Université Paris 6 – Pierre et Marie Curie
Master 2 « Probabilités et Modèles Aléatoires »

Mémoire encadré par Nathael Gozlan

Propriétés d'isopérimétrie et de concentration gaussiennes

Julien Bureaux

Soutenu à Marne-la-Vallée, le 23 septembre 2011

Table des matières

Remerciements

Je souhaiterais tout d'abord remercier mon directeur de mémoire Nathael Gozlan, qui recevant un e-mail sans me connaître a eu la gentillesse d'accepter de m'encadrer, de bien vouloir rechercher non pas un, mais trois sujets que je pourrais traiter, et qui a eu la patience de diriger mon travail jusqu'à ce jour. Les remarques et corrections qu'il m'a faites, les idées et références qu'il m'a indiquées m'ont toutes été d'une aide précieuse sans laquelle la rédaction de ce texte n'aurait très certainement pas été possible.

De manière plus personnelle, mes pensées vont également vers mes parents, vers ma sœur et vers mes amis qui ont su me soutenir et me supporter pendant les moments les plus difficiles de l'écriture de ce mémoire. Merci à eux tous pour leur présence au quotidien et leur bonne humeur. Je remercie plus particulièrement Jeremy qui a bien voulu relire les épreuves de ce texte. Enfin, merci à Delphine pour m'avoir relu et pour tout le reste.

Avant-propos

L'isopérimétrie et le phénomène de concentration de la mesure sont deux notions mathématiques proches quoique singulièrement distinctes. Proches parce qu'elles s'intéressent toutes les deux à un même objet, l'espace métrique mesuré, avec l'objectif commun de mieux comprendre sa géométrie (au sens large). Pour cela, elles considèrent toutes les deux la mesure des boréliens en essayant de contrôler ses accroissements. Elles diffèrent cependant dans la taille des accroissements qui les préoccupent : l'isopérimétrie s'intéresse à des accroissements infinitésimaux (ce qui correspond au périmètre) tandis que la concentration de la mesure contrôle plutôt des accroissements suffisamment gros. Le fil conducteur de ce texte sera la mise en relation de ces deux approches dans le cadre du modèle gaussien, l'un des mieux compris et des plus intéressants du point de vue des applications. On sait en effet depuis les débuts de l'étude du phénomène de la concentration de la mesure qu'une inégalité isopérimétrique de type gaussien entraîne automatiquement une propriété de concentration gaussienne pour la mesure considérée. Des travaux de recherche récents ont par ailleurs montré que la réciproque était également vraie sous une certaine hypothèse de courbure.

Dans cet objectif, ce mémoire s'ouvre sur un chapitre introductif présentant les notions d'isopérimétrie et de concentration en donnant les premières définitions. Le deuxième chapitre se concentre plus particulièrement sur l'isopérimétrie gaussienne et sera l'occasion de développer les premiers outils permettant d'établir un tel type d'isopérimétrie. Dans le troisième chapitre, on introduit pour le cas gaussien les notions de semi-groupe et de générateur infinitésimal qui ont permis d'importantes avancées dans la compréhension de l'isopérimétrie et de la concentration de la mesure. On montrera enfin dans le dernier chapitre comment ces méthodes se généralisent pour sortir du modèle purement gaussien. C'est dans ce contexte qu'on établira l'équivalence entre les propriétés d'isopérimétrie et de concentration gaussiennes évoquée plus haut.

Notations et conventions

Nous regroupons ici pour plus de commodité les principales notations introduites par ailleurs dans le corps du texte.

Sauf précision, nous nous placerons toujours dans un espace euclidien $\mathbb{R}^d$ de dimension $d \geqslant 1$ muni de son produit scalaire canonique noté $\cdot$ et de la norme euclidienne associée : si $(x_1, \ldots, x_d)$ sont les coordonnées d'un vecteur x dans la base canonique,

$$|x| = \sqrt{|x_1|^2 + \cdots + |x_d|^2}.$$

Pour toute matrice carré $M = (m_{i,j})$ de taille d à coefficients réels, sa norme de Hilbert-Schmidt est définie par

$$\|M\|_2 = \sqrt{\sum_{i=1}^{d} \sum_{j=1}^{d} (m_{i,j})^2}.$$

On assimilera toujours une matrice S symétrique réelle de taille d avec l'application bilinéaire symétrique définie pour $x, y \in \mathbb{R}^d$ par

$$S(x, y) = {}^t x S y.$$

Pour toute partie A de $\mathbb{R}^d$, la distance à A est l'application 1-lipschitzienne définie pour tout $x \in \mathbb{R}^d$ par

$$d(x, A) = \inf_{a \in A} |x - a|.$$

Le voisinage d'ordre $r > 0$ de A est alors l'ouvert $A_r = \{x \in \mathbb{R}^d : d(x, A) < r\}$.

Si μ est une mesure (positive) borélienne sur $\mathbb{R}^d$, on associe à tout borélien A une « mesure de bord » $\mu^+(A)$ définie par la formule du contenu (inférieur) de Minkowski :

$$\mu^+(A) = \liminf_{r \to 0} \frac{\mu(A_r) - \mu(A)}{r}.$$

Pour toute fonction localement lipschitzienne $f \colon \mathbb{R}^d \to \mathbb{R}$, on définit en tout point $x \in \mathbb{R}^d$ sa « norme de gradient » comme la constante de Lipschitz locale optimale

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

Chapitre 1

Isopérimétrie et concentration

Après une brève présentation historique de la notion d'isopérimétrie, nous en introduisons une généralisation de nature probabiliste et plus spécifiquement sa version gaussienne. Cette dernière fournit un modèle particulièrement intéressant d'isopérimétrie, tant du point de vue théorique que pour les applications. Nous aborderons dans cet esprit la question du phénomène de concentration de la mesure et de la concentration gaussienne.

1.1 Le problème isopérimétrique classique

On désigne sous le terme d'isopérimétrie une grande famille de problèmes partageant la même nature géométrique, dont les plus anciens et les plus célèbres intéressaient déjà les savants grecs.

Le problème « classique » consiste à déterminer, parmi toutes les formes du plan de même périmètre, celles qui possèdent la plus grande surface [*]. Si des traces de ce problème et de sa solution apparaissent très tôt (on peut par exemple citer le mythe de la fondation de Carthage par la reine Didon), les premières démonstrations sont attribuées à Zénodore au IIe siècle av. J.-C. qui montre que, parmi les polygones à n côtés de périmètre fixé, ce sont les polygones réguliers qui réalisent une surface maximale. Le problème plus général qui consiste à ne plus se restreindre aux polygones mais à considérer toutes les courbes fermées régulières, et dont les solutions sont données par les cercles, est resté longtemps hors de portée des mathématiciens, bien qu'intuitivement compris. Il faut attendre le XIXe siècle pour que les méthodes de symétrisation de Steiner d'une part, puis les travaux de Weierstrass sur le calcul variationnel et ceux de Minkowski sur la géométrie des convexes d'autre part, permettent enfin d'établir rigoureusement l'unicité

$*$. De façon duale, il est bien sûr équivalent de chercher à minimiser le périmètre à surface constante. C'est d'ailleurs plutôt dans ce sens-ci qu'on travaillera ici.

et l'existence de la solution.

D'autres démonstrations sont apparues par la suite (notamment celle de Hurwitz utilisant les séries de Fourier), mais c'est une généralisation de l'approche de Minkowski qui permet en 1937 à Aleksandrov et Fenchel de prouver le théorème isopérimétrique général dans les espaces euclidiens de dimension arbitrairement grande munis de la mesure de Lebesgue :

> Soit A un borélien de $\mathbb{R}^d$ de bord ∂A suffisamment régulier, et soit B une boule de volume $\mathrm{Vol}(B) = \mathrm{Vol}(A)$. Alors les mesures de surfaces respectives vérifient l'inégalité $\sigma(\partial A) \geqslant \sigma(\partial B)$.

Il est remarquable que ce théorème reste valable pour certaines géométries non euclidiennes. Le premier résultat de ce type est celui de Lévy sur la sphère $\mathbb{S}^d$, qu'il a ensuite généralisé aux surfaces à courbure positive. Gromov l'a finalement étendu au cadre de la géométrie riemannienne.

Si la notion de périmètre est assez intuitive dans le plan, elle devient plus délicate en dimension supérieure. Une démarche initiée par les travaux de Minkowski consiste à interpréter la surface d'un borélien $A \subset \mathbb{R}^d$ comme la zone par laquelle il grandit. Plus précisément, si on l'élargit de manière uniforme en l'ouvert $A_r = \{x \in \mathbb{R}^d : d(x, A) < r\}$, il est naturel de voir le comportement des accroissements de $\mathrm{Vol}(A_r)$ comme une certaine « mesure » de la taille du bord de A. Ceci permet de reformuler le théorème isopérimétrique en s'affranchissant des questions de régularité et de construction de la mesure de surface.

> Soit A un borélien de $\mathbb{R}^d$, et soit B une boule de volume $\mathrm{Vol}(B) = \mathrm{Vol}(A)$. Alors, pour tout $r > 0$ on a l'inégalité $\mathrm{Vol}(A_r) \geqslant \mathrm{Vol}(B_r)$.

Dans cet esprit, les accroissement infinitésimaux permettent de définir une notion faible de périmètre, appelée contenu de Minkowski [*].

$$\mathrm{Vol}^+(A) = \liminf_{r \to 0} \frac{\mathrm{Vol}(A_r) - \mathrm{Vol}(A)}{r}.$$

Le fait de pouvoir ainsi définir une notion de périmètre de manière intrinsèque à partir de la seule conception de volume nous permettra de généraliser l'isopérimétrie à d'autres types de mesure. Notons $V_d(r)$ (respectivement $S_d(r)$) le volume (respectivement la surface) d'une boule de rayon $r > 0$ dans $\mathbb{R}^d$. Le théorème peut alors se réécrire comme une inégalité entre le volume et la surface :

> Pour tout A borélien de $\mathbb{R}^d$, on a $\mathrm{Vol}^+(A) \geqslant (S_d \circ V_d^{-1})(\mathrm{Vol}(A))$.

$*$. Pour être tout à fait précis il faudrait parler de contenu *inférieur*, mais nous n'utiliserons jamais la version avec la limite supérieure dans ce mémoire.

1.2 Un autre point de vue sur l'isopérimétrie ?

Alors que les généralisations de Lévy et Gromov se concentrent sur la géométrie de l'espace, on peut aussi être tenté de remplacer les considérations de volume et de surface opérées par la mesure de Lebesgue par d'autres idées de « masse ».

Considérons par exemple une variable aléatoire X définie sur un espace abstrait de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans un espace métrique (S, d) muni de sa tribu borélienne. La loi μ de la variable aléatoire X fournit une mesure de probabilité sur l'espace S dont l'étude permet de mieux comprendre le phénomène aléatoire décrit par X. Nous adopterons ce point de vue en nous concentrant sur l'espace mesuré (S, μ), sans généralement faire référence à l'espace de probabilité initial.

1.2.1 Profil isopérimétrique

Soit (S, d) un espace métrique muni de sa tribu borélienne et μ une mesure de probabilité sur S. On définit la mesure de bord $\mu^+(A)$ de tout borélien $A \subset S$ à l'aide de la formule du contenu de Minkowski

$$\mu^+(A) = \liminf_{r \to 0} \frac{\mu(A_r) - \mu(A)}{r},$$

où $A_r = \{x \in S : d(x, A) < r\}$ est le r-voisinage ouvert de A dans S.

Remarque. Si $\mu(\bar{A}) > \mu(A)$, alors $\mu^+(A) = +\infty$. En effet, puisque $\bar{A} = \cap_{r>0} A_r$ on a

$$\frac{\mu(A_r) - \mu(A)}{r} \geqslant \frac{\mu(\bar{A}) - \mu(A)}{r} \xrightarrow[r \to 0]{} +\infty.$$

Définition 1. Soit μ une mesure de probabilité borélienne sur S. Le *profil isopérimétrique* de la mesure μ est la fonction $I_\mu \colon [0\,;1] \to [0\,;\infty]$ définie * par

$$I_\mu(v) = \inf\{\mu^+(A) : A \subset S,\ \mu(A) = v\}.$$

Remarque. On peut se contenter de prendre l'infimum seulement sur les parties fermées de S. En effet, si A est un borélien, on a toujours $\mu(A) \leqslant \mu(\bar{A})$, mais $\mu(\bar{A}_r) = \mu(A_r)$ car $\bar{A}_r = A_r$ pour tout $r > 0$. En passant à la limite, on obtient $\mu^+(A) \geqslant \mu^+(\bar{A})$. Par ailleurs, on peut supposer que $\mu(\bar{A}) = \mu(A)$ car sinon $\mu^+(A) = +\infty$ en vertu de la remarque précédente. Cette remarque est également valable quand il s'agit de vérifier une inégalité isopérimétrique de la forme $\mu^+(A) \geqslant I(\mu(A))$ sur tous les boréliens.

Le profil I_μ fournit automatiquement une inégalité isopérimétrique sous la mesure μ : pour tout borélien $A \subset S$, on a

$$\mu^+(A) \geqslant I_\mu(\mu(A)).$$

$*$. Avec la convention $\inf \emptyset = +\infty$.

On n'essayera pas (dans ce mémoire du moins) de déterminer précisément le profil isopérimétrique I_μ mais plutôt de le comparer à la situation de référence que nous présentons maintenant.

1.2.2 Isopérimétrie gaussienne

Le théorème isopérimétrique gaussien traite le cas fondamental des mesures de probabilité gaussiennes. Il a été démontré indépendamment par C. Borell ([Bor75]) d'une part, V. Sudakov et B. Tsirel'son ([ST78]) d'autre part. De manière peut-être surprenante, leur preuves reposent en fait sur une utilisation astucieuse de l'inégalité isopérimétrique classique de Lévy sur la sphère, à laquelle ils associent un argument limite de Poincaré bien connu qui consiste à voir la mesure gaussienne sur $\mathbb{R}^d$ comme une limite de projections orthogonales de mesures uniformes sur des sphères.

Soit γ_d la mesure gaussienne standard sur $\mathbb{R}^d$,

$$\gamma_d(dx) = \left(\frac{1}{2\pi}\right)^{d/2} e^{-\frac{1}{2}|x|^2}\, dx.$$

et γ_d^+ la mesure de bord associée. On note Φ la fonction de répartition de la mesure gaussienne γ_1 sur $\mathbb{R}$

$$\Phi(u) = \gamma_1(]-\infty\,;u]) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-x^2/2}dx, \qquad u \in \mathbb{R}.$$

Comme annoncé, la mesure gaussienne vérifie une inégalité isopérimétrique sur $\mathbb{R}^d$ que l'on ne prouvera pas avant le prochain chapitre.

Théorème 1. *Pour tout borélien $A \subset \mathbb{R}^d$ on a $\gamma_d^+(A) \geqslant (\Phi' \circ \Phi^{-1})(\gamma_d(A))$.*

Par la suite, on notera $\mathcal{I} = \Phi' \circ \Phi^{-1}$ cette fonction. On peut tout de suite vérifier que $\mathcal{I}$ coïncide avec le profil isopérimétrique I_{γ_d} en conséquence de la proposition suivante qui exprime que les demi-espaces sont *extrémaux* pour la mesure gaussienne.

Proposition 1. *Soit $H = \{x \in \mathbb{R}^d : x \cdot u \leqslant h\}$ un demi-espace de $\mathbb{R}^d$ avec $u \in \mathbb{R}^d$ unitaire et $h \in \mathbb{R}$. On a l'égalité $\gamma_d^+(H) = \mathcal{I}(\gamma_d(H))$.*

Remarque. La réciproque est vraie mais requiert des considérations de symétrie assez délicates [Ehr83].

Démonstration. Comme la mesure gaussienne est laissée invariante par les isométries vectorielles, on peut supposer que $u = e_1$ est le premier vecteur de la base canonique de $\mathbb{R}^d$. Il vient alors

$$\gamma_d(H) = \int_{x\cdot u \leqslant h} \gamma_d(dx) = \int_{x_1 \leqslant h} \gamma_d(dx) = \int_{-\infty}^{h} e^{-x_1^2/2}\frac{dx_1}{\sqrt{2\pi}} = \Phi(h).$$

La distance à un demi-espace étant atteinte pour le projeté orthogonal, on a que $H_r = \{x \in \mathbb{R}^d : \langle x, u\rangle < h + r\}$ pour tout $r > 0$, de sorte que

$$\gamma_d^+(H) = \liminf_{r\to 0} \frac{\Phi(h+r) - \Phi(h)}{r} = \Phi'(h).$$

□

Définition 2. Soit μ une mesure borélienne de probabilité sur $\mathbb{R}^d$. On dit que μ vérifie une *inégalité isopérimétrique gaussienne* de constante $c > 0$ si pour tout borélien $A \subset \mathbb{R}^d$, on a l'inégalité

$$\mu^+(A) \geqslant c\mathcal{I}(\mu(A)).$$

Remarque. Ceci revient à dire que le profil isopérimétrique I_μ est minoré par $c\mathcal{I}$.

Remarque. Ce n'est pas parce qu'on ne considère dans ce mémoire que l'isopérimétrie gaussienne qu'il faut croire que ce soit la seule digne d'intérêt. On peut par exemple noter que la loi exponentielle bilatère sur $\mathbb{R}$ de densité $d\mu/dx = \frac{1}{2}e^{-|x|}$ vérifie une inégalité isopérimétrique de profil $I_\mu(v) = \min\{v, 1-v\}$.

1.3 Concentration gaussienne

On introduit quelques définitions et propriétés générales sur le phénomène de concentration de la mesure avant d'aborder la question de la concentration gaussienne et de ses relations avec l'isopérimétrie.

1.3.1 Généralités sur la concentration de la mesure

Soit μ une mesure de probabilité borélienne sur un espace métrique (S, d). Alors pour tout partie $A \subset S$ non vide, on a

$$S = \bigcup_{r>0} A_r = \bigcup_{n\geqslant 1} A_n$$

avec $A_n \overset{\text{déf.}}{=} \{x \in S : d(x, A) < n\} \subset A_{n+1}$, et donc *a fortiori*,

$$\lim_{r\to+\infty} \mu(A_r) = \lim_{n\to+\infty} \mu(A_n) = \mu(S) = 1.$$

Par ailleurs, pour certaines mesures, on observe une vitesse de convergence uniformément rapide : pour tout A borélien de masse $\mu(A) \geqslant 1/2$ * les voisinages A_r remplissent rapidement tout l'espace au sens de la mesure, ce qu'on nomme communément *phénomène de concentration de la mesure*. La définition suivante permet de contrôler la vitesse de convergence uniformément.

*. On pourrait aussi bien prendre $\mu(A) \geqslant \varepsilon$ pour un $\varepsilon > 0$ fixé.

Définition 3. Soit μ une mesure de probabilité borélienne sur un espace métrique (S, d). La *fonction de concentration* de μ, notée α_μ est définie pour tout $r > 0$ par

$$1 - \alpha_\mu(r) = \inf\{\mu(A_r) : \mu(A) \geqslant 1/2\}.$$

Remarque. De cette manière, la fonction de concentration fournit une *inégalité de concentration* : pour tout borélien $A \subset S$ tel que $\mu(A) \geqslant 1/2$, on a

$$\mu(A_r) \geqslant 1 - \alpha_\mu(r).$$

Remarque. Bien entendu, on a toujours $\alpha_\mu(r) \leqslant 1/2$ et α_μ est décroissante.

Il est intéressant de constater que cette fonction de concentration donne un contrôle uniforme sur tous les boréliens de mesure supérieure à $1/2$ et tend *toujours* vers 0.

Lemme 1. *La fonction $\alpha_\mu(r)$ tend vers 0 lorsque r tend vers $+\infty$.*

Démonstration. Soit $\varepsilon \in \,]0\,; 1/2[$ et $x \in S$ un point fixé. Comme on l'a déjà vu, $\mu(\{x\}_r) = \mu(B(x, r))$ tend vers 1 quand $r \to +\infty$, et donc, il existe une boule $B = B(x, r_0)$ (centrée en x) telle que $\mu(B) \geqslant 1 - \varepsilon$. Ainsi, pour tout borélien A tel que $\mu(A) \geqslant 1/2$, a-t-on $\mu(A \cap B) \geqslant 1/2 - \varepsilon > 0$, et en particulier $A \cap B \neq \emptyset$. Pour tout $r \geqslant 2r_0$ on a alors automatiquement $B \subset A_r$, d'où $\mu(A_r) \geqslant 1 - \varepsilon$. □

Notons que la fonction de concentration donne aussi un contrôle sur les boréliens dont la masse est petite au moyen du lemme suivant.

