*

$$\mathrm{Ent}_{\gamma_n}(f^2) \leq 2\int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n.$$

Démonstration. En procédent de manière analogue à la démonstration de l'inégalité de Poincaré, on obtient le développement suivant de l'entropie d'une fonction f le long du semi-groupe d'Ornstein-Uhlenbeck :

$$\mathrm{Ent}_{\gamma_n}(f) = \int_0^\infty \int_{\mathbb{R}^n} \frac{|\nabla P_t f|^2}{P_t f} d\gamma_n dt \tag{9.4.13}$$

Il suffit ensuite d'utiliser l'item (9.4.5) et l'invariance du semi-groupe par rapport à la mesure gaussienne pour obtenir

$$\begin{aligned} \mathrm{Ent}_{\gamma_n}(f) &\leq \int_0^\infty \int_{\mathbb{R}^n} e^{-2t} P_t\left(\frac{|\nabla f|^2}{f}\right) d\gamma_n dt \\ &= \frac{1}{2}\int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} d\gamma_n. \end{aligned}$$

Le résultat s'ensuit en remplaçant f par f^2.

□

Il est possible de formuler l'inégalité de Sobolev logarithmique comme une décroissance exponentielle de l'entropie le long du semi-groupe.

Proposition 9.4.4. *L'inégalité de Sobolev logarithmique satisfaite par la mesure gaussienne est équivalente à l'inégalité suivante*

$$H(\nu_t|\gamma_n) \leq e^{-2t} H(\nu_0|\gamma_n) \quad \textit{pour tout} \quad t \geq 0$$

avec $d\nu_t = P_t(f)d\mu$ *pour tout* $t \geq 0$.

Remarque. Il est possible de démontrer un résultat similaire, au niveau de la variance, lorsqu'une inégalité de Poincaré est satisfaite.

9.4.3 Aparté : convergence vers la mesure invariante

Cette décroissance entropique est particulièrement utile lorsqu'elle est combinée à l'inégalité de Pinsker-Cizsar-Kullback. Cette inégalité que nous donnons ci-dessus n'est pas spécifique à la mesure gaussienne et reste valable pour toutes mesures de probabilité μ.

Proposition 9.4.5 (Pinsker)**.** *Soit* μ *une mesure de probabilité alors*

$$\|\mu - \nu\|_{TV}^2 \leq \frac{1}{2} H(\nu, \mu) \quad \textit{pour toutes mesures de probabilité} \quad \nu \tag{9.4.14}$$

avec $\|\mu - \nu\|_{TV} = \sup_{A \in \mathcal{F}} |\mu(A) - \nu(A)|$ *la distance en variation totale entre* μ *et* ν.

Démonstration. Si $\frac{d\nu}{d\mu} = f \geq 0$ alors $\int_E |1 - \frac{d\nu}{d\mu}| d\mu = 2\|\mu - \nu\|_{TV}$. Ainsi, l'inégalité de Pinsker est équivalente à

$$\Big(\int_E |1 - f| d\mu\Big)^2 \leq 2\mathrm{Ent}_\mu(f).$$

Posons alors $f_s = sf + 1 - s$ pour tout $s \in [0,1]$ et étudions les variations de la fonction $\Psi(s) = 2\mathrm{Ent}_\mu(f_s) - \big(\int_E |1 - f_s| d\mu\big)^2$. Il n'est pas difficile de montrer que

$$\Psi(s) = 2\mathrm{Ent}_\mu(f_s) - s^2\Big(\int_E |1 - f| d\mu\Big)^2$$

d'où $\Psi(0) = \Psi'(0) = 0$ et

$$\Psi''(s) = 2\int_E \frac{(1-f)^2}{f_s} d\mu - 2\Big(\int_E |1 - f| d\mu\Big)^2 \quad \text{avec} \quad s \in [0,1].$$

Alors, d'après l'inégalité de Cauchy-Schwarz $\Psi''(0) \geq 0$; en conséquence, $\Psi(s) \geq 0$ pour tout $s \in [0,1]$. En particulier, $\Psi(1) \geq 0$ ce qui achève la démonstration. $\square$

Ainsi, s'il est possible de contrôler $H(\nu_0|\mu)$ l'inégalité de Sobolev logarithmique (dans sa version entropique (9.4.13)) combinée à (9.4.14) permet d'obtenir un contrôle en variation totale sur la convergence de $(\nu_t)_{t \geq 0}$ vers μ.

9.4.4 Hypercontractivité

Introduisons une nouvelle identité permettant de relier l'entropie à l'information de Fisher. Cette nouvelle quantité, provenant de la théorie de l'information, n'est qu'une nouvelle formulation de l'energie de Dirichlet d'une fonction après avoir procédé à un changement de variable (en remplaçant f par $\sqrt{f}$) dans une inégalité de Sobolev logarithmique. Ce faisant, le membre de droite de l'inégalité (9.1.3) devient

$$\int_E \frac{|\nabla f|^2}{f} d\mu$$

et nous désignerons cette nouvelle quantité par $I_\mu(f)$, l'information de Fisher. La proposition suivante montre qu'il est possible de relier cette quantité avec la dérivée de l'entropie (le long du semi-groupe).

Proposition 9.4.6 (De Bruijn)**.** *Pour toutes fonctions f suffisamment régulières*

$$\frac{d}{dt}\mathrm{Ent}_\mu(P_t f) = -I_\mu(P_t f). \tag{9.4.15}$$

Démonstration. En utilisant la représentation de l'entropie le long du semi-groupe et les propriétés liées à l'opérateur L (intégration par partie,...), nous avons

$$\begin{aligned} \frac{d}{dt}\mathrm{Ent}_\mu(P_t f) &= \int_E (1+\log P_t f) L P_t f d\mu \\ &= -\int_E \nabla P_t f \cdot \nabla(1+\log P_t f) d\mu \\ &= -\int_E \frac{|\nabla P_t f|^2}{P_t f} d\mu = -I_\mu(P_t f) \end{aligned}$$

□

Remarque. L'identité de Bruijn, permet d'établir l'équivalence entre inégalité de Sobolev logarithmique et décroissance entropique (9.4.13). Dans le cas gaussien, l'inégalité de Sobolev logarithmique est équivalente à

$$\mathrm{Ent}_{\gamma_n}(f) \le I_{\gamma_n}(f) \quad \Longleftrightarrow \quad \Psi(t) \le -\Psi'(t) \quad \text{pour tout} \quad t \ge 0$$

avec $\Psi(t) = \mathrm{Ent}_{\gamma_n}(P_t f)$. L'implication réciproque s'obtient en dérivant l'inégalité précédente en $t = 0$.

