CHAPITRE 5

REMARQUES SUR LES INÉGALITÉS DE POINCARÉ, DE CHEEGER ET DE SOBOLEV LOGARITHMIQUE

5.1. Généralités

Toutes les probabilités considérées seront supposées être absolument continues par rapport à la mesure de Lebesgue.

On se donne une probabilité μ sur $\mathbb{R}$. On dit que μ satisfait une inégalité de Poincaré s'il existe une constante $c \geq 0$ telle que pour toute fonction f C^1 à support compact, (ou simplement C^1 ou plus généralement localement lipschitzienne) on a

$$\mathrm{Var}_\mu(f) := \int \left(f - \textstyle\int f\,d\mu\right)^2 \leq c \int |\nabla f|^2\,d\mu.$$

La meilleure constante c, c'est à dire la plus petite, pour laquelle cette inégalité est valable s'appelle la constante de Poincaré, et on la note $\mathrm{c_P}(\mu)$. La constante $\mathrm{c_P}(\mu)^{-1}$ s'appelle aussi la constante de trou spectral, pour la raison suivante. Supposons que μ ait une densité, que l'on note e^{-V}. Indoduisons l'opérateur différentiel

$$Lf = \Delta f - \nabla V \cdot \nabla f.$$

Alors, pour des fonctions C^2 à support compact on a

$$-\int (Lf) g\,d\mu = \int \nabla f \cdot \nabla g\,d\mu.$$

Ainsi, l'opérateur L est autoajoint pour le produit scalaire sur $L^2(\mu)$, d'abord sur le sous-espace des fonction régulières à support compact. On peut étendre L à un domaine maximal fermé, pour lequel l'intégration par partie ci-dessous reste valable, et on aura alors un opérateur auto adjoint non borné (mais on va éviter de rentrer dans ce genre de difficultés). L'important est d'avoir une algèbre de fonctions suffisamment riche sur lequel étudier L (on donnera des exemples plus loin). Quoi qu'il en soit, on voit que L est nul sur les fonction constantes, et la formule d'intégration par partie nous dit que ce sont les seuls éléments du noyau. Par ailleurs, on voit aussi que $-L$ est positif. Donc 0 est la plus petite valeur propre : $\lambda_0 = 0$. L'inégalité de Poincaré dit qu'il existe $c > 0$ tel qu'il n'y ait aucune valeur propre entre 0 et $\frac{1}{c}$, puisque $f - \int f\,d\mu$ décrit l'orthogonal aux fonctions constantes, sur lequel on a $-\int (Lf) f\,d\mu \geq \frac{1}{c} \int f^2\,d\mu$. On dit que L admet un trou spectral. Ainsi, la plus petite valeur propre après 0 est strictement positive, et vaut

$$\lambda_1(\mu) = \frac{1}{\mathrm{c_P}(\mu)}.$$

Soit (X, μ) un espace de probabilité. Une fonction positive $f : X \longrightarrow \mathbb{R}^+$ est dite d'entropie finie si f et $f \log f$ sont dans $L_1(\mu)$, et on définit alors l'entropie de f par rapport à μ par

$$\begin{aligned}\mathrm{Ent}_\mu(f) &:= \int f \log(f)\, d\mu - \left(\int f\, d\mu\right) \log\left(\int f\, d\mu\right) \\ &= \int f \log\left(\frac{f}{\int f\, d\mu}\right) d\mu \\ &= \int f \log(f)\, d\mu \quad \text{si } \int f\, d\mu = 1.\end{aligned}$$

Puisque la fonction $s \longrightarrow s\log(s)$ est strictement convexe sur $\mathbb{R}^+$, on a

$$\mathrm{Ent}_\mu(f) \geq 0 \quad \text{et} \quad \mathrm{Ent}_\mu(f) = 0 \Leftrightarrow f \equiv \int f\, d\mu \ \ (\mu - pp).$$

On veut justement utiliser l'entropie pour quantifier que f est proche de sa moyenne sur un ensemble de grande mesure.