Lemme 2. *Soit r_μ la fonction définie pour $\varepsilon > 0$ par $r_\mu(\varepsilon) = \inf\{r > 0 : \alpha_\mu(r) < \varepsilon\}$. Pour tout borélien $A \subset S$ avec $\mu(A) \geqslant \varepsilon > 0$, pour tout $r_0 > r_\mu(\varepsilon)$ et pour tout $r > 0$, on a la minoration*

$$\mu(A_{r_0+r}) \geqslant 1 - \alpha_\mu(r).$$

Démonstration. Il suffit de voir que $(A_{r_0})_r \subset A_{r_0+r}$ (par inégalité triangulaire) et que $\mu(A_{r_0}) \geqslant 1/2$. Soit B le complémentaire de A_{r_0}. On a $A \subset S \setminus B_{r_0}$ et si on avait $\mu(B) \geqslant 1/2$, cela entraînerait

$$\mu(A) \leqslant 1 - \mu(B_{r_0}) \leqslant \alpha_\mu(r_0) < \varepsilon.$$

□

On peut changer légèrement notre façon de voir la concentration. Plutôt que de considérer directement la mesure μ, on observe les mesures images $\mu \circ f^{-1}$ où $f : S \to \mathbb{R}$ est une fonction lipschitzienne*. Ce point de vue est particulièrement pertinent en vue d'applications en théorie des probabilités : f est alors perçue comme une variable

*. On verra dès le prochain chapitre que cette approche est également très féconde dans le cadre de l'isopérimétrie.

aléatoire à valeurs réelles de loi $\mu \circ f^{-1}$. Les inégalités de concentration se traduisent alors automatiquement en inégalités de déviation pour f, et ces inégalités caractérisent la fonction de concentration.

Proposition 2. *La fonction de concentration α_μ est la plus petite fonction positive α définie sur $]0\,;+\infty[$ vérifiant : pour toute $f\colon S \to \mathbb{R}$ fonction 1-lipschitzienne, et pour tout m tel que $\mu\{f > m\} \leqslant 1/2$, on a*

$$\forall r > 0, \quad \mu\{f \geqslant m + r\} \leqslant \alpha(r).$$

Démonstration. Commençons par montrer que α_μ vérifie la propriété annoncée. Pour cela, fixons $f\colon S \to \mathbb{R}$ fonction 1-lipschitzienne, et soit m tel que $\mu\{f > m\} \leqslant 1/2$. Considérons alors le fermé $A = \{f \leqslant m\}$ de masse $\mu(A) \geqslant 1/2$. Puisque f est 1-lipschitzienne, on a $A_r \subset \{f < m + r\}$ pour tout $r > 0$, et donc par définition de α_μ,

$$\mu\{f \geqslant m + r\} = 1 - \mu\{f < m + t\} \leqslant 1 - \mu(A_r) \leqslant \alpha_\mu(r).$$

Réciproquement, prenons une fonction α qui vérifie cette propriété. Soit $A \subset S$ un borélien de masse $\mu(A) \geqslant 1/2$ et soit $f = d(\cdot, A)$. On a

$$\mu\{f > 0\} = 1 - \mu(\bar{A}) \leqslant 1 - \mu(A) \leqslant 1/2$$

et la propriété s'applique donc pour $m = 0$:

$$\forall r > 0, \quad \mu(A_r) = \mu\{f < r\} = 1 - \mu\{f \leqslant r\} \geqslant 1 - \alpha(r).$$

On obtient alors l'inégalité $\alpha_\mu \leqslant \alpha$ par définition de α_μ. □

Remarque. Si f est une fonction K-lipschitzienne, elle vérifie également une inégalité de déviation (quitte à renormaliser f par K) :

$$\mu\{f \geqslant m + r\} \leqslant \alpha_\mu(r/K).$$

Remarque. Soit $f\colon S \to \mathbb{R}$ une fonction 1-lipschitzienne et m un réel tel que $\mu\{f > m\} \leqslant 1/2$ et $\mu\{f < m\} \leqslant 1/2$ (on dit que m est *une* médiane de f sous μ). L'inégalité de déviation pour f s'écrit

$$\forall r > 0, \quad \mu\{f \geqslant m + r\} \leqslant \alpha_\mu(r).$$

On a aussi une inégalité de déviation pour $-f$ car $\mu\{-f > -m\} \leqslant 1/2$,

$$\forall r > 0, \quad \mu\{f \leqslant m - r\} \leqslant \alpha_\mu(r).$$

Puisque $\mu\{|f - m| \geqslant r\} \leqslant \mu\{f \geqslant m + r\} + \mu\{f \leqslant m - r\}$, on en déduit une inégalité de concentration pour la loi de f sous μ de la forme

$$\forall r > 0, \quad \mu\{|f - m| \geqslant r\} \leqslant 2\alpha_\mu(r).$$

Cette inégalité s'étend bien sûr à f seulement k-lipschitzienne en renormalisant.

1.3.2 Liens entre isopérimétrie gaussienne et concentration

De nombreuses méthodes ont été développées afin d'établir des propriétés de concentration : on peut notamment citer les travaux de M. Talagrand dans les espaces produits, l'utilisation d'inégalités de Poincaré, d'inégalités de Sobolev logarithmiques, d'inégalités de transport, etc (voir par exemple [Led01]). L'approche que nous suivons ici et qui se base sur l'isopérimétrie gaussienne fournit un type de concentration particulièrement robuste et des constantes optimales.

On se place dorénavant dans l'espace euclidien $S = \mathbb{R}^d$.

Proposition 3. *Soit μ une mesure de probabilité borélienne sur $\mathbb{R}^d$ qui vérifie une inégalité isopérimétrique de constante $c > 0$. Alors*

$$\alpha_\mu(r) \leqslant 1 - \Phi(cr) \leqslant \frac{1}{r\sqrt{2\pi}} \exp\left(-\frac{r^2}{2c^2}\right), \qquad \forall r > 0.$$

Démonstration. On verra au chapitre 2 (théorème 2) qu'une inégalité isopérimétrique gaussienne peut se mettre sous la forme intégrée suivante : pour tout borélien $A \subset \mathbb{R}^d$ et pour tout $r > 0$, on a

$$\mu(A_r) \geqslant \Phi(a + cr)$$

où a est pris tel que $\Phi(a) = \mu(A)$. En particulier, si $\mu(A) \geqslant 1/2$, on a $a = \Phi^{-1}(\mu(A)) \geqslant 0$, et donc par monotonie

$$1 - \mu(A_r) \leqslant 1 - \Phi(cr).$$

On en déduit que $\alpha_\mu(r) \leqslant 1 - \Phi(cr)$ par définition. La dernière inégalité provient de l'encadrement classique de la queue gaussienne

$$\frac{e^{-r^2/2}}{r\sqrt{2\pi}}\left(1 - \frac{1}{r^2}\right) \leqslant 1 - \Phi(r) \leqslant \frac{e^{-r^2/2}}{r\sqrt{2\pi}}.$$

□

Remarque. Pour la mesure gaussienne γ_d elle-même, on est capable de déterminer précisément la fonction de concentration. En effet, quitte à admettre le théorème 1, on a l'inégalité $\alpha_{\gamma_d} \leqslant 1 - \Phi$ d'après ce qui précède (avec $c = 1$). De plus, on sait que les demi-espaces vérifient le cas d'égalité dans l'inégalité isopérimétrique pour γ_d. Si $H = \{x \in \mathbb{R}^d : x \cdot e_1 \leqslant 0\}$, on a déjà vu que $H_r = \{x \in \mathbb{R}^d : x \cdot e_1 < r\}$ pour tout $r > 0$ et que $\gamma_d(H_r) = \Phi(r)$. On en déduit alors *l'égalité*

$$\alpha_{\gamma_d}(r) = 1 - \Phi(r), \qquad \forall r > 0.$$

Remarque. Il est intéressant de remarquer que l'isopérimétrie gaussienne conduit à une situation dans laquelle la fonction de concentration est indépendante de la dimension dimension d de l'espace sur lequel μ est définie. Cette propriété s'avère très puissante en pratique, comme le montre l'exemple que nous présentons maintenant.

Une application standard

Soient $X_1, \dots, X_d$ des variables aléatoires indépendantes de loi $\mathcal{N}(m, \sigma^2)$ et soit $f : \mathbb{R}^d \to \mathbb{R}$ une fonction K-lipschitzienne telle que $Z = f(X_1, \dots, X_d)$ vérifie $\mathbb{E}[|Z|] < \infty$. Alors le vecteur

$$\left(\frac{X_1 - m}{\sigma}, \frac{X_2 - m}{\sigma}, \dots, \frac{X_d - m}{\sigma}\right)$$

est un vecteur gaussien de loi γ_d. On obtient donc une inégalité de concentration pour la loi de Z (on admet ici qu'on peut remplacer la médiane par la moyenne [Led01]) :

$$\mathbb{P}(|Z - \mathbb{E}[Z]| \geqslant r) \leqslant 2 \exp\left(-\frac{r^2}{2K^2\sigma^2}\right), \qquad r > 0$$

ce qu'on peut traduire en intervalle de confiance pour Z :

$$\mathbb{P}\left(|Z - \mathbb{E}[Z]| < \sigma K \sqrt{2 \log \frac{2}{\alpha}}\right) \geqslant 1 - \alpha, \qquad \alpha \in \left]0\,;1\right[.$$

En particulier pour la moyenne empirique $Y = (X_1 + \cdots + X_d)/d$,

$$\mathbb{P}(|Y - m| \geqslant r) \leqslant 2 \exp\left(-\frac{r^2 d}{2\sigma^2}\right), \qquad r > 0$$

$$\mathbb{P}\left(|Y - m| < \sigma \sqrt{\frac{2}{d} \log \frac{2}{a}}\right) \geqslant 1 - \alpha, \qquad \alpha \in \left]0\,;1\right[.$$

De telles inégalités sont à mettre en parallèle avec les résultats asymptotiques fournis par la loi des grands nombres ou le théorème central limite. À leur différence, on a ici un contrôle précis, et *à temps fini*. Notons qu'on aurait pu obtenir l'inégalité pour la moyenne empirique de manière purement élémentaire avec la méthode de Chernoff.

Comme pour l'isopérimétrie, le cas gaussien fournit une situation de référence par rapport à laquelle on repère les autres.

Définition 4. Une mesure de probabilité borélienne μ sur un métrique (S, d) vérifie une *propriété de concentration gaussienne* s'il existe une constante $C > 0$ telle que pour tout $r > 0$,

$$\alpha_\mu(r) \leqslant C e^{-r^2/C}.$$

Comme on l'a vu plus haut, une inégalité isopérimétrique gaussienne donne de manière automatique une propriété de concentration gaussienne. Nous montrerons plus tard dans le mémoire (au chapitre 4) que sous une certaine hypothèse portant sur la mesure μ (de convexité), une propriété de concentration gaussienne pour μ entraîne à son tour une propriété d'isopérimétrie gaussienne.

Chapitre 2

Formulation fonctionnelle de l'isopérimétrie gaussienne

Nous développons dans ce chapitre une approche fonctionnelle de l'isopérimétrie gaussienne initiée par Bobkov [Bob96], et présentons notamment une inégalité fonctionnelle qui possède d'intéressantes propriétés de tensorisation et qui apparaîtra dès lors comme un outil de première importance. En guise d'exemple, nous utiliserons ces outils pour donner une première preuve élémentaire de l'inégalité isopérimétrique pour la mesure de Gauss.

2.1 Forme intégrée de l'isopérimétrie gaussienne

Rappelons que $\mathcal{I} = \Phi' \circ \Phi^{-1}$ désigne la fonction intervenant dans l'inégalité isopérimétrie pour la mesure gaussienne, appelée *profil isopérimétrique gaussien.*

Définition 5. Un mesure borélienne de probabilité μ sur $\mathbb{R}^d$ vérifie une *inégalité isopérimétrique gaussienne* s'il existe une constante $c > 0$ telle que $\mu^+(A) \geqslant c\mathcal{I}(\mu(A))$ pour tout borélien $A \subset \mathbb{R}^d$.

Il est naturel (notamment en vue d'applications à la concentration de la mesure) de se demander si, comme dans le cas du problème isopérimétrique classique, on peut intégrer l'inégalité isopérimétrique gaussienne en une inégalité de la forme $\mu(A_h) \geqslant R_h(\mu(A))$. La réponse est affirmative et on a même une équivalence.

Théorème 2. *Une mesure borélienne de probabilité μ sur $\mathbb{R}^d$ vérifie une inégalité isopérimétrique gaussienne de constante $c > 0$ si et seulement si, pour tout borélien $A \subset \mathbb{R}^d$ et tout $h > 0$, on a $\mu(A_h) \geqslant \Phi(a + ch)$ avec $a \in \mathbb{R}$ tel que $\Phi(a) = \mu(A)$.*

Remarque. Par monotonie de Φ (et Φ^{-1}), il suffit de vérifier l'inégalité intégrée sur les fermés. En effet, supposons que c'est le cas, et prenons A un borélien quelconque de $\mathbb{R}^d$.

Alors puisque pour tout $h > 0$ on a $A_h = \bar{A}_h$, on obtient

$$\mu(A_h) = \mu(\bar{A}_h) \geqslant \Phi(\Phi^{-1}(\mu(\bar{A})) + ch) \geqslant \Phi(\Phi^{-1}(\mu(A)) + ch).$$

Preuve qui échoue. Le raisonnement suivant n'est pas valide mais permet de mieux comprendre le résultat et la démonstration qui va suivre. Soit A un fermé de $\mathbb{R}^d$. Supposons que la fonction $h \mapsto \mu(A_h)$ soit dérivable sur $]0\,;\infty[$ de dérivée $\mu^+(A_h)$ (ce serait le cas si la limite inférieure était une vraie limite). Comme A est fermé, $\lim_{h\to 0}\mu(A_h) = \mu(A)$, et elle est donc de plus continue en 0. Il en est de même de la composée $u\colon h \mapsto \Phi^{-1}(\mu(A_h))$ de dérivée

$$u'(h) = \frac{\mu^+(A_h)}{\mathcal{I}(\mu(A_h))} \geqslant c.$$

Pour tout $r > 0$, le théorème des accroissements finis montre alors que $u(r) - u(0) \geqslant cr$, c'est à dire $\mu(A_r) \geqslant \Phi(a + cr)$. □

Nous donnons maintenant une preuve plus technique, mais rigoureuse.

Démonstration. Nous commençons par montrer que l'inégalité intégrée, entraîne l'isopérimétrie gaussienne. Fixons un borélien $A \subset \mathbb{R}^d$ et $a \in \mathbb{R}$ tel que $\Phi(a) = \mu(A)$. Alors, pour tout $r > 0$ on a

$$\frac{\mu(A_r) - \mu(A)}{r} \geqslant \frac{\Phi(a + cr) - \Phi(a)}{r}.$$

En faisant tendre r vers 0, il vient que $\mu^+(A) \geqslant c\,\Phi'(a) = c\,\mathcal{I}(\mu(A))$.

Passons à la réciproque. Afin d'alléger la preuve on suppose que $c = 1$, et on introduit une famille de fonctions R_h définies par $R_h(p) = \Phi(\Phi^{-1}(p) + h)$ pour $h \geqslant 0$, $p \in]0\,;1[$ et $R_h(1) = 1 - R_h(0) = 1$. Les fonctions R_h sont strictement croissantes, continues et vérifient la propriété de semi-groupe $R_{h+h'} = R_h \circ R_{h'}$. On veut montrer que pour tout borélien $A \subset \mathbb{R}^d$, on a

$$\mu(A_h) \geqslant R_h(\mu(A)).$$

On commence par fixer un borélien $A \subset \mathbb{R}^d$ de mesure $\mu(A) \in]0\,;1[$ (si $\mu(A) \in \{0,1\}$, l'inégalité est triviale au vu la définition de R_h). De plus, plutôt que de la prouver directement, on l'affaiblit légèrement sous la forme

$$\mu(A_h) \geqslant R^\sigma_h(\mu(A)), \quad \text{où } \sigma > 1, \text{ et } R^\sigma_h \overset{\text{déf.}}{=} R_{h/\sigma}.$$

Introduisons l'ensemble Δ_A des $h > 0$ tels que cette inégalité est vérifiée pour tout $h' \in]0\,;h]$, et soit $h^* \overset{\text{déf.}}{=} \sup \Delta_A$ (avec la convention $\inf \emptyset = 0$). Par définition, Δ_A est nécessairement de la forme $]0\,;h^*]$ ou $]0\,;h^*[$. On procède alors en plusieurs étapes :

1. On montre que $h^* > 0$.
2. On montre que $h^* \notin \Delta$.
3. On montre que $h^* = +\infty$.

Ceci prouvera que $\Delta_A =]0\,;+\infty[$.

1. Par définition de μ^+, on a pour h au voisinage de 0^+,

$$\mu(A_h) \geqslant \mu(A) + h\mu^+(A) + o(h).$$

Par ailleurs, un développement de Taylor-Young de $h \mapsto R_h^\sigma(\mu(A))$ donne

$$\begin{aligned} R_h^\sigma(\mu(A)) &= \mu(A) + \frac{h}{\sigma}\Phi'(\Phi^{-1}(\mu(A))) + o(h) \\ &= \mu(A) + \frac{h}{\sigma}\mathcal{I}(\mu(A)) + o(h). \end{aligned}$$

Comme par hypothèse $\mu^+(A) \geqslant \mathcal{I}(\mu(A))$, ceci montre que $\mu(A_h) \geqslant R_h^\sigma(\mu(A))$ pour tout h assez petit (on a $\mu^+(A) > 0$ car $0 < \mu(A) < 1$), et donc $h^* > 0$ car Δ_A n'est pas vide.

2. Supposons que $h^* \in \Delta_A$ et posons $B = A_{h^*}$, qui vérifie donc $\mu(B) \geqslant R_{h^*}^\sigma(A)$. Ici encore on peut supposer que $0 < \mu(B) < 1$. Par définition de μ^+ et par un développement de Taylor-Young, pour h au voisinage de 0^+, on a

$$\begin{aligned} \mu(B_h) &\geqslant \mu(B) + h\mu^+(B) + o(h) \\ R_h^\sigma(\mu(B)) &= \mu(B) + \frac{h}{\sigma}\mathcal{I}(\mu(B)) + o(h). \end{aligned}$$

En utilisant l'inégalité $\mu^+(B) \geqslant \mathcal{I}(\mu(B))$ on obtient $\mu(B_h) \geqslant R_h^\sigma(\mu(B))$ pour tout $h > 0$ assez petit. Pour un tel h, on a donc

$$\mu(B_h) \geqslant R_h^\sigma(\mu(B)) \geqslant R_h^\sigma(R_{h^*}(A)) = R_{h+h^*}^\sigma(A).$$

Mais comme $B_h = (A_{h^*})_h \subset A_{h^*+h}$ (d'après l'inégalité triangulaire), on a alors $\mu(A_{h^*+h}) \geqslant R_{h+h^*}^\sigma(A)$, ce qui montre que $h + h^* \in \Delta_A$ et contredit la définition de h^*.

3. Supposons que $\Delta_A =]0\,;h^*[$ avec $h^* \in]0\,;+\infty[$, et considérons une suite $(h_n) \subset \Delta$ qui croît vers h^*. Pour tout entier n, on a

$$\mu(A_{h_n}) \geqslant R_{h_n}^\sigma(\mu(A)).$$

Or $\lim_{n\to\infty} R_{h_n}^\sigma(\mu(A)) = R_h^\sigma(\mu(A))$ par continuité, et comme $A_{h^*} = \bigcup_n A_{h_n}$, on a $\lim_{n\to\infty} \mu(A_{h_n}) = \mu(A_{h^*})$. En passant à la limite dans l'inégalité précédente, on voit alors que $h^* \in \Delta_A$, ce qui contredit l'hypothèse.