Comme signifié auparavant, il est intéressant de considérer un semi-groupe $(P_t)_{t\ge 0}$ comme une famille d'opérateurs de $L^p(\mu), p \ge 1$, dans lui-même, où μ est la mesure invariante du semi-groupe. L'inégalité de Jensen nous affirme que ces opérateurs sont des contractions de $L^p(\mu)$, autrement dit pour toute fonction $f \in L^p(\mu)$ et tout $t \ge 0$,

$$\|P_t f\|_p \le \|f\|_p$$

Parfois, le semi-groupe peut vérifier une propriété plus forte dite d'hypercontractivité. C'est le contenu de la définition ci-dessous.

Définition 9.4.1. *Considérons un semi-groupe $(P_t)_{t\geq 0}$ et sa mesure invariante μ. On dit que le semi-groupe $(P_t)_{t\geq 0}$ est hypercontractif si, pour tout $t \geq 0$, il existe des nombres $q > p > 1$ (pouvant dépendre de t) tels que pour tout $f \in Ł^q(\mu)$:*

$$\|P_t f\|_q \leq \|f\|_p.$$

Lorsque la mesure invariante μ satisfait une inégalité de Sobolev logarithmique, cela induit une propriété plus forte sur le semi-groupe sous-jacent. Il s'agit du théorème de Gross [64] qui démontre l'équivalence entre l'hypercontractivité du semi-groupe et le fait que la mesure invariante satisfasse une inégalité de Sobolev logarithmique. Cette observation est survenue la première fois dans la théorie quantique des champs.

Nous énonçons ce résultat dans le cas de la mesure gaussienne γ_n sur $\mathbb{R}^n$ et du semi-groupe d'Ornstein-Uhlenbeck associé.

Théorème 9.4.1 (Nelson)**.** *Soient $1 < p < q < \infty$ alors les assertions suivantes sont satisfaites.*

- *Si $q-1 \leq e^{2t}(p-1)$, alors l'opérateur P_t de Ornstein-Uhlenbeck est hypercontractif de $L^p(\gamma_n)$ dans $L^q(\gamma_n)$ C'est à dire que pour toutes fonctions $f \in \mathcal{D}(L)$, on a :*

$$\|P_t f\|_q \leq \|f\|_p. \tag{9.4.16}$$

- *Si $q-1 > e^{2t}(p-1)$, alors l'opérateur P_t n'est pas continu de $L^p(\gamma_n)$ dans $L^q(\gamma_n)$.*

Démonstration. Sans que cela complique la démonstration, nous allons traiter un cas plus général. Plus précisément, nous supposons que $d\mu = e^{-V}dx$ satisfait une inégalité de Sobolev logarithmique de constante $C > 0$: pour toutes fonctions $f : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulières

$$i.e. \quad \mathrm{Ent}_\mu(f^2) \leq 2C \int_{\mathbb{R}^n} |\nabla f|^2 d\mu$$

A toute fin utile, rappelons que μ est la mesure invariante et réversible de l'opérateur $L = \Delta - \nabla V \cdot \nabla$ et désignons le semi groupe associé par $P_t = e^{tL}$ pour $t \geq 0$. vérifie l'équation de la chaleur associée à L : i.e. $\partial_t P_t = LP_t$. De plus, une intégration par partie est satisfaite : pour toutes fonctions $f, g : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulières, $\int_{\mathbb{R}^n} f(-Lg)d\mu = \int_{\mathbb{R}^n} \nabla f \cdot \nabla g d\mu$.

Soit $p > 1$ et considérons une fonction $t \mapsto q(t)$ telle que $q(0) = p$. Posons ensuite $\Psi(t) = \|P_t f\|_{q(t)}^{q(t)}$ et étudions la fonction $H(t) = \frac{1}{q(t)} \log\big(\Psi(t)\big)$. Nous allons montrer que si μ vérifie une inégalité de Sobolev logarithmique alors $t \mapsto H(t)$ est une fonction décroissante. En particulier

$$H(t) \leq H(0) \quad \Longleftrightarrow \quad \|P_t f\|_{q(t)} \leq \|f\|_p$$

Observons que

$$H'(t) = -\frac{q'(t)}{q^2(t)} \log\big(\Lambda(t)\big) + \frac{1}{q(t)} \frac{\Psi'(t)}{\Psi(t)}$$

De plus

$$\begin{aligned}\Psi'(t) &= \frac{d}{dt}\int_{\mathbb{R}^n}(P_tf)^{q(t)}d\mu = \int_{\mathbb{R}^n}\big[q'(t)\log P_tf + q(t)\partial_tP_t\big]\times(P_tf)^{q(t)}qt\\ &= q'(t)\int_{\mathbb{R}^n}(P_tf)^{q(t)}\log P_tfd\mu - q(t)\big(q(t)-1\big)\int_{\mathbb{R}^n}(P_tf)^{q(t)-2}|\nabla P_tf|^2d\mu\end{aligned}$$

où, pour obtenir la dernière égalité, nous avons utilisé le fait que le semi groupe $(P_t)_{t\geq 0}$ vérifie l'équation de la chaleur associée à L : i.e. $\partial_tP_t = LP_t$; de plus, une intégration par partie est satisfaite : pour toutes fonctions $f,g : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulières, $\int_{\mathbb{R}^n} f(-Lg)d\mu = \int_{\mathbb{R}^n}\nabla f\cdot\nabla gd\mu$.