On va maintenant revenir à une mesure de probabilité μ sur $\mathbb{R}^n$ équipé de sa structure euclidienne usuelle. Rappelons (théorème de Rademacher) qu'une fonction localement lipschitzienne est différentiable presque partout et sa différentielle (presque partout) est une fonction localement intégrable qui coïncide avec la différentielle au sens des distributions. En particulier, si f est localement lipschitzienne avec $|\nabla f| \in L_1(\gamma_n)$, alors $P_t(\nabla f) \longrightarrow \nabla f$ lorsque $t \to 0$ où P_t est un semi-groupe gaussien (voir chapitre suivant).

***Definition* (LSI(ρ_0)).** — *On dit qu'une probabilité μ vérifie une inégalité de Sobolev logarithmique avec constante $\rho > 0$ (en abrégé « LSI(ρ) ») si pour toute fonction $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ suffisamment régulière (i.e. g^2 d'entropie finie et g est de classe C^1 ou localement lipschitzienne) on a :*

$$\mathrm{Ent}_\mu(g^2) \leq \frac{2}{\rho} \int |\nabla g|^2\, d\mu. \tag{1}$$

La meilleure constante ρ (c'est-à-dire la plus grande) s'appelle la constante de log-Sobolev, et on la note $\lambda_{\mathrm{LS}}(\mu)$.

On préfère parfois énoncer le résultat de la manière équivalente suivante : pour toute fonction positive $f : \mathbb{R}^n \longrightarrow \mathbb{R}^+$ d'entropie finie avec $\sqrt{f}$ de classe C^1 (ou localement lipschitzienne), on a

$$\mathrm{Ent}_\mu(f) \leq \frac{1}{2\rho_0} \int \frac{|\nabla f|^2}{f}\, d\mu. \tag{2}$$

L'intégrale de droite s'appelle *l'information de Fisher de f par rapport à μ*

Remarque. — *Notons que si $g := \sqrt{f}$ est de classe C^1, alors,*

$$2g\nabla g = \nabla f$$

et donc $\frac{|\nabla f|^2}{f}$ est interprété, en un point où f s'annule, comme le prolongement (par continuité) $4|\nabla g|^2$. Par ailleurs, puisque g est positive, on a $f(x) = 0 \Leftrightarrow g(x) = 0 \Rightarrow \nabla g(x) = 0$ puisque 0 est un minimum local, et donc, on conclut qu'on a toujours

$$\int \frac{|\nabla f|^2}{f}\, d\mu = \int_{f>0} \frac{|\nabla f|^2}{f}\, d\mu.$$

Le résultat reste vrai lorsque $g = \sqrt{f}$ *est localement lipschitzienne, puisqu'on a alors* $\nabla g = 0$ *presque partout sur l'ensemble* $\{g = 0\}$.

L'inégalité de Sobolev logarithmique est plus forte que l'inégalité de Poincaré.

Proposition. — *Si* μ *satisfait une inégalité de log-sob, alors elle satisfait une inégalité de Poincaré, avec*

$$\lambda_1(\mu) := c_P(\mu)^{-1} \geq \lambda_{LS}(\mu).$$

Démonstration. — Soit g C^1 à support compact d'intégrale nulle. On considère la fonction lipschitzienne $f = f_\varepsilon$ donnée par

$$f_\varepsilon(x) = 1 + \varepsilon g(x)$$

avec ε petit. On remarque que $\int f \, d\mu = 1$ et que $\log f_\varepsilon = \varepsilon g - \varepsilon^2 g^2/2 + o(\varepsilon^2)$ uniformément en x sur $\mathbb{R}^n$ lorsque $\varepsilon \to 0$ (par développement de Taylor avec reste). On a donc

$$\int f_\varepsilon \log(f_\varepsilon) \, d\mu = \varepsilon^2 \int \left[g^2 - \frac{1}{2} g^2\right] d\mu + O(\varepsilon^3),$$

et de même

$$\int \frac{|\nabla f_\varepsilon|^2}{f_\varepsilon} \, d\mu = \varepsilon^2 \int |\nabla g|^2 \, d\mu + O(\varepsilon^3).$$

En comparant les termes en ε^2, on trouve l'inégalité de Poincaré souhaitée. □

Pour nous, l'exemple fondamental sera donné par la mesure gaussienne, mais il existe de nombreux autre exemple.