À ce stade, on a montré que pour tout borélien $A \subset \mathbb{R}^d$ de mesure $\mu(A) \in]0\,;1[$ et pour tout $h > 0$, $\sigma > 1$, on a

$$\mu(A_h) \geqslant \Phi\left(\Phi^{-1}(\mu(A)) + \frac{h}{\sigma}\right).$$

On peut alors conclure en faisant tendre σ vers 1 et en utilisant la continuité de Φ. $\square$

2.2 Deux inégalités isopérimétriques fonctionnelles

Plutôt que de travailler directement sur les boréliens de $\mathbb{R}^d$, l'approche de S. Bobkov consiste à « observer » l'espace $\mathbb{R}^d$ (ou plus généralement un espace métrique mesuré) au travers des fonctions (localement) lipschitziennes. L'inégalité isopérimétrique gaussienne se transforme alors en une inégalité fonctionnelle pour ces applications lipschitziennes. Il est assez naturel de penser que quitte à pouvoir approcher correctement les fonctions indicatrices de boréliens, l'information apportée par ce point de vue est au moins aussi riche que celle donnée initialement. On a besoin de la définition ci-dessous pour formaliser cette idée.

Définition 6. Soit $f\colon \mathbb{R}^d \to \mathbb{R}$ une fonction (localement) lipschitzienne. En tout $x \in \mathbb{R}^d$, on définit une *norme de gradient* en x, notée $|\nabla f|(x)$, comme la constante de Lipschitz locale optimale au voisinage de x :

$$|\nabla f|(x) \overset{\text{déf.}}{=} \limsup_{y \to x} \frac{|f(x) - f(y)|}{|x - y|} < +\infty.$$

Remarque. Lorsque f est différentiable au voisinage du point a, la norme de gradient $|\nabla f|(a)$ coïncide avec la norme $|\nabla f(a)|$ du gradient usuel de f en a, ce qui justifie la notation. En effet, on a un développement limité

$$f(x + a) = f(a) + \nabla f(a) \cdot x + o(x)$$

pour x au voisinage de 0. Donc, d'après l'inégalité de Cauchy-Schwarz,

$$|f(x + a) - f(a)| \leqslant |\nabla f(a) \cdot x| + o(x) \leqslant |\nabla f(a)|\,|x| + o(x).$$

D'où $|\nabla f|(a) \leqslant |\nabla f(a)|$. Réciproquement, en prenant $x = h\,\nabla f(a)$, on a que

$$|\nabla f|(a) \geqslant \lim_{h \to 0} \frac{|f(a + h\,\nabla f(a)) - f(a)|}{|h||\nabla f(a)|} = \frac{\nabla f(a) \cdot \nabla f(a)}{|\nabla f(a)|} = |\nabla f(a)|.$$

Remarque. La fonction $x \in \mathbb{R}^d \mapsto |\nabla f|(x)$ est mesurable. En effet, on peut écrire

$$|\nabla f|(x) = \lim_{n \to \infty} S_n(x), \qquad S_n(x) \overset{\text{déf.}}{=} \sup_{y \in B(x, 1/n)} \frac{|f(x) - f(y)|}{|x - y|},$$

et les S_n sont semi-continues inférieurement donc mesurables.

Le théorème suivant donne, sous la forme d'une inégalité fonctionnelle sur les fonctions lipschitziennes, une condition suffisante qui entraîne l'isopérimétrie gaussienne. Quitte à admettre temporairement l'inégalité isopérimétrique pour la mesure gaussienne (théorème 1), il y a même équivalence.

Théorème 3. *Soit μ une mesure borélienne de probabilité sur $\mathbb{R}^d$. Pour que μ vérifie une inégalité isopérimétrique gaussienne de constante $c > 0$, il faut et il suffit que pour toute fonction $f\colon \mathbb{R}^d \to [0\,;1]$ lipschitzienne on ait*

$$\mathcal{I}\left(\int f\,d\mu\right) \leqslant \int \left(\mathcal{I}(f) + \frac{1}{c}|\nabla f|\right) d\mu.$$

Démonstration de la condition suffisante. Supposons que l'inégalité fonctionnelle est vérifiée pour toute fonction lipschitzienne à valeurs dans $[0\,;1]$. Il suffit de vérifier l'inégalité isopérimétrique pour A fermé de $\mathbb{R}^d$. Considérons la famille des fonctions à valeurs dans $[0\,;1]$ définies pour tous $h > 0$ et $x \in \mathbb{R}^d$ par

$$f_h(x) = \left(1 - \frac{d(x, A_{h^2})}{h}\right)_+$$

et dont on vérifie qu'elle converge en tout point vers l'indicatrice de A lorsque h tend vers 0. Rappelons de plus que la distance à une partie est une fonction 1-lipschitzienne. Pour tous $x, y \in \mathbb{R}^d$, on a donc

$$1 - \frac{1}{h}d(x, A_{h^2}) \leqslant 1 - \frac{1}{h}d(y, A_{h^2}) + \frac{1}{h}|x-y|,$$

ce qui entraîne que $f_h(x) \leqslant f_h(y) + \frac{1}{h}|x-y|$ par croissance de $u \in \mathbb{R} \mapsto u_+$ et d'après l'inégalité triangulaire. Ceci montre (x et y jouent des rôles symétriques) que chaque f_h est $1/h$-lipschitzienne, c'est à dire qu'on a $|\nabla f_h| \leqslant 1/h$ uniformément sur $\mathbb{R}^d$. De plus, $|\nabla f_h| = 0$ en dehors de l'adhérence de A_{h+h^2} et sur A_{h^2}, donc

$$\mathcal{I}\left(\int f_h \, d\mu\right) - \int \mathcal{I}(f_h)\, d\mu \leqslant \frac{1}{c}\frac{\mu(\overline{A_{h+h^2}}) - \mu(A_{h^2})}{h} \leqslant \frac{1}{c}\frac{\mu(\overline{A_{h+h^2}}) - \mu(A)}{h}.$$

Puisque $\mathcal{I}$ est continue avec $\mathcal{I}(0) = \mathcal{I}(1) = 0$, le théorème de convergence dominée de Lebesgue montre que le terme de gauche tend vers $\mathcal{I}(\mu(A))$ lorsque h tend vers 0. Alors, en passant à la limite inférieure au moyen du lemme suivant, on obtient

$$\mathcal{I}(\mu(A)) \leqslant \frac{1}{c}\mu^+(A).$$

□

Lemme 3. *Pour tout borélien $A \subset \mathbb{R}^d$, on a*

$$\mu^+(A) = \liminf_{h\to 0} \frac{\mu(\overline{A_{h+o(h)}}) - \mu(A)}{h}.$$

Démonstration. Soit $A \subset \mathbb{R}^d$ un borélien quelconque, et soit $\sigma > 1$ un paramètre fixé. Alors, pour tout h assez petit, on a $A_h \subset \overline{A_{h+o(h)}} \subset A_{\sigma h}$, et donc

$$\frac{\mu(A_h) - \mu(A)}{h} \leqslant \frac{\mu(\overline{A_{h+o(h)}}) - \mu(A)}{h} \leqslant \sigma \times \frac{\mu(A_{\sigma h}) - \mu(A)}{\sigma h}.$$

En passant à la limite inférieure quand $h \to 0$, il vient

$$\mu^+(A) \leqslant \liminf_{h\to 0} \frac{\mu(\overline{A_{h+o(h)}}) - \mu(A)}{h} \leqslant \sigma\mu^+(A), \qquad \forall \sigma > 1.$$

On conclut alors en faisant tendre σ vers 1. □

Afin de prouver la seconde implication du théorème, on commence par établir un lemme qui relie l'intégrale de la norme de gradient et μ^+.

Lemme 4 (Inégalité de la co-aire)**.** *Soit $f\colon \mathbb{R}^d \to [0\,;1]$ une fonction lipschitzienne. On a*

$$\int |\nabla f|\, d\mu \geqslant \int_0^1 \mu^+\{f > t\}\, dt.$$

Démonstration. Considérons la famille de fonctions $(f_h)_{h>0}$ définies par

$$f_h(x) \stackrel{\text{déf.}}{=} \sup_{|x-y|<h} f(y), \qquad x \in \mathbb{R}^d.$$

Notons par ailleurs $A^t = \{f > t\} = \{x \in \mathbb{R}^d : f(x) > t\}$. Pour tout $t \geqslant 0$, l'ensemble $\{f_h > t\} = (A^t)_h$ est un ouvert, ce qui montre que les f_h sont semi-continues inférieurement. De plus, le théorème de Fubini-Tonelli montre que

$$\int f_h\, d\mu = \int_0^1 \mu\{f_h > t\}\, dt = \int_0^1 \mu(A^t_h)\, dt,$$

et de même

$$\int f\, d\mu = \int_0^1 \mu(A^t)\, dt,$$

de sorte que

$$\int \frac{f_h - f}{h}\, d\mu = \int_0^1 \frac{\mu(A^t_h) - \mu(A^t)}{h}\, dt.$$

D'après le lemme de Fatou et par définition de μ^+ on a alors

$$\int_0^1 \mu^+(A^t)\, dt \leqslant \liminf_{h\to 0} \int \frac{f_h - f}{h}\, d\mu.$$

Or, par définition de f_h, on a pour tout $x \in \mathbb{R}^d$

$$\frac{f_h(x) - f(x)}{h} = \sup_{|x-y|<h} \frac{|f(y) - f(x)|}{h} \leqslant \sup_{|y-x|<h} \frac{|f(y) - f(x)|}{|x-y|} \underset{h\to 0}{\longrightarrow} |\nabla f|(x)$$

avec majoration par la constante de Lipschitz de f. Le théorème de convergence dominée montre alors que

$$\int_0^1 \mu^+(A^t)\, dt \leqslant \liminf_{h\to 0} \int \sup_{|y-x|<h} \frac{|f(x) - f(y)|}{|x-y|}\, d\mu(x) = \int |\nabla f|\, d\mu.$$

$\square$

Démonstration de la condition nécessaire. Supposons que μ vérifie une inégalité isopérimétrique gaussienne de constante $c > 0$, c'est à dire que pour tout borélien A on a

$\mu^+(A) \geqslant c\mathcal{I}(\mu(A))$. Soit $f\colon \mathbb{R}^d \to [0\,;1]$ une fonction K-lipschitzienne fixée ($K > 0$). D'après l'inégalité de la co-aire, on a alors

$$\int |\nabla f|\, d\mu \geqslant \int_0^1 \mu^+\{f > t\}\, dt \geqslant c \int_0^1 \mathcal{I}(\mu\{f > t\})\, dt.$$

Soit alors $\nu = \mu \circ f^{-1}$ la mesure sur $[0\,;1]$, image de μ par f. Quitte à régulariser ν (voir par exemple [BL96]), on peut supposer qu'elle est absolument continue par rapport à la mesure de Lebesgue, de densité n strictement positive sur $[0\,;1]$. Soit alors $N(t) = \nu([0\,;t])$ sa fonction de répartition et $k\colon \mathbb{R} \to [0\,;1]$ la fonction $k = N^{-1} \circ \Phi$ prolongée par $k(-\infty) = 0$ et $k(+\infty) = 1$ qui envoie la mesure gaussienne γ_1 sur la mesure ν. Comme par hypothèse, μ vérifie une inégalité isopérimétrique gaussienne de constante c, pour le borélien $A = \{f \leqslant t\}$ qui vérifie $A_s \subset \{f \leqslant t + Ks\}$ et si s est assez petit, on a

$$\frac{N(t+Ks) - N(t)}{s} \geqslant \frac{\mu(A_s) - \mu(A)}{s} \geqslant \frac{\Phi(\Phi^{-1}(\mu(A)) + cs) - \Phi(\Phi^{-1}(\mu(A)))}{s}$$

En faisant tendre s vers 0 on obtient alors que $n(t) \geqslant \frac{c}{K}\mathcal{I}(N(t))$ pour tout $t \in \,]0\,;1[$. En particulier, ceci indique que pour $t = k(x)$,

$$k'(x) = \frac{\Phi'(x)}{n(N^{-1}(\Phi(x)))} = \frac{\Phi'(\Phi^{-1}(N(t)))}{n(t)} \leqslant \frac{K}{c}.$$

La fonction k est donc lipschitzienne et on peut lui appliquer l'inégalité isopérimétrique pour γ_1 (temporairement admise pour cette démonstration) :

$$\mathcal{I}\left(\int k\, d\gamma_1\right) \leqslant \int (\mathcal{I}(k) + k')\, d\gamma_1.$$

On vérifie alors que

$$\int k\, d\gamma_1 = \int_0^1 t\, d\nu(t) = \int f\, d\mu$$

et de la même manière,

$$\int \mathcal{I}(f)\, d\gamma_1 = \int_0^1 \mathcal{I}\, d\nu = \int \mathcal{I}(f)\, d\mu.$$

Enfin, en faisant le changement de variable $t = k(x)$ et en utilisant le fait que $\mathcal{I}(N(t)) = \mathcal{I}(\mu\{f \leqslant t\}) = \mathcal{I}(\mu\{f > t\})$ par symétrie de $\mathcal{I}$, on obtient

$$\int k'\, d\gamma_1 = \int k'(x)\Phi'(x)\, dx = \int_0^1 \Phi'(k^{-1}(t))\, dt = \int_0^1 \mathcal{I}(N(t))\, dt = \int_0^1 \mathcal{I}(\mu\{f > t\})\, dt,$$

ce qui donne finalement

$$\mathcal{I}\left(\int f\, d\mu\right) \leqslant \int \left(\mathcal{I}(f) + \frac{1}{c}|\nabla f|\right) d\mu.$$

$\square$

Remarque (importante). Au passage, on a montré que si une mesure borélienne de probabilité μ vérifie une inégalité isopérimétrique gaussienne de constante $c > 0$, alors l'image de μ par toute application K-lipschitzienne à valeurs dans $\mathbb{R}$ est une mesure qui s'obtient comme image de la mesure gaussienne γ_1 par une application K/c-lipschitzienne.

En particulier, toute mesure de probabilité μ sur $\mathbb{R}$ qui vérifie l'inégalité isopérimétrique gaussienne de constante $c = 1$ est une contraction de la mesure gaussienne. Bien entendu, la réciproque est vraie : si $\mu = \gamma_1 \circ f$ avec $f : \mathbb{R} \to \mathbb{R}$ une fonction 1-lipschitzienne, alors pour tout borélien $A \subset \mathbb{R}$, on a $\{f \in A\}_h \subset \{f \in A_h\}$ et donc pour tout $h > 0$,

$$\frac{\mu(A_h) - \mu(A)}{h} \geqslant \frac{\gamma_1\{f \in A\}_h - \gamma_1\{f \in A\}}{h},$$

ce qui donne à la limite $\mu^+(A) \geqslant \gamma_1^+\{f \in A\} \geqslant \mathcal{I}(\mu(A))$. On a donc caractérisé les mesures de probabilités sur $\mathbb{R}$ qui vérifient l'inégalité isopérimétrique gaussienne. Il faudrait se garder de croire que la situation est aussi simple en dimension supérieure, comme on le verra au chapitre 4.

On introduit maintenant une seconde inégalité fonctionnelle, cette fois-ci quadratique et apparemment plus forte, qui entraîne elle aussi l'isopérimétrie gaussienne.

Corollaire 1. *Soit μ une mesure borélienne de probabilité sur $\mathbb{R}^d$. Pour que μ vérifie une inégalité isopérimétrique gaussienne de constante $c > 0$, il suffit que pour toute fonction $f : \mathbb{R}^d \to [0\,;1]$ lipschitzienne, on ait*

$$\mathcal{I}\left(\int f\,d\mu\right) \leqslant \int \sqrt{\mathcal{I}^2(f) + \frac{1}{c^2}|\nabla f|^2}\,d\mu.$$

Remarque. Cette formulation fonctionnelle est en fait toujours équivalente à l'inégalité isopérimétrique gaussienne (voir par exemple [BM00]).

Démonstration. Il suffit d'utiliser la majoration grossière $\sqrt{u+v} \leqslant \sqrt{u} + \sqrt{v}$ et d'appliquer la proposition précédente. □

On verra rapidement dans la suite que cette inégalité fonctionnelle se comporte bien vis à vis de la tensorisation (contrairement à la première), ce qui la rendra plus facile à manipuler.

2.3 Étude du profil isopérimétrique gaussien

Cette section est dévolue à l'étude de la fonction $\mathcal{I}$ intervenant dans l'inégalité isopérimétrique pour la mesure gaussienne. Les propriétés que l'on va établir seront bien utiles dans la preuve de l'inégalité isopérimétrique de la section suivante. Considérons donc la fonction $\mathcal{I}$ de $[0\,;1]$ dans $\mathbb{R}$ définie par :

$$\mathcal{I}(x) = \Phi' \circ \Phi^{-1}(x) \quad \text{où} \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-t^2/2}\,\mathrm{d}t = \gamma(]-\infty\,;x]),$$

c'est-à-dire la fonction obtenue en prenant la densité de l'inverse de la fonction de distribution de la mesure gaussienne γ.

Proposition 4. *On a les propriétés élémentaires suivantes :*

(i) $\mathcal{I}$ *s'annule en* 0 *et en* 1 *et reste strictement positive sur* $]0\,;1[$.

(ii) $\mathcal{I}$ *est concave sur* $[0\,;1]$ *et de classe* $\mathscr{C}^\infty$ *sur* $]0\,;1[$.

(iii) $\mathcal{I}$ *vérifie l'équation différentielle* $\mathcal{I}'' \times \mathcal{I} = -1$.

(iv) La fonction $(\mathcal{I}')^2$ *est strictement convexe sur* $]0\,;1[$.

Démonstration. Dans ce qui suit on notera $\phi = \Phi'$ la densité gaussienne. Commençons par remarquer qu'on peut prolonger Φ en un homéomorphisme de $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ vers $[0\,;1]$ en posant $\Phi(-\infty) = 0$ et $\Phi(+\infty) = 1$. La fonction $\mathcal{I}$ est donc bien définie et continue sur $[0\,;1]$ avec $\mathcal{I}(0) = \phi(-\infty) = 0$ et $\mathcal{I}(1) = \phi(+\infty) = 0$. Les théorèmes classiques d'analyse réelle à une variable montrent également que Φ est un $\mathscr{C}^\infty$-difféomorphisme de $\mathbb{R}$ vers $]0\,;1[$, ce qui permet de conclure la preuve des points (i) et (ii). La fonction $\phi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$ vérifie $\phi'(x) = -x\phi(x)$ et on a $(\Phi^{-1})' \times \Phi' \circ \Phi^{-1} = 1$, d'où

$$\mathcal{I}'(x) = \phi' \circ \Phi^{-1}(x) \times (\Phi^{-1})'(x) = -\Phi^{-1}(x), \quad \text{et} \quad \mathcal{I}''(x) = -\frac{1}{\mathcal{I}(x)},$$

ce qui donne (iii). Enfin, pour montrer le point (iv), on dérive deux fois : $(\mathcal{I}'^2)' = 2\mathcal{I}'' \times \mathcal{I}' = -2\mathcal{I}'/\mathcal{I}$, puis $(\mathcal{I}'^2)'' = 2(1+\mathcal{I}'^2)/\mathcal{I}^2$ qui reste donc strictement positif sur $]0\,;1[$. □

L'inégalité qui suit est nettement moins facile à obtenir, ce qui n'est pas très étonnant quand on sait (on le saura en fait à la prochaine section) qu'elle contient déjà en substance l'inégalité isopérimétrique gaussienne dans sa forme générale.