Tout ceci permet d'exprimer $\Psi'(t)$ sous la forme

$$\Psi'(t) = \frac{q'(t)}{q(t)}\big[\mathrm{Ent}_\mu(g^2) + \Psi(t)\log\Psi(t)\big] - q(t)\big(q(t)-1\big)\int_{\mathbb{R}^n}(P_tf)^{q(t)-2}|\nabla P_tf|^2d\mu$$

avec $g^2 = (P_tf)^{q(t)}$. En outre, d'après l'inégalité de Sobolev logarithmique, nous avons

$$\mathrm{Ent}_\mu(g^2) \leq \frac{C}{2}q^2(t)\int_{\mathbb{R}^n}(P_tf)^{q(t)-2}|\nabla P_tf|^2d\mu.$$

Ainsi,

$$H'(t) \leq \frac{C}{2}\frac{q'(t)}{\Lambda(t)}\int_{\mathbb{R}^n}(P_tf)^{q(t)-2}|\nabla P_tf|^2d\mu - \frac{q(t)-1}{\Psi(t)}\int_{\mathbb{R}^n}(P_tf)^{q(t)-2}|\nabla P_tf|^2d\mu$$

C'est pourquoi

$$H'(t)\leq 0 \quad\Longleftrightarrow\quad \frac{C}{2}q'(t) - q(t) + 1 \leq 0 \quad\Longleftrightarrow\quad q'(t) \leq \frac{2}{C}\big[q(t)-1\big].$$

Autrement dit, en résolvant l'inégalité différentielle précédente,

$$H'(t)\leq 0 \quad\Longleftrightarrow\quad \frac{q(t)-1}{p-1} \leq e^{\frac{2}{C}t}$$

Puisque dans le cas gaussien $C = 1$, cela conclut la démonstration. Réciproquement, il suffit de dériver en $t = 0$ l'inégalité $\|P_tf\|_{q(t)} \leq \|f\|_p$ pour obtenir une inégalité de Sobolev logarithmique de constante $C > 0$. □

A noter que ce résultat (9.4.16) ne dépend absolument pas de la dimension n. Ceci fait partie d'une des remarquables propriétés qui fait de l'hypercontractivité un outil puissant. Nous donnerons dans la section suivante une application de l'hypercontractivité pour obtenir une inégalité de Talagrand [108]. Cette inégalité, qui améliore l'inégalité de Poincaré, sera l'un des outils majeurs dans l'étude de la superconcentration.

9.5 Inégalité courbure-dimension

Cette partie montre comment étendre les démonstrations d'inégalités fonctionnelles par interpolation à des exemples pour lesquels le semi-groupe n'admet pas de formule explicite. Soulignons que de nombreuses propriétés du semi-groupe d'Ornstein-Uhlenbeck pouvaient être obtenues grâce à la formule de Mehler (9.4.1) ; en général, une telle expression n'est pas disponible. Nous verrons alors que les inégalités de courbure dimension provenant de la théorie de Bakry-Émery sont le substitut adéquat permettant de reproduire le principe des démonstrations développées plus tôt.

9.5.1 Introduction

Pour plus de détails, nous renvoyons le lecteur vers [16, 18] ou le cours [76].

Suivant Meyer, Bakry et Emery ont introduit les opérateurs « carré du champ » et leurs itérés de la manière suivante :

$\Gamma_0(f,g) = fg$ et $\Gamma_n(f,g) = \frac{1}{2}\Big[L\Gamma_{n-1}(f,g) - \Gamma_{n-1}(f,Lg) - \Gamma_{n-1}(Lf,g)\Big]$ pour $n \geq 1$. Afin d'alléger les notations, nous utiliserons $\Gamma_n(f)$ pour désigner $\Gamma_n(f,f)$, $n \geq 0$.

Certaines de ces formes bilinéaires symétriques jouent un rôle important pour obtenir, de manière élémentaire, des inégalités fonctionnelles. Par exemple, une inégalité du type $\Gamma_2 \geq \rho\Gamma_1$ avec $\rho > 0$ permet d'obtenir facilement, lorsqu'une hypothèse supplémentaire d'ergodicité est satisfaite, une inégalité de Sobolev logarithmique de constante $1/\rho$ pour la mesure invariante d'une diffusion (cf. [16] pour plus de détails). Dans la littérature, ce genre de relation, entre Γ_2 et Γ_1, est appelée : une inégalité de courbure-dimension $CD(\rho,\infty)$.

Définition 9.5.1. *Un couple (L,μ) satisfait une inégalité de courbure dimension infinie si, pour tout $f \in \mathcal{D}(L)$ nous avons*

$$\Gamma_2(f) \geq \rho\Gamma(f)$$

avec $\rho \in \mathbb{R}$. Nous noterons ceci de manière abrégé en disant que (L,μ) vérifie un critère $CD(\rho,+\infty)$.

Remarque.

1. Nous avons déjà briévement mentionné ce genre de résultat dans le chapitre un, en présentant, de manière équivalente, l'inégalité de Poincaré satisfaite par la mesure gaussienne, en terme d'inégalité entre Γ_1 et Γ_2 (cf. (9.4.12)).

2. Les mesures de la forme $d\mu = e^{-V}dx$ tel que le potentiel V soit strictement convexe (i.e. il existe $\rho > 0$ tel que $\nabla\nabla V \geq \rho I_d$) sont invariantes pour le semi-groupe engendré par l'opérateur de diffusion $L = \Delta - \nabla V \cdot \nabla$ vérifie un critère de courbure dimension $CD(\rho,\infty)$; soulignons que ce type de semi-groupe (hormis le cas gaussien) n'admettent pas de représentation explicite permettant de faire les calculs aussi simplement que pour le semi-groupe d'Ornstein-Uhlenbeck.

3. Il est possible d'introduire un paramètre de dimension $n \geq 0$ ainsi qu'un critère $CD(\rho,n)$ donné par
$$\Gamma_2(f) \geq \rho\Gamma(f) + \frac{1}{n}(Lf)^2.$$
Ce critère, plus général, permet d'atteindre des inégalités fonctionnelles plus fortes que celles de Poincaré ou de Sobolev logarithmique mais nous ne développerons pas cet aspect dans ce cours et renvoyons le lecteur vers [16] pour plus de détails.