Lemme 5. *Pour tous réels* a *et* b *dans* $[0\,;1]$, *on a*

$$\mathcal{I}\left(\frac{a+b}{2}\right) \leqslant \frac{1}{2}\sqrt{\mathcal{I}(a)^2 + \left|\frac{a-b}{2}\right|^2} + \frac{1}{2}\sqrt{\mathcal{I}(b)^2 + \left|\frac{a-b}{2}\right|^2}. \tag{2.1}$$

Démonstration. On peut toujours supposer que $a \neq b$ (sinon il n'y a rien à démontrer). Posons alors $c = (a+b)/2 \in\,]0\,;1[$ et introduisons la fonction $g(x) = \mathcal{I}(c+x)^2 + x^2$. Si on pose $x = (a-b)/2 \neq 0$, l'inégalité (5) se réécrit sous la forme

$$\sqrt{g(0)} \leqslant \frac{1}{2}\sqrt{g(x)} + \frac{1}{2}\sqrt{g(-x)},$$

ou encore, en passant aux carrés,

$$4g(0) - (g(x) + g(-x)) \leqslant 2\sqrt{g(x)g(-x)}.$$

Si le terme de gauche est négatif on obtient immédiatement

$$\sqrt{g(0)} \leqslant \sqrt{\frac{g(x)+g(-x)}{4}} \leqslant \frac{1}{2}\sqrt{g(x)} + \frac{1}{2}\sqrt{g(-x)},$$

et sinon, l'inégalité est équivalente à

$$16g(0)^2 + (g(x)-g(-x))^2 \leqslant 8g(0)(g(x)+g(-x)).$$

En posant $h(x) = g(x) - g(0) = \mathcal{I}(c+x)^2 + x^2 - \mathcal{I}(c)^2$, ceci se réécrit sous la forme

$$(h(x)-h(-x))^2 \leqslant 8\mathcal{I}(c)^2(h(x)+h(-x)).$$

Commençons par minorer le second membre : $h(x)+h(-x)$. On remarque que la fonction $R(x) = h(x)+h(-x)-2\mathcal{I}'(c)^2x^2$ (qui est un grand $O(x^3)$) est convexe en x. En effet, on a $R'(x) = 2\mathcal{I}(x+c)\mathcal{I}'(x+c) - 2\mathcal{I}(x-c)\mathcal{I}'(x-c) + 4x - 2\mathcal{I}'(c)^2x$ et donc

$$R''(x) = 4\left(\frac{\mathcal{I}'(c+x)^2 - \mathcal{I}'(c-x)^2}{2} - \mathcal{I}'(c)^2\right) \geqslant 0$$

car $(\mathcal{I}')^2$ est convexe. Comme R est paire, on a en particulier $R(x) \geqslant R(0) = 0$, d'où

$$h(x)+h(-x) \geqslant 2\mathcal{I}'(c)^2x^2.$$

Il reste à majorer le terme en $u(x) = h(x)-h(-x) = \mathcal{I}(c+x)^2 - \mathcal{I}(c-x)^2$. Comme $\mathcal{I}$ est symétrique par rapport à $1/2$, quitte à remplacer a et b par $1-a$ et $1-b$, et quitte à les échanger, on peut supposer que $c \leqslant 1/2$ et que $x > 0$. On a alors $\mathcal{I}(c+x) \geqslant \mathcal{I}(c-s)$ par symétrie car $\mathcal{I}$ est croissante sur $[0\,;1/2]$ et décroissante sur $[1/2\,;1]$. Inversement, comme elle est convexe (et toujours symétrique), la fonction $(\mathcal{I}')^2$ est décroissante sur $]0\,;1/2]$ et croissante sur $[1/2\,;0[$, ce qui montre que $\mathcal{I}'(c+x)^2 \leqslant \mathcal{I}'(c-x)^2$. On en déduit que $u''(x) = 2(\mathcal{I}'(c+x)^2 - \mathcal{I}'(c-x)^2) \leqslant 0$, c'est à dire que u est concave sur $[0\,;c]$. On en déduit que $x \in\,]0\,;c] \mapsto (u(x)-u(0))/x$ est décroissante. Or $u(0) = 0$ et $u'(0) = 4\mathcal{I}(c)\mathcal{I}'(c)$, donc

$$\left(\frac{h(x)-h(-x)}{x}\right)^2 \leqslant 16\mathcal{I}(c)^2\mathcal{I}'(c)^2 \leqslant 8\frac{h(x)+h(-x)}{x^2},$$

ce qui nous permet de conclure. □

2.4 De l'isopérimétrie discrète à l'isopérimétrie gaussienne

Dans toute cette section, $(\Omega, \mathscr{F}, \mathbb{P})$ désignera un espace de probabilité sur lequel seront définies les variables aléatoires. Nous commençons par interpréter la dernière inégalité obtenue sous un jour plus probabiliste. Soit X une variable aléatoire sur $(\Omega, \mathscr{F}, \mathbb{P})$ à valeur dans $\{-1,1\}$ telle que $\mathbb{P}(X=1)=\mathbb{P}(X=-1)=1/2$. Pour toute fonction $f\colon\{-1,1\}\to[0\,;1]$ on a alors (en prenant $a=f(1)$ et $b=f(-1)$) l'inégalité

$$I\left(\mathbb{E}[f(X)]\right)\leqslant\mathbb{E}\left[\sqrt{I(f(X))^2+|\nabla_1 f|^2(X)}\right],\tag{2.2}$$

où la valeur absolue du gradient discret $|\nabla_1 f|$ est la quantité définie par $|\nabla_1 f|=\frac{|f(1)-f(-1)|}{2}$. Si on note $\mu=(\delta_1+\delta_{-1})/2$ la loi de X, cette inégalité peut aussi se réécrire sous la forme

$$I\left(\int f\,\mathrm{d}\mu\right)\leqslant\int\sqrt{I(f)^2+|\nabla_1 f|^2}\,\mathrm{d}\mu.$$

Ainsi reconnaît-on une forme fonctionnelle de l'isopérimétrie (pour la mesure μ) dont il a déjà été question. Notre objectif se ramène donc à voir comment étendre cette inégalité aux mesures gaussiennes. Nous commençons par traiter le cas (toujours discret) des fonctions définies sur l'hypercube $\{-1,1\}^n$. Pour une telle fonction $f\colon\{-1,1\}^n\to[0\,;1]$, on pose

$$|\nabla_n f|^2(x_1,\dots,x_n)=\frac{1}{4}\sum_{i=1}^n|f(x_1,\dots,x_i,\dots,x_n)-f(x_1,\dots,-x_i,\dots,x_n)|^2.$$

L'inégalité (2.2) possède une intéressante propriété d'additivité par tensorisation de l'espace de probabilité (due essentiellement au fait qu'on travaille avec le carré de la norme euclidienne). C'est ce qui rend possible le passage d'une inégalité sur l'espace à deux points, à une égalité sur un hypercube de dimension supérieure. C'est l'objet du lemme suivant.

Lemme 6. *Soit $I\colon[0\,;1]\to\mathbb{R}_+$ qui vérifie l'inégalité (2.2) pour toute fonction de $\{-1,1\}$ dans $[0\,;1]$. Soient $n\geqslant 1$ un entier, et $X_1,\dots,X_n$ des variables aléatoires indépendantes et de même loi μ. Alors, pour toute fonction $f:\{-1,1\}^n\to[0\,;1]$,*

$$I\left(\mathbb{E}[f(X_1,\dots,X_n)]\right)\leqslant\mathbb{E}\left[\sqrt{I(f(X_1,\dots,X_n))^2+|\nabla_n f|^2(X_1,\dots,X_n)}\right],\tag{2.3}$$

Remarque. La conclusion de ce lemme exprime ici encore une forme fonctionnelle de l'isopérimétrie pour la mesure produit $\mu^{\otimes n}$ sur l'espace $\{-1,1\}^n$.

Démonstration. On montre le lemme par récurrence sur l'entier $n\geqslant 1$, le cas $n=1$ correspondant bien sûr à l'hypothèse. On va voir que c'est aussi cette hypothèse qui donne l'hérédité de la propriété.

Supposons donc le résultat démontré au rang $n-1$ avec $n\geqslant 2$. Soient $X_1,\dots,X_n$ des variables aléatoires indépendantes et soit $f\colon\{-1,1\}^n\to[0\,;1]$ une fonction. Par indépendance, on a

$$\mathbb{E}[f(X_1,\dots,X_n)\mid X_n]=h(X_n)\quad\mathbb{P}-\text{p.s.}\quad\text{avec}\quad h(x)=\mathbb{E}[f(X_1,\dots,X_{n-1},x)].$$

L'inégalité (2.2), c'est à dire le cas $n = 1$ appliqué avec la fonction $h\colon \{-1,1\} \to [0\,;1]$, donne

$$I(\mathbb{E}[f(X_1,\dots,X_n)]) = I(\mathbb{E}[h(X_n)]) \leqslant \mathbb{E}\left[\sqrt{I(h(X_n))^2 + |\nabla_1 h|^2(X_n)}\right]. \qquad (*)$$

De même, en travaillant conditionnellement à X_n, l'hypothèse de récurrence au rang $n-1$ appliquée à la fonction $g_{X_n}\colon (x_1,\dots,x_{n-1}) \mapsto f(x_1,\dots,x_{n-1},X_n)$ définie sur $\{-1,1\}^{n-1}$ nous donne

$$I(h(X_n)) \leqslant \mathbb{E}\left[\sqrt{I(g_{X_n}(X_1,\dots,X_{n-1}))^2 + |\nabla_{n-1} g_{X_n}|^2(X_1,\dots,X_{n-1})} \mid X_n\right]. \qquad (**)$$

Il ne reste plus qu'à remarquer que $g_{X_n}(X_1,\dots,X_{n-1}) = f(X_1,\dots,X_n)$, que

$$|\nabla_1 h|^2(X_n) = \frac{1}{4}\,\mathbb{E}\left[f(X_1,\dots,X_{n-1},X_n) - f(X_1,\dots,X_{n-1},-X_n) \mid X_n\right]^2$$

et enfin que

$$|\nabla_{n-1} g_{X_n}|^2(X_1,\dots,X_{n-1}) = \frac{1}{4}\sum_{i=1}^{n-1}\left(f(X_1,\dots,X_i,\dots,X_n) - f(X_1,\dots,-X_i,\dots,X_n)\right)^2.$$

Finalement, en combinant $(*)$ et $(**)$, puis en appliquant l'inégalité de Minkowski [*] $\mathbb{E}[\sqrt{u}]^2 + \mathbb{E}[\sqrt{v}]^2 \leqslant \mathbb{E}[\sqrt{u+v}]^2$ pour l'espérance conditionnelle, on obtient

$$I(\mathbb{E}[f(X_1,\dots,X_n)]) \leqslant \mathbb{E}\left[\mathbb{E}\left[\sqrt{I(f(X_1,\dots,X_n))^2 + |\nabla_n f|^2(X_1,\dots,X_n)} \mid X_n\right]\right],$$

ce qui donne le résultat au rang n et permet de conclure la preuve. □

Remarque (importante). La seule hypothèse requise sur la mesure μ dans ce lemme est l'inégalité (2.2). Ainsi pourrait-on par-exemple énoncer le même résultat en remplaçant $\{-1,1\}$ par un espace euclidien, et en prenant des fonctions continument différentiables, voire lipschitziennes.

Ce résultat discret étant établi pour le profil isopérimétrique gaussien $\mathcal{I}$, nous allons maintenant voir comment une méthode d'approximation permet de l'étendre au cas continu.

∗. La « norme » $L^{1/2}$ n'est pas une norme mais vérifie tout de même une inégalité triangulaire renversée. Il suffit pour le voir d'intégrer l'inégalité de concavité

$$\sqrt{\frac{u+v}{\mathbb{E}[\sqrt{u}]^2 + \mathbb{E}[\sqrt{v}]^2}} \geqslant \alpha\sqrt{\frac{u}{\mathbb{E}[\sqrt{u}]^2}} + (1-\alpha)\sqrt{\frac{v}{\mathbb{E}[\sqrt{v}]^2}}, \text{ où } \alpha = \frac{\mathbb{E}[\sqrt{u}]^2}{\mathbb{E}[\sqrt{u}]^2 + \mathbb{E}[\sqrt{v}]^2}.$$

Théorème 4 (Isopérimétrie gaussienne, version fonctionnelle)**.** *Soit $n \geqslant 1$ un entier. Pour toute fonction lipschitzienne $f \colon \mathbb{R}^n \to [0\,;1]$,*

$$I\left(\int f\,\mathrm{d}\gamma_n\right) \leqslant \int \sqrt{I(f)^2 + |\nabla f|^2}\,\mathrm{d}\gamma_n. \tag{2.4}$$

De manière équivalente, si $X = (X_1, \ldots, X_n)$ est un vecteur gaussien $\mathcal{N}(0, I_n)$,

$$I(\mathbb{E}[f(X)]) \leqslant \mathbb{E}\left[\sqrt{I(f(X))^2 + |\nabla f|^2(X)}\right]. \tag{2.5}$$

Démonstration. Nous prouvons le résultat dans sa version probabiliste. Rappelons que ∇_m désigne le gradient discret pour les fonction définies sur $\{-1,1\}^m$ tel que défini précédemment. Commençons par supposer que $n = 1$ et que f est une fonction bornée de classe $\mathscr{C}^2$ avec ses dérivées partielles premières et secondes bornées. Soient $\xi_1, \xi_2, \ldots$ une suite de variables aléatoires indépendantes de loi commune donnée par $\mathbb{P}(\xi_1 = 1) = \mathbb{P}(\xi_1 = -1) = 1/2$. On pose

$$Z_k = \frac{\xi_1 + \cdots + \xi_k}{\sqrt{k}},$$
$$f_k(x_1, \ldots, x_k) = f\left(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\right),$$

de sorte que $f(Z_k) = f_k(\xi_1, \ldots, \xi_k)$. Le résultat discret sur l'hypercube s'applique à f_k, et on en tire donc que

$$I\left(\mathbb{E}[f(Z_k)]\right) \leqslant \mathbb{E}\left[\sqrt{I(f(Z_k))^2 + |\nabla_k f_k|^2(\xi_1, \ldots, \xi_k)}\right].$$

Par ailleurs, le théorème central limite assure la convergence en loi de la suite (Z_k) vers une variable aléatoire X de loi normale centrée réduite $\mathcal{N}(0,1)$. En particulier,

$$I(\mathbb{E}[f(Z_k)]) \underset{k\to\infty}{\longrightarrow} I(\mathbb{E}[f(X)]) \quad \text{et}$$
$$\forall \varepsilon > 0, \quad \mathbb{E}\left[\sqrt{I(f(Z_k))^2 + |\nabla f(Z_k)|^2 + \varepsilon}\right] \underset{k\to\infty}{\longrightarrow} \mathbb{E}\left[\sqrt{I(f(X))^2 + |\nabla f(X)|^2 + \varepsilon}\right].$$

Il reste à contrôler l'approximation du gradient usuel par le gradient discret. On a

$$\begin{aligned}
|\nabla_k f_k(x_1, \ldots, x_k)|^2 &= \frac{1}{4}\sum_{i=1}^{k}\left(f\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\Big) - f\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}} \pm \frac{2}{\sqrt{k}}\Big)\right)^2 \\
&= \left|f'\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\Big)\right|^2 + O\Big(\frac{1}{\sqrt{k}}\Big)
\end{aligned}$$

uniformément en x, en utilisant le développement de Taylor-Lagrange (uniforme en x)

$$\left|f\left(x \pm \frac{2}{\sqrt{k}}\right) - f(x)\right| = \frac{2}{\sqrt{k}}\,|f'(x)| + O\Big(\frac{1}{k}\Big).$$

On en déduit en passant aux limites dans l'inégalité discrète que pour tout $\varepsilon > 0$,

$$I(\mathbb{E}[f(X)]) = \limsup_{k\to\infty} I(\mathbb{E}[f(Z_k)]) \leqslant \mathbb{E}\left[\sqrt{I(f(X))^2 + |\nabla f(X)|^2 + \varepsilon}\right],$$

ce qui donne finalement par convergence monotone l'inégalité

$$I(\mathbb{E}[f(X)]) \leqslant \mathbb{E}\left[\sqrt{I(f(X))^2 + |\nabla f(X)|^2}\right].$$

Pour passer au cas multidimensionnel [*], il suffit alors d'invoquer la version du lemme d'additivité (lemme 6) discutée dans la remarque qui le suit.

Enfin, si la fonction $f\colon \mathbb{R}^n \to [0\,;1]$ est seulement supposée lipschitzienne, alors quitte à la tronquer, on peut supposer que son support est compact (on peut trouver des compacts de $\mathbb{R}^d$ de probabilité arbitrairement grande). À partir de là, un argument standard de régularisation par convolution (avec un noyau gaussien par exemple) permet de se ramener au cas traité dans la preuve ci-dessus. □

$*$. Dans l'article original [Bob97], S. Bobkov utilise directement la version multidimensionnelle du théorème central limite. Notre argument a le mérite de montrer une nouvelle fois l'intérêt du lemme 6.

Chapitre 3

Méthode de semi-groupe : Ornstein-Uhlenbeck

Nous commençons par rappeler quelques définitions et propriétés générales concernant les semi-groupes de Feller et leurs générateurs infinitésimaux. On pourra en trouver une présentation beaucoup plus détaillée dans [Lig10] et [RY99] dont nous nous sommes très fortement inspirés.

3.1 Semi-groupes et processus de Markov

Soient $B(\mathbb{R}^d)$ l'espace des fonctions boréliennes bornées de $\mathbb{R}^d$ dans $\mathbb{R}$ et $C_0 = C_0(\mathbb{R}^d)$ le sous-espace des fonctions continues de $\mathbb{R}^d$ dans $\mathbb{R}$ qui s'annulent à l'infini [*]. On munit ces deux espaces de la norme de la convergence uniforme définie pour toute fonction $f \in B(\mathbb{R}^d)$ par

$$|f| = \sup_{x \in \mathbb{R}^d} |f(x)|,$$

qui leur confère une structure d'espace de Banach.

Définition 7. Un *semi-groupe de transition* (ou de Markov) est une famille $\{P_t, t \geqslant 0\}$ d'opérateurs linéaires positifs sur l'espace tels que

(i) $P_0 = \mathrm{Id}$ et $\|P_t\| \leqslant 1$ pour tout $t \geqslant 0$; [†]

(ii) $P_{t+s} = P_t \circ P_s$ pour tous $t, s \geqslant 0$;

(iii) $P_t \mathbf{1} = \mathbf{1}$ pour tout $t \geqslant 0$;

(iv) $(t, x) \mapsto P_t f(x)$ est $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}^d)$-mesurable pour toute $f \in B(\mathbb{R}^d)$.

*. On dit que $f \colon \mathbb{R}^d \to \mathbb{R}$ s'annule à l'infini si $\lim_{|x| \to +\infty} f(x) = 0$.

†. Où $\|\cdot\|$ désigne la norme d'opérateur sur $B(\mathbb{R}^d)$.

Les semi-groupes de transition peuvent être naturellement associés à une classe de processus aléatoires très importante.

Définition 8. Soit $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ un espace de probabilité filtré. Un *processus de Markov (homogène)* X sur $\mathbb{R}^d$ par rapport à $(\mathscr{F}_t)$ de *semi-groupe de transition* (P_t) est un processus adapté tel que, pour toute fonction borélienne bornée $f\colon \mathbb{R}^d \to \mathbb{R}$ et pour tous réels $t, s \geqslant 0$,

$$\mathbb{E}[f(X_{t+s}) \mid \mathscr{F}_s] = P_t f(X_s) \quad \mathbb{P}\text{-p.s.}$$

Grâce au théorème d'extension de Kolmogorov, on est toujours capable construire de une version canonique d'un processus de Markov sur $\mathbb{R}^d$ de semi-groupe de transition (P_t) donné (à l'avance) : pour toute probabilité *initiale* ν sur $\mathbb{R}^d$, il existe une unique probabilité $\mathbb{P}_\nu$ sur l'espace des trajectoires à valeurs dans $\mathbb{R}^d$ sous laquelle le processus canonique $\{X_t\colon \omega \mapsto \omega(t), t \geqslant 0\}$ est de Markov par rapport à la filtration naturelle avec (P_t) pour fonction de transition, et X_0 de loi ν. Un semi-groupe de transition peut donc toujours être vu comme le semi-groupe de transition associé à un processus de Markov. Cette remarque nous permettra d'établir des propriétés sur le semi-groupe par des raisonnements de nature probabiliste et vice-versa. Pour le choix de loi initiale $\nu = \delta_x$ avec $x \in \mathbb{R}^d$, on notera (X_t^x) le processus associé, et

$$\mathbb{E}_x[f(X_t)] \overset{\text{déf.}}{=} \mathbb{E}[f(X_t^x)] = P_t f(x).$$

On distingue maintenant un type de semi-groupe de transition qui joue un rôle important en théorie des processus.

Définition 9.

1. Un *semi-groupe de Feller* $\{P_t, t \geqslant 0\}$ est un semi-groupe de transition vérifiant
   (i) $P_t f \in C_0$ pour tous $f \in C_0$ et $t \geqslant 0$;
   (ii) $\lim_{t\to 0} |P_t f - f| = 0$ pour tout $f \in C_0$;
2. Un *processus de Feller* est un processus de Markov donc le semi-groupe de transtition est de Feller.

Remarque. La condition (ii) de continuité forte pour la norme $|\cdot|$ peut être remplacée de manière équivalente par la propriété (qui ne demande qu'une convergence ponctuelle) : pour tout $f \in C_0$ et pour tout $x \in \mathbb{R}^d$, $\lim_{t\to 0} P_t f(x) = f(x)$. On pourra par exemple consulter [RY99], chapitre III, proposition (2.4). Notons également qu'une famille d'opérateurs de C_0 qui vérifie toutes les conditions pour être un semi-groupe de Feller *a priori* seulement sur C_0 s'étend en fait de manière unique en un semi-groupe de transition.

Exemple. Le mouvement brownien (B_t) dans $\mathbb{R}^d$ est un processus de Feller par rapport à sa filtration naturelle $\mathscr{F}_t^B = \sigma(B_u, u \leqslant t)$ dont le semi-groupe de transition est le *semi-groupe de la chaleur* défini par

$$P_t f(x) = \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} f(y) \exp -\frac{|y-x|^2}{2t} dy$$

3.1.1 Générateur infinitésimal

De part sa structure additive en temps, toute « l'information » d'un semi-groupe est contenue dans des intervalles de temps arbitrairement petits. Parallèlement, il est plus facile de comprendre et décrire le comportement d'un processus à partir de son évolution infinitésimale qu'à partir de son semi-groupe. Les notions qui suivent correspondent à ce point de vue.

Définition 10. Soit $\{P_t, t \geqslant 0\}$ un semi-groupe de Feller. Une fonction $f \in C_0$ appartient au *domaine* $\mathcal{D}(L)$ du générateur infinitésimal de P_t si la limite

$$Lf \overset{\text{déf.}}{=} \lim_{t \to 0} \frac{1}{t}(P_t f - f)$$

existe dans C_0. L'opérateur linéaire $L \colon \mathcal{D}(L) \to C_0$ ainsi défini est appelé le *générateur infinitésimal* du semi-groupe P_t.

Si X est un processus de Feller par rapport $(\mathscr{F}_t)$ de semi-groupe P_t, et si f est dans le domaine $\mathscr{D}(L)$ on a par définition

$$\mathbb{E}[f(X_{t+h}) - f(X_t) \mid \mathscr{F}_t] = P_h f(X_t) - f(X_t) = hLf(X_t) + o(h).$$

Le générateur permet donc de décrire (au moins partiellement) la manière dont le processus évolue au voisinage d'un point sur des intervalles de temps infinitésimaux.

On énonce quelques propriétés du générateur L sans les démontrer (on pourra trouver les preuves dans [RY99], chapitre VII, propositions (1.2) et (1.3)).

Proposition 5. *Soit $f \in \mathcal{D}(L)$, alors*

(i) pour tout $t \geqslant 0$, $P_t f \in \mathcal{D}(L)$;

(ii) la fonction $t \mapsto P_t f$ est (fortement) dérivable dans C_0 et

$$\frac{d}{dt} P_t f = L P_t f = P_t L f \ ;$$

(iii) pour tout $t \geqslant 0$, $P_t f - f = \int_0^t P_s L f \, ds = \int_0^t L P_s f \, ds$;

(iv) l'espace $\mathcal{D}(L)$ est dense dans C_0 et L est un opérateur fermé.*

Exemple. Pour le semi-groupe de la chaleur en dimension $d = 1$, le domaine $\mathcal{D}(L)$ est exactement l'espace C_0^2 des fonctions deux fois différentiables sur $\mathbb{R}$ telles que f, f' et f'' sont dans C_0, et on a $Lf = \frac{1}{2} f''$ sur C_0^2. Si $d > 1$, il n'est plus vrai que $\mathcal{D}(L) = C_0^2$, mais on peut montrer que $\mathcal{D}(L)$ est le sous-espace des fonctions $f \in C_0$ telles que Δf (au sens des distributions) est dans C_0, et on a alors $Lf = \frac{1}{2} \Delta f$. Cette discussion montre que même sur des exemples simples, le domaine $\mathcal{D}(L)$ est souvent difficile à exhiber précisément. Nous n'aurons cependant pas besoin de le faire dans le cadre des applications que nous avons en vue.

*. Opérateur linéaire dont le graphe est fermé.

La proposition suivante permet d'identifier des éléments de $\mathcal{D}(L)$ ainsi que leur image à l'aide de martingales définies pour le processus de Markov associé.

Proposition 6. *Soit X un processus de Feller par rapport à $(\mathscr{F}_t)$ de semi-groupe (P_t) et à trajectoires continues***. Pour toutes fonctions $f, g \in C_0(\mathbb{R}^d)$, les deux conditions suivantes sont équivalentes :*

(i) On a $f \in \mathcal{D}(L)$ et $Lf = g$.

(ii) Pour tout $x \in \mathbb{R}^d$, le processus

$$M_t^f \overset{\text{déf.}}{=} f(X_t) - f(X_0) - \int_0^t Lf(X_s)\, ds$$

est une $(\mathscr{F}_t, \mathbb{P}_x)$-martingale.

Remarque. Si, l'une de ces conditions est vérifiée, le processus M_t^f est en fait une martingale sous $\mathbb{P}_\mu$ pour toute loi initiale μ.

Cette proposition fournit une nouvelle manière de déterminer le générateur infinitésimal en utilisant par exemple les outils fournis par le calcul stochastique.

Exemple. D'après la formule d'Itô, pour toute $f \in C_0^2$,

$$f(B_t) - f(B_0) - \frac{1}{2}\int_0^t \Delta f(B_s)\, ds = \int_0^t \nabla f(B_s) \cdot dB_s$$

est une martingale locale bornée sur tout segment $[0\,; T]$, donc une vraie martingale. On a ainsi retrouvé le fait que $C_0^2 \subset \mathcal{D}(L)$ et que $L = \frac{1}{2}\Delta$ sur C_0^2.

3.1.2 Mesures invariantes et réversibles

Considérons un processus de Markov (X_t) de semi-groupe de transition (P_t).

Définition 11. Une mesure de probabilité borélienne μ sur $\mathbb{R}^d$ est dite *invariante* pour le semi-groupe (P_t) si, pour tout $t \geqslant 0$ et pour toute fonction f borélienne bornée, on a

$$\int P_t f\, d\mu = \int f\, d\mu.$$

Remarque. Supposons que le semi-groupe soit de Feller. Si μ est une mesure invariante pour le semi-groupe (P_t), on voit en dérivant à l'aide du théorème de convergence dominée que pour toute fonction $f \in \mathcal{D}(L)$,

$$\int Lf\, d\mu = \lim_{t \to 0} \int \frac{P_t f - f}{t}\, d\mu = 0.$$

$*$. De manière à ce que les intégrales qui apparaissent aient un sens. En fait, il suffirait de travailler avec une version càdlàg de (X_t) (voir [RY99]).

En particulier, on en déduit si X_0 suit la loi μ, alors $f(X_t)$ est une $\mathbb{P}_\mu$-martingale. Réciproquement, il est facile de voir que si on a

$$\int Lf\, d\mu = 0,$$

pour tout f appartenant à un sous-ensemble de $\mathcal{D}(A)$ dense dans C_0, alors pour tout f dans ce sous-ensemble on a

$$\forall t \geqslant 0, \quad \int P_t f\, d\mu = \int f\, d\mu,$$

et que ceci s'étend à tout $f \in C_0$.

Remarque. Supposons que μ soit une mesure de probabilité invariante pour le semi-groupe (P_t) et remarquons par ailleurs que pour toute fonction f borélienne bornée et pour tout $p \in [1\,;\infty[$, l'inégalité de Jensen indique que pour tout $x \in \mathbb{R}^d$,

$$|P_t f(x)|^p = |\mathbb{E}_x[f(X_t)]|^2 \leqslant \mathbb{E}_x[|f(X_t)|^p] = P_t|f|^p(x).$$

En utilisant l'invariance de μ on voit alors que

$$\|P_t f\|_{L^p(\mu)} = \int |P_t f|^p\, d\mu \leqslant \int P_t|f|^p\, d\mu = \int |f|^p\, d\mu = \|f\|_{L^p(\mu)}\,,$$

et cette inégalité reste vraie pour $p = \infty$. Dès lors, (P_t) se prolonge de manière unique en un semi-groupe d'opérateurs de contraction de $L^p(\mu)$ dans lui-même. Par ailleurs, en prenant $f = \mathbf{1}$, on voit que l'opérateur induit sur $L^p(\mu)$ est de norme égale à 1.

Pour ce semi-groupe sur l'espace de Banach $L^p(\mu)$, on peut définir un générateur A et un domaine $\mathcal{D}(A)$: une fonction $f \in L^p(\mu)$ est dans $\mathcal{D}(A)$ si et seulement si la limite

$$Af \overset{\text{déf.}}{=} \lim_{t \to 0} \frac{1}{t}(P_t f - f)$$

existe dans $L^p(\mu)$. Lorsque le semi-groupe est de Feller, on a bien sûr $\mathcal{D}(L) \subset \mathcal{D}(A)$ et $Af = Lf$ pour tout $f \in \mathcal{D}(L)$. En particulier, $\mathcal{D}(A)$ est dense dans $L^p(\mu)$ puisque $\mathcal{D}(L)$ est dense dans C_0 qui est lui-même dense dans $L^p(\mu)$. De plus, on voit que $t \mapsto P_t f \in L^p(\mu)$ est continue et ceci entraîne notamment que A est un opérateur fermé. Pour plus de détails sur ces notions, on pourra consulter [Roy99].

Voici maintenant une condition qui entraîne l'invariance et qui a son intérêt propre.

Définition 12. Une mesure borélienne de probabilité μ sur $\mathbb{R}^d$ est *symétrique* par rapport au semi-groupe (P_t) si pour toutes fonctions f, g boréliennes bornées et pour tout $t \geqslant 0$,

$$\int g\, P_t f\, d\mu = \int f\, P_t g\, d\mu.$$

Remarque. Si μ est symétrique, alors elle est invariante. En effet, il suffit de prendre $g = \mathbf{1}$ dans l'équation précédente.

Remarque. Dans le cadre L^p introduit plus haut, cette propriété de symétrie se prolonge et indique que les opérateur P_t induits sont auto-adjoints dans l'espace de Hilbert $L^2(\mu)$. En dérivant dans $L^2(\mu)$, on voit par ailleurs que ceci revient à dire que le générateur A est auto-adjoint sur son domaine $\mathcal{D}(A)$.

Les notions de mesure invariante et de mesure symétrique se révèlent être de précieux outils pour l'étude de processus de Markov. Un autre point de vue est possible, et c'est celui que nous adoptons ici : on étudie le semi-groupe de transition ou le processus associé, et on en déduit des propriétés pour la mesure invariante.

3.2 Processus et semi-groupe d'Ornstein-Uhlenbeck

Le *semi-groupe d'Ornstein-Uhlenbeck* est l'outil central de la démonstration de l'inégalité isopérimétrique pour la mesure gaussienne de M. Ledoux et D. Bakry [BL96]. On peut le définir comme la famille d'opérateurs $(P_t)_{t\geqslant 0}$ agissant sur les fonctions boréliennes bornées de $\mathbb{R}^d$ dans $\mathbb{R}$ par

$$P_t f(x) = \int_{\mathbb{R}^d} f\Big(\mathrm{e}^{-t/2}x + (1-\mathrm{e}^{-t})^{1/2}y\Big)\,\gamma_d(\mathrm{d}y),$$

où on rappelle que $\gamma_d = \gamma_1^{\otimes d}$ désigne la mesure gaussienne canonique sur $\mathbb{R}^d$.

Nous donnons dans cette section une présentation peut-être plus « naturelle » de ce semi-groupe conformément aux notions que nous venons d'introduire. À partir de là, nous étudions quelques unes de ses propriétés principales qui nous servirons à établir l'inégalité isopérimétrique pour la mesure gaussienne.

3.2.1 Nature markovienne du semi-groupe d'Ornstein-Uhlenbeck

On considère l'équation différentielle stochastique dite de *Langevin* (qui correspond intuitivement à un mouvement brownien multidimensionnel pour lequel on a ajouté sur chaque coordonnée une force de rappel proportionnelle à sa distance à l'origine),

$$\mathrm{d}U_t = \mathrm{d}B_t - \frac{1}{2}U_t\mathrm{d}t,$$

où $(B_t)_{t\geqslant 0} = (B_t^1, \ldots, B_t^d)_{t\geqslant 0}$ est un mouvement brownien standard dans $\mathbb{R}^d$. Les coefficients de cette équation différentielle stochastique étant lipschitziens, les théorèmes généraux * nous permettent d'affirmer l'existence et l'unicité trajectorielle d'une solution (U_t) définie pour tout $t \in [0\,;\infty[$, processus de Markov (même fortement markovien) à trajectoires continues.

*. Consulter par exemple [RY99].

Définition 13. Le *processus d'Ornstein-Uhlenbeck* est l'unique processus de Markov (U_t) solution de l'équation différentielle stochastique $dU_t = dB_t - \frac{1}{2}U_t dt$.

En fait, on peut même obtenir une expression explicite de ce processus à l'aide d'une intégrale stochastique : sous la condition initiale $U_0 = x$, on a

$$U_t = \mathrm{e}^{-t/2}\left[x + \int_0^t \mathrm{e}^{s/2}\,\mathrm{d}B_s\right].$$

La vérification de ceci est un exercice facile à partir de la formule d'Itô (ou d'intégration par parties ici). On écrit que

$$U_t - U_0 = \int_0^t \mathrm{e}^{-s/2}\mathrm{e}^{s/2}\,\mathrm{d}B_s + \int_0^t \left(-\frac{1}{2}\right) U_s\,\mathrm{d}s,$$

le terme d'Itô s'annulant car $\mathrm{e}^{-t/2}$ est à variations finies. Dans la suite, on notera (U_t^x) la solution trajectorielle issue de $x \in \mathbb{R}^d$. Le terme $\mathrm{e}^{-t/2}\int_0^t \mathrm{e}^{s/2}\mathrm{d}B_s = U_t^0$ suit une loi normale d-dimensionnelle de moyenne nulle et de matrice de covariance donnée par

$$\mathbb{E}\left[\left(\mathrm{e}^{-t/2}\int_0^t \mathrm{e}^{s/2}\,\mathrm{d}B_s^i\right)\left(\mathrm{e}^{-t/2}\int_0^t \mathrm{e}^{s/2}\,\mathrm{d}B_s^j\right)\right] = \mathbb{E}\left[\mathrm{e}^{-t}\int_0^t \mathrm{e}^{s}\,\mathrm{d}\langle B^i, B^j\rangle_s\right],$$

c'est à dire $(1-\mathrm{e}^{-t})I_d$. On voit donc directement que pour toute fonction $f : \mathbb{R}^d \to \mathbb{R}$ borélienne et bornée, on a

$$\mathbb{E}_x\left[f(U_t)\right] = \mathbb{E}\left[f(U_t^x)\right] = P_t f(x),$$

où (P_t) est le semi-groupe d'Ornstein-Uhlenbeck tel que défini précédemment.

Proposition 7. *Le processus d'Ornstein-Uhlenbeck est un processus de Feller. De plus, la loi de U_t converge étroitement vers γ_d quelle que soit la loi initiale de U_0.*

Démonstration. Soit $f \in C_0(\mathbb{R}^d)$. Une telle fonction étant bornée, le théorème de convergence dominée de Lebesgue permet d'affirmer directement que $P_t f$ est continue et que $\lim_{|x|\to\infty} P_t f(x) = 0$. De même, on vérifie que pour toute fonction $f \in C_b(\mathbb{R}^d)$, et pour toute loi initiale $\mu = \mathscr{L}(U_0)$, on a $\lim_{t\to\infty} \mathbb{E}[f(U_t)] = \mathbb{E}[f(\mathcal{N}(0, I_d))]$. $\square$

Remarquons qu'on aurait pu être encore un peu plus précis dans la description de la loi de (U_t^0). Comme ces accroissements sont indépendants et suivent des lois gaussiennes, c'est un processus gaussien continu dont on a déjà calculé la fonction de convariance. Il a donc la même loi qu'un mouvement brownien changé de temps $(\beta_{1-\mathrm{e}^{-t}})_{t\geqslant 0}$, où β est un mouvement brownien standard d-dimensionnel.

3.2.2 Générateur infinitésimal et mesure invariante

Nous déterminons maintenant le générateur infinitésimal L du processus (respectivement semi-groupe) d'Ornstein-Uhlenbeck sur un sous-espace dense de C_0 au moyen des méthodes déjà mises en avant.

Proposition 8. *Pour toute fonction $f \in C_c^2(\mathbb{R}^d)$ *, on a $f \in \mathcal{D}(L)$ et*

$$Lf(x) = \frac{1}{2}[\Delta f(x) - x \cdot \nabla f(x)], \qquad x \in \mathbb{R}^d.$$

Démonstration. On applique la formule d'Itô au processus d'Ornstein-Uhlenbeck :

$$f(U_t) - f(U_0) + \frac{1}{2}\int_0^t U_s \cdot \nabla f(U_s)ds - \frac{1}{2}\int_0^t \Delta f(U_s)ds = \int_0^t \nabla f(U_s) \cdot dBs,$$

ce qui montre que le processus défini par

$$M_t^f \overset{\text{déf.}}{=} f(U_t) - f(U_0) - \int_0^t Lf(U_s)ds = \int_0^t \nabla f(U_s) \cdot dB_s$$

est une martingale (c'est une martingale locale dont le crochet est borné sur tout intervalle de temps compact) relativement à la filtration brownienne, et on conclut à l'aide de la proposition 6. □

L'importance accordée au semi-groupe et au processus d'Ornstein-Uhlenbeck dans ce chapitre provient essentiellement de la proposition suivante, qui montre qu'ils admettent la mesure gaussienne canonique comme mesure symétrique et invariante.

Proposition 9. *La mesure gaussienne $\gamma = \gamma_d$ est symétrique (donc invariante) par rapport au semi-groupe d'Ornstein-Uhlenbeck.*

Démonstration. Soient $f, g\colon \mathbb{R}^d \to \mathbb{R}$ deux fonctions boréliennes bornées et $t \geqslant 0$ fixé. Par définition, on a

$$\int fP_tg\,d\gamma = \mathbb{E}[f(X)g(c_tX + s_tY)]$$

où X et Y sont deux variables aléatoires indépendantes de loi γ et $c_t = e^{-t/2}$, $s_t = (1-e^{-t})^{1/2}$ sont deux réels positifs tels que ${c_t}^2 + {s_t}^2 = 1$. Les couples $(X, c_tX + s_tY)$ et $(c_tX + s_tY, X)$ sont deux vecteurs gaussiens dont les matrices de covariance coïncident et valent

$$\begin{pmatrix} 1 & c_t \\ c_t & 1 \end{pmatrix}.$$

Ils ont donc la même loi et il vient en particulier

$$\int fP_tg\,d\gamma = \mathbb{E}[f(c_tX + s_tY)g(X)] = \int gP_tf\,d\gamma.$$

□

*. On note $C_c^2(\mathbb{R}^d)$ l'espace des fonction de $\mathbb{R}^d$ dans $\mathbb{R}$ de classe C^2 et à support compact.

Remarque. La convergence étroite de $\mathscr{L}(U_t)$ vers γ_d permettait aussi d'obtenir directement l'invariance pour les fonctions continues bornées, mais pas la propriété de symétrie.