9.5.2 Courbure-dimension et inégalité de poincaré

Dans le cas du semi-groupe d'Ornstein-Uhlenbeck et son générateur infinitésimal L, il est facile de montrer que, pour toute fonction f suffisamment régulière,

$$\Gamma_1(f) = |\nabla f|^2 \quad \Gamma_2(f) = \|\mathrm{Hess} f\|_2^2 + |\nabla f|^2.$$

Il est alors évident que $\Gamma_2 \geq \Gamma_1$. Il se trouve que ces opérateurs apparaissent naturellement (cf. [16]) dans la formule de représentation de la variance (9.4.11) que nous avons employée durant une majeure partie de notre travail :

$$\mathrm{Var}_{\gamma_n}(f) = 2\int_0^\infty \int_{\mathbb{R}^n} |\nabla P_t f|^2 d\gamma_n dt,$$

avec γ_n la mesure gaussienne standard dans $\mathbb{R}^n$. Posons, pour tout $t \geq 0$,

$$\Psi(t) = \int_{\mathbb{R}^n} |\nabla P_t f|^2 d\mu.$$

Au vu de ce qui précède, on remarque que

$$\Psi(t) = \int_{\mathbb{R}^n} \Gamma_1(P_t f) d\gamma_n$$

et

$$\Psi'(t) = -2\int_{\mathbb{R}^n} \Gamma_2(P_t f) d\gamma_n.$$

Il est alors possible d'exprimer le critère de courbure dimension $\Gamma_2 \geq \Gamma_1$, le long du semi-groupe, sous une forme intégrée, à l'aide des fonctions $t \mapsto \Psi(t)$ et $t \mapsto \Psi'(t)$. Autrement dit

$$2\Psi + \Psi' \leq 0 \Leftrightarrow \int_{\mathbb{R}^n} \Gamma_2(P_t f)\mathrm{d}\mu \geq \int_{\mathbb{R}^n} \Gamma(P_t f) d\mu$$

Remarque. Notons le fait suivant, en intégrant l'inégalité $2\Psi + \Psi' \leq 0$ entre s et t (pour $0 \leq s < t$) nous obtenons

$$\Psi(t)e^{2t} \leq \Psi(s)e^{2s}.$$

Il est alors classique de choisir $s = 0$, ceci permet d'obtenir facilement l'inégalité de Poincaré gaussienne. Observons qu'il également possible de garder s fixé et faire tendre t vers l'infini. Puisque

$$i.e. \quad \lim_{t\to+\infty} e^{2t}\Psi(t) = \left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2,$$

nous en déduisons

$$\Psi(s) \geq e^{-2s} \left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2 \quad \text{pour tout} \quad s \geq 0.$$

Ce qui permet de retrouver l'inégalité de Poincaré inverse satisfaite par la mesure gaussienne $\mathrm{Var}_{\gamma_n}(f) \geq \left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2$. D'autres résultats dans ce genre ont modestement été obtenus dans [116].

De manière générale, nous avons l'équivalence suivante :

Proposition 9.5.1. *Dans le cadre décrit précédemment, lorsque $\rho > 0$, nous avons*

$$CD(\rho,\infty) \quad \Longleftrightarrow \quad \Gamma(P_t f) \le e^{-2\rho t} P_t(\Gamma f) \quad \text{pour tout} \quad t \ge 0 \quad \text{et} \quad f \in \mathcal{D}(L). \tag{9.5.1}$$

En particulier, μ satisfait une inégalité de Poincaré de constante $C_P = \frac{1}{\rho}$.

Remarque. Ceci est à comparer avec l'inégalité (9.4.4) obtenue pour le semi-groupe d'Orstein-Uhlenbeck.

Démonstration. Soient $f \in \mathcal{D}(L)$ et $t > 0$ fixés, considérons alors $\Psi(s) = e^{-2\rho s} P_s(\Gamma(P_{t-s} f))$ pour tout $s \in [0,t]$. Par définition de Γ_2, nous avons

$$\Psi'(s) = 2e^{-2\rho s}\big[-\rho P_s\big(\Gamma(P_{t-s}f)\big) + P_s\big(\Gamma_2(P_{t-s}f)\big)\big].$$

Le critère $CD(\rho,+\infty)$ nous assure alors que $s \mapsto \Psi(s)$ est décroissante sur $[0,t]$ et le résultat s'ensuit. Pour la réciproque, observons que l'inégalité (9.5.1) devient une égalité lorsque $t = 0$. C'est pourquoi nous avons

$$0 \le \lim_{t\to 0} \frac{1}{2t}\big[e^{-2\rho t} P_t(\Gamma f) - \Gamma(P_t f)\big] = \Gamma_2(f) - \rho\Gamma(f)$$

et donc le critère $CD(\rho,+\infty)$ est satisfait. □

Remarque. Par simplicité, nous avons choisis de nous restreindre à l'étude des mesures invariantes. Cependant, il est possible d'étendre ceci en considérant des inégalités ponctuelles impliquant les mesures induites par les noyaux de transitions $(p_t)_{t\ge 0}$ afin d'obtenir des inégalités de Poincaré locale. Ceci permet notamment d'affaiblir la condition $\rho > 0$ en $\rho \in \mathbb{R}$ et celle-ci entraine l'inégalité suivante

$$P_t(f^2) - (P_t f)^2 \le 2c(t) P_t(\Gamma f) \quad \text{pour tout} \quad t \ge 0$$

avec $c(t) = \frac{1-e^{-2\rho t}}{\rho}$. Lorsque $\rho = 0$, $c(t) = 2t$. Le schéma de preuve est semblable à celui introduit précédemment.

$$\begin{aligned} P_t(f^2) - (P_t f)^2 &= \int_0^t \frac{d}{ds} P_s\big(P_{t-s}(f)\big)^2 ds \\ &= 2\int_0^t P_s\big(\Gamma(P_{t-s}f\big) ds \end{aligned}$$

Ensuite, d'après la Proposition 9.5.1, $\Gamma(P_{t-s}f) \le e^{-2\rho(t-s)} P_{t-s}(\Gamma f)$ ce qui implique

$$P_t(f^2) - (P_t f)^2 \le 2\int_0^t e^{-2\rho(t-s)} P_t(\Gamma f) ds = c(t) P_t(\Gamma f).$$

Cette dernière inégalité s'apparente bien à une inégalité de Poincaré pour la famille de mesures $(\mu_t)_{t\ge 0}$ correspondant à la loi du processus associé à la famille de semi-groupes $(P_t)_{t\ge 0}$. Lorsque $\rho > 0$, il est possible de faire tendre $t \to +\infty$ et de retrouver, par ergodicité, une inégalité de Poincaré de constante $\frac{1}{\rho}$ pour la mesure invariante μ.

9.5.3 Courbure dimension et inégalités de Sobolev logarithmiques

Nous allons voir que l'hypothèse de diffusion permet de renforcer l'inégalité (9.5.1) lorsque nous avons affaire à des diffusions. A cet effet, nous rappelons que L est un opérateur de diffusion sur $\mathbb{R}^n$ s'il s'écrit

$$L = \sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x)\frac{\partial}{\partial x_i},$$

Dans un tel cadre, Bakry et Emery ont démontré le résultat suivant.

Proposition 9.5.2. *Supposons que L soit une diffusion et que le couple (L, μ) vérifie un critère de courbure dimension $(\rho, +\infty)$ alors l'inégalité suivante est satisfaite :*

$$\sqrt{\Gamma(P_t f)} \le e^{-\rho t}\sqrt{P_t \Gamma(f)} \quad t \ge 0. \tag{9.5.2}$$

Remarque. Ceci est à comparer avec l'inégalité (9.4.4) obtenue pour le semi-groupe d'Orstein-Uhlenbeck. Dans un tel cadre, un des points cruciaux de la preuve est l'obtention de l'inégalité suivante :

$$\Gamma(\Gamma_2 - \rho\Gamma) \ge \frac{1}{4}\Gamma(\Gamma)$$

qui renforce l'inégalité $\Gamma_2 - \rho\Gamma \ge 0$ qui était employée dans la démonstration de l'inégalité de Poincaré.

En conséquence de la proposition 9.5.2 nous avons le résultat suivant.

Proposition 9.5.3. *Supposons que L soit une diffusion et que le couple (L, μ) vérifie un critère de courbure dimension $(\rho, +\infty)$ alors*

$$C(\rho, t)\frac{\Gamma(P_t f)}{P_t f} \le P_t(f \log f) - P_t f - P_t f \log P_t f \le D(\rho, t) P_t\left(\frac{\Gamma(f)}{f}\right) \quad \textit{pour tout} \quad t \ge 0.$$

avec $C(\rho, t) = \int_0^t e^{2\rho s} ds$ et $D(\rho, t) = \int_0^t e^{-2\rho s} ds$.

Démonstration. Le principe de démonstration est semblable à celui employé pour l'inégalité de Poincaré. En effet,

$$P_t(f \log f) - P_t f \log P_t f = \int_0^t \frac{d}{ds} P_s(P_{t-s} f \log P_{t-s} f) ds = \int_0^t P_s\left(\frac{\Gamma(P_{t-s} f}{P_{t-s} f}\right) ds$$

En outre, puisque $\Gamma_2 \ge \rho\Gamma$, l'application $s \mapsto e^{-2\rho s} P_s(P_{t-s} f \Gamma(\log P_{t-s} f))$ est croissante sur l'intervalle $[0, t]$, il est possible d'utiliser la propriété de commutation (9.5.2) afin d'obtenir

$$e^{2\rho(t-s)}\frac{\Gamma(P_t f)}{P_t f} = e^{-2\rho(t-s)}\frac{\Gamma(P_s(P_{t-s} f))}{P_s(P_{t-s} f)} \le P_s\left(\frac{\Gamma(P_{t-s} f)}{P_{t-s} f}\right) \le e^{-2\rho s} P_t\left(\frac{\Gamma(f)}{f}\right).$$

Le résultat s'ensuit après intégration par rapport à la variable s. A nouveau, lorsque $\rho > 0$, il est possible d'obtenir une inégalité de Sobolev logarithmique en faisant tendre $t \to +\infty$ dans l'inégalité de droite après avoir remplacer f par f^2. Pour cela, il convient d'utiliser l'ergodicité du semi-groupe et de constater que $\frac{\Gamma(f^2)}{f^2} = 4\Gamma(f)$. $\square$

9.5.4 Autres inégalités fonctionnelles

Les outils introduit précédemment peuvent également être utilisés pour obtenir des inégalités plus fortes. A titre d'exemple, si une inégalité de courbure dimension $CD(\rho, N)$ avec $\rho > 0$ et $N \in]2, +\infty[$ est vérifiée

$$\Gamma_2(f) \geq \rho\Gamma(f) + \frac{1}{N}(Lf)^2$$

alors (cf. [16]) la mesure μ sous-jacente vérifie une inégalité de Sobolev : pour f suffisament régulière, nous avons

$$\|f\|_p^2 \leq \|f\|_2^2 + \frac{4}{n(n-2)}\frac{n-1}{\rho}\mathcal{E}_\mu(f) \quad \text{avec} \quad p = \frac{2n}{n-2}.$$

Ce genre de résultat permet aussi de retrouver, dans un contexte généralisé, le Théorème de Bonnet-Myers : dans le cadre d'une variété Riemannienne M de dimension N dont la courbure de Ricci est minorée par ρ, celui-ci assure que le diamètre de M est borné par $\pi\frac{\sqrt{2pB}}{p-2}$ avec B une constante dépendant de n. Nous renvoyons le lecteur vers [76] pour plus de détails.

Il est aussi possible d'obtenir un Théorème de comparaison de profil isopérimétrique dans l'esprit du Théorème de Gromov. Pour décrire ceci, il est nécessaire d'introduire les notions que nous avons occultées dans le chapitre 6 ; aussi, dans ce qui suit (E, d, μ) désigne un espace métrique mesuré (avec μ une mesure de probabilité).