Notons enfin l'importante propriété d'intégration par parties vérifiée par le générateur infinitésimal du processus d'Ornstein-Uhlenbeck.

Proposition 10. *Pour toutes fonctions $f, g \in C_c^2$, on a*

$$\int fLg\, d\gamma = -\frac{1}{2} \int \nabla f \cdot \nabla g\, d\gamma.$$

Démonstration. Il suffit de faire le calcul. Par définition,

$$2 \int fLg\, d\gamma = \int f(x) \Delta g(x)\, d\gamma(x) - \int f(x) \left[x \cdot \nabla g(x)\right] d\gamma(x).$$

On fait une intégration par parties sur le premier terme :

$$\frac{1}{(2\pi)^{d/2}} \sum_{i=1}^{d} \int \frac{\partial^2 g}{\partial x_i^2}(x) f(x) e^{-|x|^2/2} dx = -\int \nabla f \cdot \nabla g\, d\gamma + \int f(x) \left[x \cdot \nabla g(x)\right] d\gamma,$$

ce qui donne le résultat. On aurait aussi pu remarquer que

$$L(fg) - fLg - gLf = \nabla f \cdot \nabla g$$

et utiliser la symétrie de la mesure μ. $\square$

Remarque. Comme on l'a déjà vu dans un cas plus général, la symétrie de la mesure γ permet de donner un sens L^2 au semi-groupe (P_t). Les opérateurs (P_t) ainsi définis forment un semi-groupe d'opérateurs de contraction auto-adjoints sur $L^2(\mu)$ dont le générateur A est auto-adjoint sur son domaine $\mathcal{D}(A)$. On peut montrer que le domaine $\mathcal{D}(A)$ est exactement formé des $f \in L^2(\mu)$ qui sont dans l'espace de Sobolev $W^{2,2}_{\mathrm{loc}}$ et tels que $Lf \in L^2(\mu)$ au sens des distributions ([Roy99], théorème 2.2.27). Nous n'aurons cependant pas besoin d'un résultat aussi précis.

Plus simplement, considérons l'algèbre $\mathcal{A}$ des fonctions $f \colon \mathbb{R}^d \to \mathbb{R}$ de classe C^∞ et telles que f et toutes ses dérivées soient *à croissance lente*, c'est à dire bornées par des polynômes. Pour une telle fonction, on vérifie directement sur la définition de P_t que $P_t f \in \mathcal{A}$ et qu'on a encore la convergence de $\frac{1}{t}(P_t f - f)$ vers $Lf \in \mathcal{A}$ (voir par exemple [An00]). De même, pour tous $f, g \in \mathcal{A}$ on a

$$\int fLg\, d\gamma = \int gLf\, d\gamma,$$

et a fortiori la formule d'intégration par parties est également vérifiée. Notons enfin que cette classe de fonctions est dense dans chaque $L^p(\mu)$ $(1 < p < \infty)$ et contient les constantes.

3.3 Théorème isopérimétrique gaussien

Les outils que l'on vient de mettre en place pour le processus d'Ornstein-Uhlenbeck permettent d'établir de nombreux résultats sur la mesure invariante γ_d au moyen de calculs portant uniquement sur le semi-groupe P_t et son générateur L. On peut notamment utiliser cette méthode pour donner une nouvelle démonstration de l'inégalité isopérimétrique gaussienne.

Théorème 5. *Pour toute fonction* $f\colon \mathbb{R}^d \to [0\,;1]$ *lipschitzienne,*

$$\mathcal{I}\left(\int f\,d\gamma_d\right) \leqslant \int \sqrt{\mathcal{I}^2(f) + |\nabla f|^2}d\gamma_d.$$

Démonstration. Comme on l'a déjà vu dans le premier chapitre, on peut supposer que f est régulière, par exemple dans C_c^∞ qui est stable par P_t et par L, ou encore dans l'algèbre $\mathcal{A}$. On peut par ailleurs se contenter de prouver le théorème dans le cas $d = 1$ * en vertu du lemme 6.

Pour commencer, nous montrons que la fonction

$$J_t = \int \sqrt{\mathcal{I}^2(P_t f) + (P_t f)'^2}d\gamma$$

est décroissante en t sur $[0\,;\infty[$. Par commodité de notation, nous poserons $f_t \overset{\text{déf.}}{=} P_t f$ pour tout $t \geqslant 0$, qui vérifie en particulier $df_t/dt = Lf_t$. En dérivant (on peut toujours supposer que $0 < \eta \leqslant f \leqslant 1-\eta$ quitte à faire tendre η vers 0 plus tard), on obtient

$$\frac{dJ}{dt} = \int \frac{\mathcal{I}\mathcal{I}'(f_t)Lf_t + f_t'(Lf_t)'}{\sqrt{\mathcal{I}^2(f_t) + f_t'^2}}d\gamma = \int \frac{\mathcal{I}\mathcal{I}'(f_t)Lf_t + f_t'(Lf_t)'}{\sqrt{K_t}}d\gamma,$$

où on a posé $K_t \overset{\text{déf.}}{=} \mathcal{I}^2(f_t) + f_t'^2$.

De plus, la formule d'intégration par parties montre que

$$\int \frac{\mathcal{I}\mathcal{I}'(f_t)Lf_t}{\sqrt{K_t}}d\gamma = -\frac{1}{2}\int \left(\frac{\mathcal{I}\mathcal{I}'(f_t)}{\sqrt{K_t}}\right)' f_t' d\gamma$$

et comme

$$\left(\frac{\mathcal{I}\mathcal{I}'(f_t)}{\sqrt{K_t}}\right)' = \frac{(\mathcal{I}'^2 + \mathcal{I}\mathcal{I}'')(f_t)f_t'}{K_t^{1/2}} - \frac{1}{2}\frac{\mathcal{I}\mathcal{I}'(f_t)(2\mathcal{I}\mathcal{I}'(f_t)f_t' + 2f_t'f_t'')}{K_t^{3/2}},$$

on obtient, en utilisant l'équation différentielle $\mathcal{I}\mathcal{I}'' = -1$,

$$\int \frac{\mathcal{I}\mathcal{I}'(f_t)Lf_t}{\sqrt{K_t}}d\gamma = -\frac{1}{2}\int \frac{f_t'^2}{K_t^{1/2}}[\mathcal{I}'^2(f_t) - 1]d\gamma + \frac{1}{2}\int \frac{\mathcal{I}\mathcal{I}'(f_t)f_t'^2}{K_t^{3/2}}[\mathcal{I}\mathcal{I}'(f_t) + f_t'']d\gamma.$$

*. En fait, on fera les calculs dans un cas plus général et multidimensionnel au prochain chapitre.

Pour l'autre terme, on remarque que $f'(Lf)' = -\frac{1}{2}f'^2 + f'Lf'$, et on applique à nouveau la formule d'intégration par parties :

$$\begin{aligned}\int \frac{f_t' L f_t'}{\sqrt{K_t}} &= -\frac{1}{2}\int \left[\frac{f_t''}{K_t^{1/2}} - \frac{f_t'(\mathcal{I}\mathcal{I}'(f_t)f_t' + f_t' f_t'')}{K_t^{3/2}}\right] f_t'' d\gamma \\ &= -\frac{1}{2}\int \frac{f_t''^2}{K_t^{1/2}} d\gamma + \frac{1}{2}\int \frac{f_t' f_t''}{K_t^{3/2}}[\mathcal{I}\mathcal{I}'(f_t)f_t' + f_t' f_t'']d\gamma.\end{aligned}$$

Finalement, on a donc

$$\frac{dJ}{dt} = -\frac{1}{2}\int \frac{1}{K_t^3/2}\left[\mathcal{I}'^2(f_t)f_t'^4 - 2\mathcal{I}\mathcal{I}'(f_t)f_t'^2 f_t'' + \mathcal{I}^2(f_t)f_t''^2\right] d\gamma,$$

et on conclut à l'aide de la factorisation

$$\mathcal{I}'^2 f'^4 - 2\mathcal{I}\mathcal{I}' f'^2 f'' + \mathcal{I}^2 f''^2 = \left(\mathcal{I}' f'^2 - \mathcal{I} f''\right)^2$$

qui montre que la dérivée est négative pour tout $t \geqslant 0$, c'est à dire que $t \mapsto J_t$ est décroissante. En particulier, on obtient que pour tout $t \geqslant 0$,

$$\int \sqrt{\mathcal{I}^2(f) + |f'|^2}\, d\gamma = J_0 \geqslant \int \sqrt{\mathcal{I}^2(f_t) + |f_t'|^2}\, d\gamma.$$

Il reste alors à voir que

$$J_t \underset{t\to\infty}{\longrightarrow} \mathcal{I}\left(\int f d\gamma\right).$$

Or on a déjà vu (du fait de la convergence étroite) que pour tout $x \in \mathbb{R}^d$,

$$\lim_{t\to\infty} P_t f(x) = \int f\, d\gamma,$$

et on vérifie sans peine à l'aide du théorème de dérivation des intégrales à paramètre que

$$(P_t f)'(x) = e^{-t/2} P_t f'(x) \underset{t\to\infty}{\longrightarrow} 0.$$

Comme l'intégrande est majoré par $\sqrt{1 + |f'|_\infty^2}$, le théorème de convergence dominée donne la limite voulue.

$\square$

Chapitre 4

Généralisation aux mesures de Boltzmann

La méthode de semi-groupe que nous avons utilisée dans le chapitre précédent afin d'établir l'inégalité isopérimétrique pour la mesure gaussienne se généralise naturellement à une classe importante de mesures de probabilité, dites de *Boltzmann*. Ces mesures sur $\mathbb{R}^d$ sont absolument continues par rapport à la mesure de Lebesgue et ont une densité de la forme $Z^{-1}e^{-W(x)}$. On va montrer que sous certaines hypothèses sur W, elles vérifient également une inégalité isopérimétrique gaussienne.

4.1 Processus de Kolmogorov et mesure de Boltzmann

Soit $W\colon \mathbb{R}^d \to \mathbb{R}_+$ une fonction de classe C^2 positive telle que

$$Z \overset{\text{déf.}}{=} \int_{\mathbb{R}^d} e^{-W(x)}\, dx < \infty.$$

Définition 14. La *mesure de Boltzmann* de potentiel W, notée μ_W (ou μ s'il n'y a pas de confusion possible) est la mesure de probabilité sur $\mathbb{R}^d$ de densité $d\mu_W(x)/dx = Z^{-1}e^{-W(x)}$ par rapport à la mesure de Lebesgue.

Remarque. Bien sûr, la mesure gaussienne est un cas particulier de mesure de Boltzmann pour le choix de potentiel $W(x) = \frac{1}{2}|x|^2$.

Nous introduisons maintenant un processus qui joue pour la mesure de Boltzmann un rôle analogue au processus d'Ornstein-Uhlenbeck pour la mesure gaussienne. Considérons pour cela une équation différentielle stochastique de Langevin modifiée :

$$dX_t = dB_t - \frac{1}{2}\nabla W(X_t)dt.$$

Sous certaines conditions (par exemple si ∇W est lipshitzien) * qu'on supposera remplies, les théorèmes généraux assurent l'existence et l'unicité trajectorielle d'une solution (X_t) appelée processus de Kolmogorov et définie pour tout $t \in [0\,;\infty[$ et à trajectoires continues quelle que soit la condition initiale. De plus, ce processus est un processus de Markov (voire même de Feller pour la condition ∇W lipschitzien). Comme précédemment, on note X_t^x pour $x \in \mathbb{R}^d$ la solution issue de $X_0 = x$.

Soit (P_t) le semi-groupe du processus de Kolmogorov défini pour toute fonction f mesurable bornée et $t \geqslant 0$ par

$$P_t f(x) \overset{\text{déf.}}{=} \mathbb{E}_x[f(X_t)] = \mathbb{E}[f(X_t^x)].$$

Contrairement au cas $W(x) = |x|^2/2$, on est généralement incapable d'expliciter le processus (X_t). En revanche, on peut déterminer la loi de X_t (et même un peu plus) à l'aide d'une transformation de Girsanov. On renvoie à [Roy99], lemme 2.2.21 pour le résultat suivant.

Lemme 7. *Pour tout $t \geqslant 0$, la loi de $\{X_s^x; 0 \leqslant s \leqslant t\}$ sur $C([0\,;t], \mathbb{R})$ est absolument continue par rapport à la loi d'un mouvement brownien (B_t) issu de x, avec densité*

$$\begin{aligned} D_t &= \exp\left\{-\frac{1}{2}\int_0^t \nabla W(B_s)\cdot dB_s - \frac{1}{8}\int_0^t |\nabla(B_s)|^2\, ds\right\} \\ &= \exp\left\{\frac{1}{2}\left[W(x) - W(B_t)\right] + \frac{1}{4}\int_0^t \Delta W(B_s)\, ds - \frac{1}{8}\int_0^t |\nabla W(B_s)|^2\, ds\right\}. \end{aligned}$$

En utilisant le fait que la loi d'un mouvement brownien est invariante par retournement du temps, on en déduit la symétrie de la mesure de Boltzmann.

Corollaire 2. *La mesure de Boltzmann μ est symétrique (donc invariante) par rapport au semi-groupe de Kolmogorov (P_t).*

Remarque. On peut donc comme on l'a déjà vu donner un sens $L^2(\mu)$ au semi-groupe.

Calculons maintenant le générateur infinitésimal L du semi-groupe de Kolmogorov.

Proposition 11. *Toute fonction $f \in C_c^2$ est dans le domaine du générateur L et pour tout $x \in \mathbb{R}^d$,*

$$Lf(x) = \frac{1}{2}(\Delta f(x) - \nabla W(x)\cdot\nabla f(x)).$$

Démonstration. Comme pour le processus d'Ornstein-Uhlenbeck, on applique la formule d'Itô à la semi-martingale X_t :

$$f(X_t) - f(X_0) = \int_0^t \nabla f(X_s)\cdot dB_s - \frac{1}{2}\int_0^t \nabla W(X_s)\cdot\nabla f(X_s) ds + \frac{1}{2}\int_0^t \Delta f(X_s) ds.$$

*. Deux autres conditions suffisantes moins fortes sont données dans [Roy99] :

1. $W(x) \to +\infty$ quand $|x| \to +\infty$ et $|\nabla W|^2 - \frac{1}{2}\Delta W$ minoré par une constante.
2. Il existe $a, b \in \mathbb{R}$ tels que $\forall x,\ x\cdot\nabla W(x) \geqslant -a|x|^2 - b$.

On voit donc qu'avec l'opérateur L proposé, le processus

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds = \int_0^t \nabla f(X_s) \cdot dB_s$$

est une martingale locale de crochet

$$\langle M^f \rangle_t = \int_0^t |\nabla f(X_s)|^2 ds$$

borné sur tout intervalle de temps compact, ce qui montre que (M_t^f) est une vraie martingale et achève de prouver la proposition. $\square$

Par analogie au cas du processus d'Ornstein-Uhlenbeck, on fait une hypothèse qui permet d'assurer la validité des calculs à venir.

Définition 15 (Algèbre standard)**.** On se place sous l'hypothèse de l'existence d'une algèbre $\mathcal{A}$ de fonctions C^∞ contenant $C_c^\infty(\mathbb{R}^d)$ et telle que
- $\mathcal{A}$ est inclue et dense dans tous les $L^p(\mu)$ pour $1 < p < \infty$;
- $\mathcal{A}$ est stable par P_t et par L ;
- $\mathcal{A}$ est stable par composition avec les fonctions C^∞ et contient les constantes.

On introduit maintenant sur cette algèbre $\mathcal{A}$ deux formes bilinéaires symétriques qui jouent un rôle important :

$$\Gamma(f,g) \overset{\text{déf.}}{=} L(fg) - fLg - gLf,$$
$$\Gamma_2(f,g) \overset{\text{déf.}}{=} L\Gamma(f,g) - \Gamma(f,Lg) - \Gamma(g,Lf).$$

Pour plus de commodité, on notera $\Gamma(f)$ et $\Gamma_2(f)$ en lieu et place de $\Gamma(f,f)$ et $\Gamma_2(f,f)$.

Lemme 8. *Pour toutes fonctions $f, g \in \mathcal{A}$, on a*

$$\Gamma(f,g) = \nabla f \cdot \nabla g \quad \textit{et} \quad \Gamma_2(f) = \|f''\|_2^2 + W''(\nabla f, \nabla f)$$

où f'' et W'' désignent respectivement les matrices hessiennes de f et W assimilées à deux formes bilinéaires symétriques, et où $\|f''\|_2$ est la norme de Hilbert-Schmidt de f'',

$$\|f''\|_2^2 = \sum_{i,j=1}^d \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]^2.$$

Démonstration. Par polarisation, il suffit de vérifier ces formules sur la diagonale, c'est-à-dire à partir de

$$\Gamma(f) = Lf^2 - 2fLf \quad \text{et} \quad \Gamma_2(f) = L\Gamma(f) - 2\Gamma(f,Lf).$$

On a $Lf = \frac{1}{2}[\Delta f - \nabla W \cdot \nabla f]$, donc puisque $\Delta f^2 = 2[|\nabla f|^2 + f\Delta f]$ et $\nabla W \cdot \nabla f^2 = 2f\,\nabla W \cdot \nabla f$, il vient

$$Lf^2 = |\nabla f|^2 + f[\Delta f - \nabla W \cdot \nabla f],$$

d'où finalement

$$\Gamma(f) = |\nabla f|^2.$$

Il reste à calculer l'opérateur Γ_2. On a $\Delta|\nabla f|^2 = 2\,\|f''\|_2^2 + 2[\nabla f \cdot \nabla \Delta f]$ et $\nabla W \cdot \nabla |\nabla f|^2 = 2f''(\nabla W, \nabla f)$, d'où

$$L\Gamma(f) = \|f''\|_2^2 + \nabla f \cdot \nabla \Delta f - f''(\nabla W, \nabla f).$$

Pour le second terme,

$$\begin{gathered}\Gamma(f, \Delta f) = \nabla f \cdot \nabla \Delta f \quad \text{et} \\ \Gamma(f, \nabla W \cdot \nabla f) = \nabla f \cdot \nabla(\nabla W \cdot \nabla f) = W''(\nabla f, \nabla f) + f''(\nabla W, \nabla f), \\ \Gamma(f, Lf) = \frac{1}{2}\left[\nabla f \cdot \nabla \Delta f - W''(\nabla f, \nabla f) - f''(\nabla W, \nabla f)\right],\end{gathered}$$

d'où finalement

$$\Gamma_2(f) = \|f''\|_2^2 + W''(\nabla f, \nabla f).$$

□

Ces calculs permettent notamment de déduire immédiatement la formule d'intégration par parties pour μ.

Proposition 12. *Pour toutes fonctions $f, g \in \mathcal{A}$, on a*

$$\int f Lg\, d\mu = -\frac{1}{2}\int \nabla f \cdot \nabla g\, d\gamma.$$

Démonstration. Ceci découle directement de la définition de Γ et du lemme précédent en utilisant la symétrie de la mesure de Boltzmann μ. □

4.2 Inégalités fonctionnelles en courbure positive

À partir de maintenant, on se place sous l'hypothèse de courbure positive $W'' \geqslant 0$, c'est-à-dire

$$W''(x, x) \geqslant 0, \qquad \forall x \in \mathbb{R}^d.$$

Le semi-groupe (P_t) vérifie alors une très importante propriété de sous-commutativité.