Définition 9.5.2. *Si $A \subset$, la quantité*

$$\mu^+(A) = \liminf_{r\to 0} \frac{1}{r}\big[\mu(A_r) - \mu(A)\big]$$

désigne la mesure de bord de Minkowski de l'ensemble A. Cette quantité décrit le mesure de surface de l'ensemble A.

Dans ce contexte, nous dirons que la mesure μ vérifie une inégalité de type isopérimétrique s'il existe une fonction (non triviale) $\mathcal{J}$ sur $\mathbb{R}_+$ telle que

$$\mathcal{J}\big(\mu(A)\big) \leq \mu^+(A)$$

pour tout ensemble (fermé et de mesure finie) A de E. Nous désignerons le profil isopérimétrique associé à μ par $\mathcal{I}_\mu$, comme étant la plus grande fonction $\mathcal{J}$ sur $[0, \mu(E)]$ vérifiant l'inégalité précédente.

Remarque. La définition de μ^+ explique pourquoi nous parlions de version « intégrée » lorsque nous avons introduit la notion de problème isopérimétrique dans le chapitre 6.

Exemple 9.5.1. Dans quelques exemples (cf. [28, 77]), il est possible de déterminer l'expression de $\mathcal{I}_\mu$.

1. Si $\sigma_R = \sigma_R^n$ désigne la mesure de Haar sur la sphère S_R^n de dimension n et de rayon r alors

$$\mathcal{I}_{\sigma_R} = \sigma'_R \circ \sigma_R^{-1}$$

2. Soit $n \geq 1$, alors dans le cas de la mesure gaussienne γ_n nous avons

$$\mathcal{I}_{\gamma_n} = \Phi' \circ \Phi^{-1}$$

avec $\Phi(t) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-x^2/2} dx$ et $t \in [-\infty, +\infty]$.

Nous pouvons maintenant énoncer le théorème de comparaison de Gromov.

Théorème 9.5.1 (Gromov)**.** *Soit $(\mathcal{M}, g)$ une variété riemannienne, compacte de dimension $n \geq 2$ équipée de l'élément de volume (normalisé) riemannien $d\mu$ telle que la courbure de Ricci est minorée par une constante $\rho > 0$ alors*

$$\mathcal{I}_\mu \geq \mathcal{I}_{\sigma_R}$$

avec R défini par $\frac{n-1}{R^2} = \rho$. En particulier, la mesure μ vérifie les mêmes propriétés de concentration que σ_R

$$i.e. \quad \alpha_{\mathcal{M},g,\mu}(r) \leq e^{-cr^2/2} \quad \textit{pour tout} \quad r > 0$$

Remarque. Ce genre de résultat souligne l'importance d'espaces modèles (comme celui de la sphère) pour lequel le problème isopérimétrique a été résolu et avec lesquels d'autres espaces peuvent être comparés. En revanche, le résultat précédent ne donne aucune information sur les ensembles extremaux du problèmes isopérimétrique associé à μ. Enfin, le cas d'égalité dans l'inégalité $\mathcal{I}_\mu \geq \mathcal{I}_{\sigma_R}$ n'est possible qu'à condition que $\mathcal{M}$ soit une sphère.

Dans [17], Bakry et Ledoux ont obtenu un résultat analogue lorsqu'un critère $CD(\rho, +\infty)$ (avec $\rho > 0$) est satisfait.

Théorème 9.5.2 (Bakry-Ledoux)**.** *Si (L, μ) vérifie un critère de courbure dimension $CD(\rho, \infty)$ avec $\rho > 0$ alors le profil isopérimétrique $\mathcal{I}_\mu$ est minorée par $\sqrt{\rho}\mathcal{I}_{\gamma_n}$. En particulier, la mesure μ vérifie, au moins, les mêmes propriétés de concentration que la mesure gaussienne γ_n.*

Remarque.

- Ce résultat s'applique notamment dans le cas de mesure log-concave $d\mu = e^{-V} dx$ lorsque le potentiel V est uniformément convexe, i.e. $\nabla\nabla V \geq \rho I_d$ avec $\rho > 0$.
- L'idée essentielle de la preuve consiste à reprendre la formulation fonctionnelle du problème isopérimétrique établie par Bobkov dans [24]. Dans le cas gaussien, celle-ci s'énonce comme suit : pour toutes fonctions f suffisamment régulières appartenant à l'intervalle $[0, 1]$, nous avons
$$\mathcal{I}_{\gamma_n}\left(\int_{\mathbb{R}^n} f d\gamma_n\right) \leq \int_{\mathbb{R}^n} \sqrt{\mathcal{I}^2_{\gamma_n}(f) + |\nabla f|} d\gamma_n.$$
Cette inégalité peut-être démontrée à l'aide d'arguments d'interpolations par semi-groupe. Bien que calculatoire, la preuve s'ensuit aisément en remarquant la relation cruciale
$$\mathcal{I}''_{\gamma_n} = -\frac{1}{\mathcal{I}_{\gamma_n}}.$$

9.6 Point de vue spectral des inégalités de Poincaré

Historiquement, les inégalités de Poincaré sont nées durant l'étude de problème spectraux. En fait, il se trouve qu'une inégalité de Poincaré peut se voir au travers du spectre d'un opérateur de diffusion L. En effet, elles correspondent à une formulation variationnelle de la première valeur propre non nulle d'un opérateur de diffusion L sur une variété riemannienne. Nous allons rapidement développer cet aspect dans cette section.

Sur $\mathbb{R}$, si une mesure μ admet un moment exponentiel

$$i.e. \quad \text{il existe} \quad \alpha > 0 \quad \text{tel que} \quad \int_{\mathbb{R}} e^{\alpha|x|} \mathrm{d}\mu(\mathrm{x}) < \infty,$$

il est possible de construire de manière canonique une base hilbertienne de $L^2(\mu)$. En effet, une telle base est obtenue en orthogonalisant, par rapport au produit scalaire de $L^2(\mu)$, la suite des polynômes $(x^k)_{k\geq 0}$ qui est alors un ensemble dense dans $L^2(\mu)$. A la normalisation et au signe près, une telle famille de polynômes orthogonaux est unique. Nous prendrons d'ailleurs pour convention le choix de normaliser (dans $L^2(\mu)$) ces polynômes et d'imposer au coefficient du terme dominant d'être strictement positif.