Lemme 9. *Pour toute fonction $f \in A$ et pour tout $t \geqslant 0$, on a $|\nabla P_t f| \leqslant P_t |\nabla f|$.*

Démonstration. On réalise une interpolation entre 0 et t de la manière suivante

$$P_t|\nabla f| - |\nabla P_t f| = P_t\sqrt{\Gamma(f)} - \sqrt{\Gamma(P_t f)} = \int_0^t \frac{d}{ds} P_s\sqrt{\Gamma(P_{t-s}f)}\, ds.$$

Or la dérivée vaut

$$\begin{aligned}\frac{d}{ds}P_s\sqrt{\Gamma(P_{t-s}f)} &= P_sL\sqrt{\Gamma(P_{t-s}f)} + P_s\frac{1}{2\sqrt{\Gamma(P_{t-s}f)}}\frac{d}{ds}\Gamma(P_{t-s}f)\\ &= P_s\left(L\sqrt{\Gamma(P_{t-s}f)} - \frac{\Gamma(P_{t-s}f, LP_{t-s})}{\sqrt{\Gamma(P_{t-s}f)}}\right).\end{aligned}$$

Pour calculer le terme de gauche, on peut s'appuyer sur les calculs précédents. On a

$$\nabla W\cdot\nabla\sqrt{\Gamma(f)} = \frac{1}{2\sqrt{\Gamma(f)}}\nabla W\cdot\nabla\Gamma(f)$$

et de même

$$\Delta\sqrt{\Gamma(f)} = \frac{1}{2\sqrt{\Gamma(f)}}\left[\Delta\Gamma(f) - \frac{1}{2\Gamma(f)}|\nabla\Gamma(f)|^2\right],$$

ce qui permet d'obtenir

$$L\sqrt{\Gamma(f)} = \frac{1}{2\sqrt{\Gamma(f)}}\left[L\Gamma(f) - \frac{1}{4\Gamma(f)}|\nabla\Gamma(f)|^2\right].$$

La dérivée initiale vaut donc

$$P_s\left(\frac{1}{2\sqrt{\Gamma(P_{t-s}f)}}\left[\Gamma_2(P_{t-s}f) - \frac{|\nabla\Gamma(P_{t-s}f)|^2}{4\Gamma(P_{t-s}f)}\right]\right) \geqslant 0.$$

En effet, d'après l'inégalité de Cauchy-Schwarz, on a

$$|\nabla\Gamma(f)|^2 = \sum_{i=1}^d\left[\sum_{j=1}^d 2\frac{\partial^2 f}{\partial x_i\partial x_j}\frac{\partial f}{\partial x_i}\right]^2 \leqslant 4\sum_{j=1}^d\left[\frac{\partial f}{\partial x_i}\right]^2\sum_{i,j=1}^d\left[\frac{\partial^2 f}{\partial x_i\partial x_j}\right]^2 = 4\Gamma(f)\,\|f''\|_2^2\,,$$

et on a montré que $\Gamma_2(f) = \|f''\|_2^2 + W''(\nabla f, \nabla f)$. □

Remarque. On peut montrer dans un cadre beaucoup plus général que cette propriété est en fait caractéristique d'un certain critère de courbure dû à Barkry et Emery (voir par exemple [An00]).

En utilisant l'inégalité de Cauchy-Schwarz, on en déduit immédiatement une propriété de sous-commutativité plus faible.

Corollaire 3. *Pour toute fonction $f\in\mathcal{A}$ et pour tout $t\geqslant 0$, on a $|\nabla P_t f|^2 \leqslant P_t|\nabla f|^2$.*

Nous sommes désormais en mesure de prouver qu'en courbure positive, le semi-groupe de Kolmogorov P_t fournit, uniformément en x, des inégalités de Poincaré et des inégalités de Sobolev logarithmiques pour tout $t\geqslant 0$.

Proposition 13 (Inégalités de Poincaré locales)**.** *Pour toute fonction $f \in \mathcal{A}$ et pour tout $t \geqslant 0$,*

$$t\,|\nabla P_t f|^2 \leqslant P_t(f^2) - (P_t f)^2 \leqslant t\,P_t|\nabla f|^2.$$

Démonstration. On fait ici encore une interpolation en 0 et t :

$$P_t(f^2) - (P_t f)^2 = \int_0^t \frac{d}{ds} P_s (P_{t-s} f)^2\,ds.$$

En dérivant, on voit alors que

$$\frac{d}{ds} P_s (P_{t-s} f)^2 = P_s L (P_{t-s} f)^2 - 2 P_s \left[(P_{t-s} f) \times (L P_{t-s} f) \right] = P_s(\Gamma(P_{t-s} f)).$$

En utilisant la propriété de sous-commutativité faible dans les deux sens et la propriété de semi-groupe de P_t, on voit alors que

$$\int_0^t |\nabla P_t f|^2\,ds \leqslant P_t(f^2) - (P_t f)^2 \leqslant \int_0^t P_t |\nabla f|^2\,ds$$

où les deux intégrandes ne dépendent plus de s. □

Proposition 14 (Inégalités de Sobolev logarithmiques locales)**.** *Pour toute fonction $f \in \mathcal{A}$ telle que $f > 0$ et pour tout $t \geqslant 0$,*

$$\frac{t}{2} \frac{|\nabla P_t f|^2}{P_t f} \leqslant P_t(f \log f) - P_t f \log P_t f \leqslant \frac{t}{2} P_t \left(\frac{|\nabla f|^2}{f} \right).$$

Démonstration. La preuve est similaire à celle des inégalités de Poincaré. On écrit

$$P_t(f \log f) - P_t f \log P_t f = \frac{1}{2} \int_0^t P_s \left[\frac{|\nabla P_{t-s} f|^2}{P_{t-s} f} \right] ds.$$

La propriété de sous-commutativité (forte) et l'inégalité de Cauchy-Schwarz donnent alors

$$|\nabla P_u g|^2 \leqslant (P_u |\nabla g|)^2 \leqslant P_u g\, P_u \frac{|\nabla g|^2}{g}$$

ce qui permet d'obtenir l'encadrement annoncé en prenant $u = t - s$ et $g = f$ pour la majoration, $u = s$ et $g = P_{t-s} f$ pour la minoration. □

De manière remarquable, les minorations dans l'inégalité de Poincaré locale et dans l'inégalité de Sobolev logarithmique locale fournissent automatiquement un contrôle de type lipschitzien sur $P_t f$ lorsque $t > 0$.

Corollaire 4. *Pour toute fonction $f \in \mathcal{A}$ telle que $0 \leqslant f \leqslant 1$ et tout $t > 0$, la fonction $P_t f$ est $1/\sqrt{t}$-lipschitzienne.*

Démonstration. Ceci résulte immédiatement de $t\,|\nabla P_t f|^2 \leqslant P_t f^2 \leqslant 1$. $\square$

Corollaire 5. *Pour toute fonction $f \in A$ telle que $0 < f \leqslant 1$ et tout $t > 0$, la fonction $\phi_t = \sqrt{-\log P_t f}$ est $(1/\sqrt{2t})$-lipschitzienne.*

Démonstration. Pour le second point, on utilise l'inégalité de Sobolev logarithmique locale qui indique que

$$|\nabla P_t f|^2 \leqslant \frac{2}{t}(P_t f)^2 \log \frac{1}{P_t f}.$$

Dès lors, comme

$$\nabla \phi_t = -\frac{1}{2\phi_t}\frac{1}{P_t f}\nabla P_t f,$$

il vient que

$$|\nabla \phi_t|^2 = \frac{|\nabla P_t f|^2}{4(P_t f)^2 \log \frac{1}{P_t f}} \leqslant \frac{1}{2t}.$$

$\square$

4.3 Une inégalité isopérimétrique gaussienne

Dans cette section, on renforce encore un peu l'hypothèse de convexité du potentiel W que nous venons de faire. On suppose qu'il existe une constante $c > 0$ telle que $W'' \geqslant c$ uniformément sur $\mathbb{R}^d$ en tant que forme bilinéaire symétrique, c'est-à-dire

$$W''(x,x) \geqslant c\,|x|^2, \qquad x \in \mathbb{R}^d.$$

La preuve que nous avons donnée de l'inégalité isopérimétrique pour la mesure gaussienne se généralise à toute mesure de Boltzmann μ dont le potentiel W vérifie cette hypothèse.

Théorème 6. *Si $W'' \geqslant c > 0$, alors pour toute fonction $f \in \mathcal{A}$ telle que $0 \leqslant f \leqslant 1$, on a*

$$\mathcal{I}\left(\int f d\mu\right) \leqslant \int \sqrt{\mathcal{I}^2(f) + \frac{1}{c}|\nabla f|^2}\, d\mu.$$

Démonstration. La preuve suit exactement le même schéma que dans le cas gaussien. Pour alléger les notations, posons $f_t \overset{\text{déf.}}{=} P_t f$ qui vérifie l'équation $\frac{d}{dt} f_t = L f_t$. On commence par montrer que la quantité

$$J(t) \overset{\text{déf.}}{=} \int \sqrt{\mathcal{I}^2(f_t) + c^{-1}|\nabla f_t|^2}\, d\mu$$

définit une fonction décroissante sur $[0\,;\infty[$. On a

$$\frac{dJ}{dt} = \int \frac{(\mathcal{I}\mathcal{I}')(f_t) L f_t + c^{-1}\nabla f_t \cdot \nabla L f_t}{K_t^{1/2}}\, d\mu,$$

où $K_t \overset{\text{déf.}}{=} \mathcal{I}^2(f_t) + c^{-1}|\nabla f_t|^2$. Notons déjà (on s'en servira plus loin) que

$$(\nabla K_t)_j = 2(\mathcal{I}\mathcal{I}')(f_t)\frac{\partial f_t}{\partial x_j} + \frac{2}{c}\nabla f_t \cdot \nabla \frac{\partial f_t}{\partial x_j},$$

d'où

$$\nabla f_t \cdot \nabla K_t = 2(\mathcal{I}\mathcal{I}')(f_t)|\nabla f_t|^2 + \frac{2}{c} f_t''(\nabla f_t, \nabla f_t),$$

et

$$\sum_{i=1}^{d} \frac{\partial f_t}{\partial x_i} \nabla \frac{\partial f_t}{\partial x_i} \cdot \nabla K_t = 2(\mathcal{I}\mathcal{I}')(f_t) f_t''(\nabla f_t, \nabla f_t) + \frac{2}{c}(f_t'')^2(\nabla f_t, \nabla f_t).$$

Or

$$Lf_t = \frac{1}{2}\left[\Delta f_t - \nabla W \cdot \nabla f_t\right] = \frac{1}{2}\sum_{i=1}^{d}\left(\frac{\partial^2 f_t}{\partial x_i^2} - \frac{\partial W}{\partial x_i}\frac{\partial f_t}{\partial x_i}\right),$$

donc

$$\begin{aligned}
\nabla f_t \cdot \nabla L f_t &= \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial f_t}{\partial x_j}\left(\frac{\partial^3 f_t}{\partial x_i^2 \partial x_j} - \frac{\partial^2 W}{\partial x_i \partial x_j}\frac{\partial f_t}{\partial x_i} - \frac{\partial W}{\partial x_i}\frac{\partial^2 f_t}{\partial x_i \partial x_j}\right) \\
&= -\frac{1}{2} W''(\nabla f_t, \nabla f_t) + \nabla f_t \cdot L \nabla f_t,
\end{aligned}$$

où $W''(u,v)$ désigne la forme bilinéaire symétrique donnée par la hessienne, et $L\nabla f_t$ désigne le vecteur obtenu en faisant agir L sur les coordonnées de ∇f (les dérivées partielles premières de f_t).

D'après la formule d'intégration par parties pour le générateur L, on a

$$\begin{aligned}
\int \frac{\nabla f_t \cdot L \nabla f_t}{K_t^{1/2}} d\mu &= \sum_{i=1}^{d} \int \frac{\frac{\partial f_t}{\partial x_i}}{K_t^{1/2}} L \frac{\partial f_t}{\partial x_i} d\mu = -\frac{1}{2}\sum_{i=1}^{d} \int \nabla \frac{\frac{\partial f_t}{\partial x_i}}{K_t^{1/2}} \cdot \nabla \frac{\partial f_t}{\partial x_i} d\mu \\
&= -\frac{1}{2}\sum_{i=1}^{d} \int \left[\frac{1}{K_t^{1/2}} \nabla \frac{\partial f}{\partial x_i} \cdot \nabla \frac{\partial f}{\partial x_i} - \frac{1}{2K^{3/2}} \frac{\partial f}{\partial x_i} \nabla K_t \cdot \nabla \frac{\partial f}{\partial x_i}\right] d\mu \\
&= -\frac{1}{2}\int \frac{\|f''\|_2^2}{K^{1/2}} + \frac{1}{2}\int \frac{\mathcal{I}\mathcal{I}' f''(\nabla,\nabla) + c^{-1}(f'')^2(\nabla,\nabla)}{K^{3/2}} \\
&= -\frac{1}{2}\int \frac{\|f''\|_2^2 \mathcal{I}^2 + c^{-1}\|f''\|_2^2 \nabla^2 - \mathcal{I}\mathcal{I}' f''(\nabla,\nabla) - c^{-1}(f'')^2(\nabla,\nabla)}{K^{3/2}} d\mu
\end{aligned}$$

Afin d'alléger les notations, nous avons noté de manière symbolique $\mathcal{I}$ et $\mathcal{I}'$ pour $\mathcal{I}(f_t)$ et $\mathcal{I}'(f_t)$, f pour f_t, ∇ pour ∇f_t, ∇^2 pour $|\nabla f_t|^2$, et enfin K pour K_t. En appliquant une seconde fois la formule d'intégration par parties, on obtient de la même manière

$$\begin{aligned}\int \frac{(\mathcal{I}\mathcal{I}')(f_t)}{K_t^{1/2}} L f_t d\mu &= -\frac{1}{2}\int \nabla \frac{(\mathcal{I}\mathcal{I}')(f_t)}{K_t^{1/2}} \cdot \nabla f_t d\mu \\ &= -\frac{1}{2}\int \left[\frac{\mathcal{I}'^2(f_t)-1}{K_t^{1/2}}|\nabla f_t|^2 - \frac{(\mathcal{I}\mathcal{I}')(f_t)}{2K_t^{3/2}}\nabla f_t\cdot\nabla K_t\right] d\mu \\ &= -\frac{1}{2}\int \left[\frac{(\mathcal{I}'^2-1)\nabla^2}{K^{1/2}} - \frac{\mathcal{I}\mathcal{I}'(\mathcal{I}\mathcal{I}'\nabla^2 + c^{-1}f''(\nabla,\nabla)}{K^{3/2}}\right] d\mu \\ &= -\frac{1}{2}\int \frac{c^{-1}\mathcal{I}'^2\nabla^4 - \mathcal{I}^2\nabla^2 - c^{-1}\nabla^4 - c^{-1}\mathcal{I}\mathcal{I}'f''(\nabla,\nabla)}{K^{3/2}} d\mu.\end{aligned}$$

Les calculs précédents montrent donc que

$$\frac{dJ}{dt} = -\frac{1}{2}\int \frac{A_t}{K_t^{3/2}} d\mu,$$

où le numérateur A_t est donné par

$$\begin{aligned}A_t = \frac{1}{c}\left[\mathcal{I}^2(f_t)\,\|f_t''\|_2^2 - 2(\mathcal{I}\mathcal{I}')(f_t)f_t''(\nabla f_t,\nabla f_t) + \mathcal{I}'^2(f_t)|\nabla f_t|^4\right] \\ + \frac{1}{c^2}\left[\|f_t''\|_2^2\,|\nabla f_t|^2 - (f_t'')^2(\nabla f_t,\nabla f_t)\right] + \left[\mathcal{I}^2(f_t) + \frac{1}{c}|\nabla f_t|^2\right]\left(\frac{W''}{c} - \mathrm{Id}\right)(\nabla f_t,\nabla f_t).\end{aligned}$$

D'après l'hypothèse sur W, le dernier terme de la somme est positif. Il en est de même des deux autres en vertu du lemme qui suit cette preuve, ce qui montre que J est décroissante, et en particulier, on a

$$\int \sqrt{\mathcal{I}^2(f) + \frac{1}{c}|\nabla f|^2}\, d\mu = J(0) \geqslant \int \sqrt{\mathcal{I}^2(f_t) + \frac{1}{c}|\nabla f_t|^2}\, d\mu$$

D'après le corollaire [XXX], on a par ailleurs $|\nabla f_t|^2 \leqslant \frac{1}{t}$ uniformément sur $\mathbb{R}^d$ et on peut montrer* que sous l'hypothèse $W'' \geqslant c$ on a convergence de $f_t(x) = P_t f(x)$ vers $\int f d\mu$ pour tout $x \in \mathbb{R}^d$. Ces deux remarques permettent de conclure en utilisant le théorème de convergence dominée. □

Lemme 10. *Soient S une matrice symétrique réelle de taille $d \geqslant 1$ et a, b deux réels. Pour tout vecteur $x \in \mathbb{R}^d$, on a*

$$\|S\|_2^2\,|x|^2 - S^2(x,x) \geqslant 0 \quad \textit{et}$$
$$a^2\,\|S\|_2^2 - 2abS(x,x) + b^2|x|^4 \geqslant 0.$$

Preuve du lemme. La première inégalité revient à écrire que $|Sx| \leqslant \|S\|_2\,|x|$ (conséquence de Cauchy-Schwarz), car $S^2(x,x) = |Sx|^2$. La seconde en découle également :

$$\begin{aligned}\left|aSx - b|x|^2x\right|^2 &= a^2|Sx|^2 - 2ab|x|^2S(x,x) + b^2|x|^6 \\ &\leqslant |x|^2\left(a^2\,\|S\|_2^2 - 2abS(x,x) + b^2|x|^4\right).\end{aligned}$$

□

*. Voir par exemple [Roy99].

4.4 Concentration et isopérimétrie en courbure positive

On souhaite dans cette section étudier les liens entre l'isopérimétrie et le phénomène de concentration pour une mesure de Boltzmann μ_W sous l'hypothèse de courbure positive (ou encore de convexité) $W'' \geqslant 0$. On a vu dès le premier chapitre qu'une inégalité isopérimétrique de la forme $I_\mu \geqslant c\,F' \circ F$ (où F est une fonction de répartition) peut s'intégrer pour fournir un contrôle de la fonction de concentration de μ de la forme $\alpha_\mu(r) \leqslant 1 - F(c\,r)$. On va ici montrer inversement comment un bonne majoration de la fonction de concentration permet de déduire une inégalité isopérimétrique pour la mesure μ. Le résultat de E. Milman (voir [Mil10]) peut s'énoncer ainsi : *

Théorème 7. *Soit μ une mesure de Boltzmann sur $\mathbb{R}^d$ de potentiel W convexe, et soit $\alpha\colon]0\,;\infty[\, \to \mathbb{R}_+$ décroissante telle que $\alpha_\mu(r) \leqslant \alpha(r)$ pour $r > 0$, $\lim_{r\to\infty} \alpha(r) = 0$ et $r(v) \overset{déf.}{=} \inf\{r > 0 : \alpha(r) < v\} > 0$ pour $v \in [0\,;1/2]$. Alors, il existe une constante $c > 0$ telle que*

$$I_\mu(v) \geqslant \frac{c}{r(v/4)}\, v \log \frac{1}{v}, \qquad v \in [0\,;1/2].$$

Avant de passer à la démonstration de ce théorème, on peut faire quelques remarques sur les résultats qu'il implique.

Remarque. Supposons par exemple que μ vérifie une propriété de concentration gaussienne. Alors on peut prendre $\alpha(r) = C\,e^{-r^2/C}$ pour une constante $C > 0$ assez grande, et donc $r(v) = \sqrt{C \log \frac{C}{v}}$. D'après le théorème, il existe une constante $c > 0$ telle que

$$I_\mu(v) \geqslant c\,v\,\frac{\log \frac{1}{v}}{\sqrt{C \log \frac{C}{v}}} \qquad v \in [0\,;1/2]$$

Quitte à modifier un peu la constante on a même

$$I_\mu(v) \geqslant c\,v\,\sqrt{\log \frac{1}{v}} \qquad v \in [0\,;1/2]$$

On peut donc trouver une constante $c' > 0$ telle que $I_\mu \geqslant c'\,\mathcal{I}$. En d'autres termes, la mesure μ vérifie une inégalité isopérimétrique gaussienne. Il y a équivalence entre la concentration gaussienne et l'isopérimétrie gaussienne sous l'hypothèse de courbure positive.

Remarque. Le théorème ne traite bien sûr pas que le cas gaussien. Par exemple si $\alpha(r) = Ce^{-r/c}$ (cas de concentration exponentielle), on a $r(v) = C \log \frac{C}{v}$ et le théorème entraîne l'existence d'une constante $c > 0$ telle que $I_\mu(v) \geqslant c\,v$ pour $v \in [0\,;1/2]$.