Dans un nombre assez restreint de cas, ces polynômes sont aussi les fonctions propres d'un opérateur de diffusion. En fait, en dimension un, il n'y a que les familles des polynômes d'Hermite, de Laguerre et de Jacobi qui vérifient une telle propriété. Nous nous focaliseront uniquement sur les polynômes d'Hermite dans cette section. Bien que n'étant pas exactement de même nature, le cas du cube discret $C_n = \{-1, 1\}^n$ et des polynômes de Walsh peut-être traiter de manière similaire. Il est à noter que les exemples des polynômes d'Hermite, Laguerre et Jacobi font partis des rares cas pour lesquels la description des suites de fonctions propres et de valeurs propres d'un opérateur de diffusion, dit de Sturm-Liouville, est complète (cf. [16]).

Nous présentons ci-dessous l'exemple de l'opérateur d'Ornstein-Uhlenbeck.

En dimension un, la famille des polynômes orthogonaux associée à l'opérateur de diffusion d'Ornstein-Uhlenbeck L et à la mesure gaussienne γ_1 est celle des polynômes d'Hermite que nous noterons $(H_k)_{k\geq 0}$. Ils sont obtenus, dans leur version normalisée, via la formule suivante

$$H_n(x) = \frac{(-1)^n}{\sqrt{n!}} e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2} \quad \text{avec} \quad x \in \mathbb{R}.$$

Rappelons que L agit sur un ensemble de fonctions f suffisament régulières par la formule suivante :

$$Lf = f'' - xf'$$

et que l'énergie de Dirichlet s'exprime simplement par $\mathcal{E}_{\gamma_1}(f) = \int_{\mathbb{R}} f'^2 d\gamma_1$, où $\mathcal{E}_{\gamma_1}(f)$ est un abus de notation pour désigner $\mathcal{E}_{\gamma_1}(f, f)$. Comme annoncé plus haut, les polynômes d'Hermite sont des fonctions propres de l'opérateur L, plus précisément nous avons la relation suivante :

$$-LH_k = kH_k, \quad k \in \mathbb{N},$$

Ainsi le spectre $(\lambda_k)_{k\geq 0}$ de l'opérateur L est discret et $\lambda_k = k, \quad \forall k \in \mathbb{N}$. En utilisant la décomposition d'une fonction de $L^2(\gamma_1)$ sur la base des polynômes d'Hermite nous obtienons la formule suivante :

$$f = \sum_{k=0}^{\infty} a_k H_k,$$

avec $a_k = \int_{\mathbb{R}} f H_k \mathrm{d}\gamma_1$. Ceci entraine facilement les relations suivantes :

1.
$$\|f\|_2^2 = \sum_{k\geq 0} a_k^2. \tag{9.6.1}$$

2. $P_t f = \sum_{k\geq 0} e^{-kt} a_k H_k$ avec $t \geq 0$.

3. $-Lf = \sum_{k\geq 0} k a_k H_k$.

Puisque l'énergie de Dirichlet s'exprime en fonction de L (via une intégration par partie) par $\mathcal{E}(f) = \int_{\mathbb{R}} f(-Lf)\mathrm{d}\gamma_1$, les formules de décomposition précédentes entrainent que, pour tout $f \in L^2(\gamma_1)$,

$$\mathcal{E}_{\gamma_1}(f) = \sum_{k\geq 0} k a_k^2. \tag{9.6.2}$$

Il est alors facile, à partir de (9.6.1) et (9.6.2), de démontrer de manière spectrale l'inégalité de Poincaré 9.4.2 pour une fonction $f \in L^2(\gamma_1)$ de moyenne nulle. En effet, si f est de moyenne nulle sous γ_1 nous avons

$$\mathrm{Var}_{\gamma_1}(f) = \|f\|_2^2 = \sum_{k\geq 1} a_k^2 \leq \sum_{k\geq 1} k a_k^2 = \mathcal{E}_{\gamma_1}(f)$$

L'utilisation de la terminologie de trou spectral servant, parfois, à désigner une inégalité de Poincaré devient alors évidente. En effet, comme mentionné plus tôt, la meilleure constante C_P obtenue telle que (9.1.1) soit vérifiée, correspond à la première valeur propre non nulle de l'opérateur L.

Nous pouvons étendre toutes ces définitions au cas multidimensionnel. Notons de nouveau par L l'opérateur d'Ornstein-Uhlenbeck dans $\mathbb{R}^n$,

$$\text{i.e.} \quad Lf = \Delta f - x \cdot \nabla f,$$

avec $f \in \mathcal{D}(L)$. Les valeurs propres et les fonctions propres de l'opérateur L sont maintenant indexées par $\mathbb{N}_+^n$.

Définition 9.6.1. *Pour chaque $k = (k_1, \ldots, k_n) \in \mathbb{N}_+^n$ nous définissons le k-ième polynôme d'Hermite H_k par :*

$$H_k(x) = \prod_{i=1}^{n} H_{k_i}(x_i),$$

où $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Nous retrouvons alors des propriétés analogues à la dimension 1, auxquelles nous ajoutons des formules d'intégrations utiles pour certains calculs.

Proposition 9.6.1. *1. Le k-ième polynôme d'Hermite H_k est une fonction propre de l'opérateur $-L$ correspondant à la valeur propre $k_1 + \cdots + k_n$.*

2. Puisque $L(x_i) = -x_i$ pour tout $1 \leq i \leq n$, pour n'importe quelle fonction $f \in \mathcal{D}(L)$,

$$\int_{\mathbb{R}^n} x_i f d\gamma_n = -\int_{\mathbb{R}^n} L(x_i) f d\gamma_n = \int_{\mathbb{R}^n} \partial_i f d\gamma_n,$$

ce genre de résultat s'obtient facilement en faisant directement une intégration par partie avec la densité gaussienne.

Plus généralement, pour f suffisamment régulière de sorte que la formule suivante fasse sens, on a

Proposition 9.6.2. *Pour tout $k \in \mathbb{N}^n$, nous avons la relation de récurrence suivante :*

$$\int_{\mathbb{R}^n} f H_k d\gamma_n = \frac{1}{\sqrt{k!}} \int_{\mathbb{R}^n} \frac{\partial^{k_1+\cdots+k_n} f}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}} d\gamma_n,$$

avec $k! = k_1! \dots k_n!$.

9.7 Références historiques

Les notes historiques qui vont suivre sont tirées de [33, 16, 77, **?**].

Les inégalités de Poincaré (dont la terminologie apparait pour la première fois dans un ouvrage de Courant et Hilbert) sont un sujet très vaste des mathématiques, ceci est notamment du au fait qu'elles apparaissent dans de nombreux domaines et nous nous sommes restreins qu'à certains aspects de celles-ci (concentration de la mesure, convergence vers l'équilibre,...).

Initialement, bien qu'elles ne portaient pas ce nom, elles apparaissent dans les travaux (à partir de la fin du 19ième siècle) de Neumann, Schwarz et Poincaré qui étudient certaines propriétés spectrales du laplacien dans le but de résoudre des équations aux dérivées partielles. Au niveau de la terminologie, nous souhaitons souligner que les inégalités de Poincaré que nous avons présentés portent plutôt le nom d'inégalité de Poincaré-Wirtinger (dans les inégalités de Poincaré, nous ne retranchons pas la moyenne de f à $\|f\|_2^2$) ; celles-ci ont bien été obtenues, sur des domaines convexes de $\mathbb{R}^n$, par Poincaré à l'aide d'un argument de duplication de la mesure, combiné à un changement de variable astucieux.

L'inégalité de Poincaré vérifiée par la mesure gaussienne, remonte aux moins aux débuts des années 30 dans la littérature physique via la décomposition d'une fonction de la base hilbertienne des polynômes d'Hermite. Wirtinger semble être à l'origine du même argument, cette fois-ci avec la base des polynômes trigonométriques. Plus tard, elle apparait aussi dans des travaux de Nash, Chernoff et Chen.

Les propriétés d'intégrabilité survenant grâce à une inégalité de Poincaré ou d'une inégalité de Sobolev logarithmique ont été observées dans des travaux de Aida et Stroock en 1994.

En 1983/1984, les travaux de Bakry et Emery sur les diffusions hypercontractives ont eu un rôle essentiel dans le développement de la théorie Γ_2 et ont influencés un nombre considérable de mathématiciens. Bien que plus simple, il fallu attendre 1997 pour que Bakry présente l'analogue de ces travaux avec Emery au niveau des inégalités de Poincaré locales (pour les noyaux de transitions associés à un semi-groupe). Le principe d'interpolation sur lequel repose ces deux études est déjà présent au 19ième siècle avec le principe de Duhamel. De nombreuses extensions ont depuis été obtenues, notamment pour des espaces de dimension infinie ou des variétés riemaniennes. Il est important de souligner le fait que les travaux de Bakry et Emery sont fortement inspirés (et peuvent être vu comme des extensions) de résultats en géométrie riemannienne, notamment par le prisme des travaux de Bochner et Lichnerowicz concernant l'opérateur de Laplace-Beltrami datant du milieu du 20ième siècle.

Les inégalités de Sobolev logarithmiques deviennent un objet d'étude important lorsque Gross identifie en 1975 l'équivalence entre celle-ci et l'hypercontractivité dans un contexte assez général. Le cas du semi-groupe d'Ornstein-Uhlenbeck avait été mis en avant quelques années plus tôt (en 1973) par Nelson en théorie quantique des champs. Ce genre de travaux était initialement motivé par l'étude de propriétés de mélange des processus de Markov ou des chaines de Markov, nous renvoyons le lecteur vers les ouvrages de survols écrit par Diaconis et Saloff-Coste en 1997 et 1998 afin d'en apprendre plus à ce sujet. Ce fut d'ailleurs ces deux auteurs qui déterminèrent en 1996 la meilleure constante apparaissant dans l'inégalité de Sobolev logarithmique d'une loi de Bernoulli générale ; ce résultat fut également obtenu, de manière indépendante, par Higuchi et Yoshida en 1995.

Le modèle (généralisé) linéaire gaussien comme nous avons présenté pour dire quelques mots à propos de la méthode du LASSO a été introduit par Birgé et Massart en 2001. L'estimateur LASSO est l'oeuvre de Tibshirani en 1996 et est devenu un outil incontournable pour les problèmes de régression en grande dimension. La littérature à ce sujet est très vaste et a été agrandi par de nombreux mathématiciens (Candès et Tao, Donoho, van de Geer,...) depuis.

Enfin, ajoutons quelques mots concernant l'étude de problèmes isopérimétriques. En 1997, Bobkov fut le premier a proposer une version fonctionnelle du problème isopérimétrique pour l'espace à deux points. Il en déduisit, grâce au théorème de la limite centrale, une nouvelle démonstration du problème isopérimétrique gaussien. A partir des travaux de Bobkov, Bakry et Ledoux purent proposer, via les critères de courbure dimension, une extension de ses résultats dans un contexte abstrait de triplet Markovien ; en particulier, ils proposèrent l'analogue (dans un contexte de dimension infinie) au théorème de comparaison de Gromov (celui-ci n'étant valable que pour des variétés riemanniennes de dimension finie). Depuis, en modifiant habilement la définition du critère de courbure dimension, ce genre d'étude a été prolongée aux espaces métriques mesurées par Mondino et Cavaletti en 2015.

Nous avons déjà expliqué de quelle manière des résultats isopérimétriques permettaient d'obtenir des propriétés de concentration. De ce fait, elles se trouvent au sommet de la hiérarchie des inégalités fonctionnelles. En 2010/2012, E. Milman fut le premier à observer qu'il était possible de renverser cette hiérarchie sous l'hypothèse de courbure positive, c'est-à-dire : obtenir des informations sur le profil isopérimétrique d'une mesure μ à partir des propriétés de concentrations vérifiées par cette dernière. En 2011, Ledoux présenta des démonstrations par semi-groupes des travaux d'E. Milman, simplifiant ainsi les preuves initiales.