Alors que la preuve de E. Milman est purement géométrique, la démonstration que nous donnons ici, due à M. Ledoux ([Led11]), se base sur la méthode de semi-groupes que nous utilisons depuis le chapitre 3. Nous la découpons en trois étapes :

$*$. À vrai dire, la constante 4 est certainement superflue ici et n'apparaît d'ailleurs pas dans le résultat original de Milman. Elle n'est cependant pas très gênante.

1. On établit en guise de lemmes préliminaires des inégalités de déviation pour P_t de manière à relier le semi-groupe et la fonction de concentration.
2. On montre que si on a une inégalité isopérimétrique pour tout v dans un voisinage de 0, on peut l'étendre à tout $v \in [0\,;1/2]$.
3. On vérifie qu'on a effectivement une inégalité isopérimétrique au voisinage de 0.

4.4.1 Deux inégalités de déviation pour $P_t f$

On sait que l'hypothèse de courbure positive entraîne automatiquement un contrôle lipschitzien sur $P_t f$. On a aussi vu dans le chapitre introductif qu'un tel contrôle se traduit naturellement en terme d'inégalité de déviation majorée par la fonction de concentration.

Lemme 11. *Pour toute fonction $f \in \mathcal{A}$ telle que $0 \leqslant f \leqslant 1$, pour tous $t > 0$ et $r > 0$,*

$$\mu\left\{P_t f \geqslant 2\int f d\mu + \frac{r}{\sqrt{t}}\right\} \leqslant \alpha_\mu(r).$$

Démonstration. Soit $m = 2\int f d\mu$ supposé non nul (sinon $\mu\{P_t f > 0\} = 0$). D'après l'inégalité de Markov et en utilisant l'invariance de μ par rapport au semi-groupe (P_t), on voit que

$$\mu\{P_t f > m\} \leqslant \frac{1}{m}\int P_t f\, d\mu = \frac{1}{2}.$$

Comme de plus, la fonction $x \mapsto P_t f(x)$ est $\frac{1}{\sqrt{t}}$-lipschitzienne, le phénomène de concentration de la mesure peut s'exprimer sous la forme donnée par l'énoncé. □

Lemme 12. *Pour toute fonction $f \in \mathcal{A}$ telle que $0 < f \leqslant 1$ et $\int f d\mu \leqslant 1/2$, pour tous $t > 0$ et $r > 0$,*

$$\mu\left\{P_t f \geqslant \left(2\int f d\mu\right)^{\frac{1}{2}} e^{r^2/2t}\right\} \leqslant \alpha_\mu(r).$$

Démonstration. On va utiliser le fait que $\phi_t = \sqrt{-\log P_t f}$ est $\frac{1}{\sqrt{2t}}$-lipschitzienne. On a nécessairement $\int f d\mu > 0$ car sinon $\mu\{f > 0\} = 0$. Posons $m = (\log \frac{1}{2\int f d\mu})^{\frac{1}{2}}$. En utilisant l'inégalité de Markov et l'invariance de μ par rapport au semi-groupe (P_t) on vérifie que

$$\mu\{\phi_t < m\} = \mu\left\{P_t f > e^{-m^2}\right\} \leqslant e^{m^2}\int f d\mu = \frac{1}{2}.$$

Dès lors, le phénomène de concentration de la mesure s'exprime pour ϕ_t sous la forme

$$\mu\left\{\phi_t \leqslant m - \frac{r}{\sqrt{2t}}\right\} \leqslant \alpha_\mu(r).$$

Ce n'est pas tout à fait l'inégalité voulue, mais elle l'implique. Afin que l'événement considéré ne soit pas vide (l'inégalité serait trivialement vérifiée), on peut en effet supposer que

$$\frac{r^2}{2t} - \frac{m^2}{2} = \log\left[\left(2\int f d\mu\right)^{\frac{1}{2}} e^{r^2/2t}\right] \leqslant 0$$

et sous cette hypothèse, on a par concavité de la racine carrée

$$\sqrt{\frac{m^2}{2} - \frac{r^2}{2t}} \leqslant m - \frac{r}{\sqrt{2t}},$$

de sorte qu'après une rapide réécriture

$$\mu\left\{P_t f \geqslant \left(2\int f d\mu\right)^{\frac{1}{2}} e^{r^2/2t}\right\} \leqslant \mu\left\{\phi_t \leqslant m - \frac{r}{\sqrt{2t}}\right\} \leqslant \alpha_\mu(r).$$

□

4.4.2 Isopérimétrie linéaire pour les boréliens assez gros

Supposons qu'on a trouvé un $a \in]0\,;1/2[$ et une constante $c_1 > 0$ telle que

$$I_\mu(v) \geqslant \frac{c_1}{r(v/4)}\, v \log\frac{1}{v} \qquad v \in [0\,;a].$$

Pour étendre cette inégalité aux $v \in [a\,;1/2]$, il suffit de trouver une constante $c_2 > 0$ telle que

$$I_\mu(v) \geqslant c_2\, v, \qquad v \in [a\,;1/2].$$

En effet, on aura alors l'inégalité isopérimétrique voulue pour tout $v \in [0\,;1/2]$ avec la constante $c = \min\{c_1, \frac{c_2 r(\frac{1}{8})}{\log\frac{1}{a}}\} > 0$. Par ailleurs, on peut se contenter de travailler au niveau quadratique pour établir une telle inégalité linéaire.

Lemme 13. *Pour toute fonction $f\colon \mathbb{R}^d \to [0\,;1]$ lipschitzienne et pour tout $t > 0$, on a*

$$\int f^2 d\mu - \int (P_t f)^2 d\mu \leqslant 2\sqrt{t}\int |\nabla f| d\mu.$$

Démonstration. On commence par supposer $f \in \mathcal{A}$. On a alors

$$\int f^2 d\mu - \int (P_t f)^2 d\mu = -\int_0^t \frac{d}{ds}\left[\int (P_s f)^2 d\mu\right] ds = \int_0^t \int |\nabla P_s f|^2\, d\mu ds.$$

Or, on a vu que $|\nabla P_s f|^2 \leqslant \frac{1}{s}$ et aussi $|\nabla P_s f| \leqslant P_s|\nabla f|$ sous la condition $W'' \geqslant 0$. Ainsi a-t-on

$$\int |\nabla P_s f|^2 d\mu \leqslant \frac{1}{\sqrt{s}}\int |\nabla P_s f| d\mu \leqslant \frac{1}{\sqrt{s}}\int |\nabla f| d\mu.$$

On obtient alors en intégrant l'inégalité

$$\int f^2 d\mu - \int (P_t f)^2 d\mu \leqslant 2\sqrt{t} \int |\nabla f| d\mu,$$

qui s'étend à f lipschitzienne par les arguments d'approximation classiques. $\square$

Soit A un fermé de $\subset \mathbb{R}^d$ (on a déjà vu que ce n'est pas restrictif) tel que $0 < a \leqslant \mu(A) \leqslant \frac{1}{2}$. En reprenant les arguments de la démonstration du théorème 3, on voit que l'inégalité fournie par le lemme ci-dessus s'étend à $f = \mathbf{1}_A$ par approximation, et s'écrit alors

$$2\sqrt{2t}\, \mu^+(A) \geqslant \mu(A) - \int (P_t \mathbf{1}_A)^2 d\mu.$$

On va pouvoir contrôler le dernier terme à l'aide de la fonction de concentration. Considérons pour cela le borélien $B = \{P_t \mathbf{1}_A \leqslant (1+\tau)\mu(A)\}$ où $\tau > 0$ sera fixé ultérieurement. En utilisant l'inégalité de Markov et l'invariance de la mesure μ par rapport au semi-groupe (P_t), on voit que

$$1 - \mu(B) = \mu\{P_t f > (1+\tau)\mu(A)\} \leqslant \frac{1}{(1+\tau)\mu(A)} \int P_t \mathbf{1}_A d\mu = \frac{1}{1+\tau},$$

c'est à dire que $\mu(B) \geqslant \rho \overset{\text{déf.}}{=} \tau/(1+\tau) > 0$. D'après le lemme 2, on sait que pour tout $r_0 > r_\mu(\rho)$ (en d'autres termes $\alpha_\mu(r_0) < \rho$) et pour tout $r > 0$, on a

$$1 - \mu(B_{r+r_0}) \leqslant \alpha_\mu(r).$$

Or puisque $P_t \mathbf{1}_A$ est $\frac{1}{\sqrt{t}}$-lipschitzienne (c'est la limite simple des $P_t f_h$ qui le sont d'après le lemme 4), le résultat du lemme 11 est valable, d'où

$$B_{r+r_0} \subset \left\{ P_t \mathbf{1}_A < (1+\tau)\mu(A) + \frac{r+r_0}{\sqrt{t}} \right\}$$

et on en déduit que

$$\mu \left\{ P_t \mathbf{1}_A \geqslant (1+\tau)\mu(A) + \frac{r+r_0}{\sqrt{t}} \right\} \leqslant \alpha_\mu(r).$$

Comme annoncé, on peut donc contrôler l'intégrale en $(P_t \mathbf{1}_A)^2$:

$$\int (P_t \mathbf{1}_A)^2 d\mu \leqslant \left[(1+\tau)\mu(A) + \frac{r+r_0}{\sqrt{t}} \right]^2 + \alpha_\mu(r)$$

et on obtient alors

$$2\sqrt{t}\, \mu^+(A) \geqslant \mu(A) - \left[(1+\tau)\mu(A) + \frac{r+r_0}{\sqrt{t}} \right]^2 - \alpha_\mu(r).$$

On peut prendre de manière arbitraire $\tau = \frac{1}{10}$, ce qui fixe $\rho = \frac{1}{11}$, et détermine également $r_0 > r_\mu(\rho)$. Fixons aussi $r > 0$ assez grand pour que $\alpha_\mu(r) \leqslant \frac{a}{8}$ (c'est à dire $r > r_\mu(a/8)$) et $t > 0$ de façon à ce que $\frac{r+r_0}{\sqrt{t}} \leqslant \frac{a}{10}$. Pour ces choix,

$$2\sqrt{t}\,\mu^+(A) \geqslant \mu(A) - \left[\frac{12}{10}\mu(A)\right]^2 - \frac{\mu(A)}{8} \geqslant \frac{\mu(A)}{8},$$

et cette inégalité est indépendante du borélien A pourvu que $\mu(A) \in [a\,;1/2]$. On a donc montré l'existence d'une constante $c_2 > 0$ (ici $c_2 = \frac{1}{16\sqrt{t}}$) telle que

$$I_\mu(v) \geqslant c_2\, v, \qquad v \in [a\,;1/2]$$

et c'est précisément ce qu'on voulait.

4.4.3 Isopérimétrie sur-linéaire pour les boréliens petits

Conformément à ce qu'on a déjà prouvé, il suffit de trouver un $a \in \,]0\,;1/2[$ et une constante $c_1 > 0$ tels que

$$\mu^+(A) \geqslant \frac{c_1}{r(\mu(A)/4)}\mu(A)\log\frac{1}{\mu(A)}, \qquad 0 \leqslant \mu(A) \leqslant a.$$

Comme la fonction dans le membre de droite est croissante en $\mu(A)$, on peut se contenter de vérifier cette inégalité pour les fermés.

Lemme 14. *Soit $\eta \in \,]0\,;1[$. Pour toute fonction $f\colon \mathbb{R}^d \to [\eta\,;1]$ lipschitzienne, et pour tout $t > 0$,*

$$\int f\log f d\mu - \int P_t f \log P_t f \leqslant 2\left(2t\log\frac{1}{\eta}\right)^{\frac{1}{2}} \int |\nabla f| d\mu.$$

Démonstration. On commence par supposer $f \in \mathcal{A}$. D'après la formule d'intégration par parties on a pour tout $t > 0$,

$$\frac{d}{dt}\left[\int P_t f \log P_t f d\mu\right] = \int [1+\log P_t f] L P_t f d\mu = -\int \frac{|\nabla P_t f|^2}{P_t f} d\mu.$$

Or on sait (en conséquence des inégalités de Sobolev logarithmiques locales) que,

$$|\nabla P_t f| \leqslant \sqrt{\frac{2}{t}(P_t f)^2 \log\frac{1}{P_t f}} \leqslant \sqrt{\frac{2}{t}\log\frac{1}{\eta}}P_t f,$$

d'où, en utilisant la propriété de commutation $|\nabla P_t f| \leqslant P_t|\nabla f|$,

$$\int \frac{|\nabla P_t f|^2}{P_t f} d\mu \leqslant \sqrt{\frac{2}{t}\log\frac{1}{\eta}} \int |\nabla P_t f| d\mu \leqslant \sqrt{\frac{2}{t}\log\frac{1}{\eta}} \int |\nabla f| d\mu.$$

En intégrant cette inégalité on obtient alors

$$\int f \log f d\mu - \int P_t f \log P_t f d\mu = \int_0^t \int \frac{|\nabla P_s f|^2}{P_s f} d\mu ds \leqslant \sqrt{\log \frac{1}{\eta}} \int |\nabla f| d\mu \int_0^t \sqrt{\frac{2}{s}} ds,$$

c'est l'inégalité recherchée. $\square$

Soit $A \subset \mathbb{R}^d$ une partie fermée. En considérant la famille de fonctions f_h lipschitziennes définies pour tout $h > 0$ et $x \in \mathbb{R}^d$et $x \in \mathbb{R}^d$ par

$$f_h(x) = \eta \vee \left(1 - \frac{d(x, A_{h^2})}{h}\right)_+ .$$

et en raisonnant comme dans le théorème 3, on voit qu'on peut étendre l'inégalité fournie par le lemme ci-dessus à $f = \eta \vee \mathbf{1}_A$. On obtient alors

$$2\left(2t \log \frac{1}{\eta}\right)^{\frac{1}{2}} \mu^+(A) \geqslant -\eta \log \frac{1}{\eta} + \int P_t f \log \frac{1}{P_t f} d\mu.$$

Posons $\beta_t(r) = \left(2 \int f d\mu\right)^{\frac{1}{2}} e^{r^2/2t} = (2[\eta + (1-\eta)\mu(A)])^{\frac{1}{2}} e^{r^2/2t}$. D'après le lemme [XXX] appliqué à f (c'est une limite de fonction $1/\sqrt{2t}$-lipschitziennes), on a l'inégalité $\mu\{P_t f \geqslant \beta_t(r)\} \leqslant \alpha_\mu(r)$ pourvu que $\eta + (1-\eta)\mu(A) \leqslant \frac{1}{2}$. Par ailleurs, pour tout $\beta > 0$ on a

$$\int P_t f \log \frac{1}{P_t f} d\mu \geqslant \int_{\{P_t f \leqslant \beta\}} P_t f \log \frac{1}{\beta} d\mu = \log \frac{1}{\beta} \left[\int P_t f d\mu - \int_{\{P_t f > \beta\}} P_t f d\mu\right]$$

et on en déduit la minoration

$$\int P_t f \log \frac{1}{P_t f} d\mu \geqslant [\eta + (1-\eta)\mu(A) - \alpha_\mu(r)] \log \frac{1}{\beta_t(r)},$$

ce qui donne

$$2\left(2t \log \frac{1}{\eta}\right)^{\frac{1}{2}} \mu^+(A) \geqslant -\eta \log \eta + [\eta + (1-\eta)\mu(A) - \alpha_\mu(r)] \log \frac{1}{\beta_t(r)}.$$

Afin d'optimiser cette quantité, nous suivons [Led11] en choisissant $\eta = \mu(A)^2$ et $t \log \frac{1}{\eta} = r^2$. On voit alors que dès que $\mu(A) < \frac{1}{16}$, on a pour tout $r > 0$

$$2\sqrt{2} r \mu^+(A) \geqslant -2\mu(A)^2 \log \frac{1}{\mu(A)} + \frac{1}{4}\left[\frac{\mu(A)}{2} - \alpha_\mu(r)\right] \log \frac{1}{16\mu(A)}.$$

En particulier, pour $r = \varepsilon + r_\mu(\mu(A)/4)$ avec $\varepsilon > 0$, on a par définition $\alpha_\mu(r) < \mu(A)/A$ et donc,

$$2\sqrt{2}\, r\, \mu^+(A) \geqslant \frac{1}{16}\mu(A) \log \frac{1}{\mu(A)} + o\left(\mu(A) \log \frac{1}{\mu(A)}\right).$$

On peut donc trouver une constante $c_1 > 0$ (par exemple $c_1 = \frac{1}{64}$) et un $a > 0$ assez petit tel que dès que $v = \mu(A) \leqslant a$,

$$(r_\mu(v/4) + \varepsilon)\, \mu^+(A) \geqslant c_1\, v \log \frac{1}{v}.$$

En faisant tendre $\varepsilon \to 0$ on a donc

$$I_\mu(v) \geqslant \frac{C}{r(v/4)} v \log \frac{1}{v}, \qquad v \in [0\,; a].$$

Ceci achève la démonstration du théorème. $\square$

Bibliographie

[An00] Cécile ANÉ et al. *Sur les inégalités de Sobolev logarithmiques*. Panoramas et Synthèses 10. Société Mathématique de France, 2000.

[BH97] Serguei BOBKOV et Christian HOUDRÉ. *Some Connections between Isoperimetric and Sobolev-type Inequalities*. T. 129. Memoirs of the American Mathematical Society 616. American Mathematical Society, 1997.

[BL96] Dominique BAKRY et Michel LEDOUX. « Lévy-Gromov's isoperimetric inequality for an infinite dimensional diffusion generator ». Dans : *Inventiones mathematicae* 123 (1996), p. 259–281.

[BM00] Frédéric BARTHE et Bernard MAUREY. « Some remarks on isoperimetry of Gaussian type ». Dans : *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 36.4 (2000), p. 419–434.

[Bob96] Sergey G. BOBKOV. « A functional form of the isoperimetric inequality for the Gaussian measure ». Dans : *Journal of Functional Analysis* 135.1 (1996), p. 36–49.

[Bob97] Sergey G. BOBKOV. « An isoperimetric inequality on the discrete cube and an elementary proof of the isoperimetric inequality in Gauss space ». Dans : *The Annals of Probability* 25.1 (1997), p. 206–214.

[Bor75] Christer BORELL. « The Brunn-Minkowski inequality in Gauss space ». Dans : *Inventiones mathematicae* 30.2 (1975), p. 207–216.

[Ehr83] Antoine EHRHARD. « Symétrisation dans l'espace de Gauss ». Dans : *Mathematica Scandinavica* 53.2 (1983), p. 281–301.

[Led01] Michel LEDOUX. *The concentration of measure phenomenon*. Mathematical Surveys and Monographs 89. American Mathematical Society, 2001.

[Led11] Michel LEDOUX. « From concentration to isoperimetry: semigroup proofs ». Dans : *Concentration, Functional Inequalities and Isoperimetry*. Sous la dir. de Christian HOUDRÉ et al. American Mathematical Society, 2011.

[Led99] Michel LEDOUX. « Concentration of measure and logarithmic Sobolev inequalities ». Dans : *Séminaire de Probabilités*. T. XXXIII. Lecture Notes in Mathematics. Berlin : Springer, 1999, p. 120–216.

[Lig10] Thomas M. LIGGETT. *Continuous Time Markov Processes*. Graduate Studies in Mathematics 113. American Mathematical Society, 2010.

[Lé51] Paul LÉVY. *Problèmes concrets d'analyse fonctionnelle*. Gauthier-Villars, 1951.

[Mil10] Emanuel MILMAN. « Isoperimetric and concentration inequalities : equivalence assuming curvature lower bound ». Dans : *Duke Mathematical Journal* 154.2 (2010), p. 207–239.

[Roy99] Guy ROYER. *Une initiation aux inégalités de Sobolev logarithmiques*. Cours Spécialisés 5. Société Mathématiques de France, 1999.

[RY99] Daniel REVUZ et Marc YOR. *Continuous Martingales and Brownian Motion*. Grundlehren der Mathematischen Wissenschaften 293. Springer-Verlag, 1999.

[ST78] V. N. SUDAKOV et B. TSIREL'SON. « Extremal properties of half-spaces for spherically invariant measures ». Dans : *Journal of Mathematical Science* 9.1 (1978), p. 9–18.