En particulier, on peut s'intéresser au cas où μ est la mesure uniforme sur un ensemble Ω de $\mathbb{R}^n$, i.e. $d\mu(x) = \frac{1_\Omega}{|\Omega|} \, dx$. Le problème spectral mentionné plus haut correspond alors au Laplacien

$$L = \Delta$$

avec condition de Neumann. L'opérateur $-L$ est autoajoint positiv dans $L^2(\Omega)$, sur son domaine inclus dans $L^2(\Omega)$, qui est ici l'espace de Sobolev $H_1(\Omega)$ avec conditions de Neumann. Dans ce cas, les constantes sont bien des vecteurs propres associés à la valeur propre 0 (noyau). La valeur propre suivante est donnée par l'infimum de

$$\frac{-\int (Lg) g \, d\mu}{\int g^2 \, d\mu} = \frac{\int |\nabla g|^2 \, d\mu}{\int g^2 \, d\mu}$$

sur les fonctions orthogonales au premier vecteur propre, c'est-à-dire sur l'espace des fonctions telles que $\int g \, d\mu = 0$.

Si Ω est un ouvert connexe relativement compact, alors on sait qu'il aura un trou spectral (Neumann) strictement positif. Mais ce qui nous intéresse, ce sont des estimations quantitatives.

L'exemple le plus simple est $[0,1] \subset \mathbb{R}$. On peut alors montrer que $\lambda_1 = \pi^2$ et que le vecteur propre correspondant est $g_1(x) = \cos(\pi x)$. Soit encore, pour g dérivable avec $\int_{[0,1]} g = 0$ on a

$$\int_{[0,1]} g^2 \leq \frac{1}{\pi^2} \int_{[0,1]} (g')^2.$$

Il existe deux manière d'établir ce résultat : soit en étudiant de manière qualitative les solution de $g'' - \lambda g = 0$, soit en se ramenant à un développement en série de Fourier.

5.2. Tensorisation

Proposition. — *Soit μ_1 et μ_2 deux probabilités sur $\mathbb{R}^{n_1}$ et $\mathbb{R}^{n_2}$ respectivement. Si elles satisfont une inégalité de Poincaré, alors $\mu_1 \otimes \mu_2$ aussi, et de plus on a*

$$\mathrm{c_P}(\mu_1 \otimes \mu_2) = \max\{\mathrm{c_P}(\mu_1), \mathrm{c_P}(\mu_2)\}.$$

Démonstration. — Soit $g : \mathbb{R}^{n_1+n_2} \to \mathbb{R}$ une fonction régulière à support compact. Introduisons les fonctions

$$g_t(y) = g(t, y) \text{ (pour } t \in \mathbb{R}^{n_1} \text{ fixé)} \quad \text{ et } \quad G(x) = \int_{\mathbb{R}^{n_2}} g(x, y)\, d\mu_2(y).$$

On a la décomposition classique de la variance :

$$V := \int_{\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}} \Big(g(x, y) - \textstyle\int g\Big)^2 d\mu_1(x) d\mu_2(y) = \displaystyle\int_{\mathbb{R}^{n_1}} \mathrm{Var}_{\mu_2}(g_x)\, d\mu_1(x) + \mathrm{Var}_{\mu_1}(G).$$

Ainsi

$$V \leq \mathrm{c_P}(\mu_2) \int_{\mathbb{R}^{n_1}} \int_{\mathbb{R}^{n_2}} |\partial_y g_x|^2 d\mu_1(y) d\mu_1(x) + \mathrm{c_P}(\mu_1) \int_{\mathbb{R}^{n_1}} |\nabla_x G|^2\, d\mu_1(x).$$

On conclut en remarquant que, par dérivation sous l'intégrale et Cauchy-Schwarz (ou Jensen), $|\nabla_x G|^2 \leq \int_{\mathbb{R}^{n_2}} |\partial_x g|^2\, d\mu_2(x)$.

On a donc établi l'inégalité $\mathrm{c_P}(\mu_1 \otimes \mu_2) \leq \max\{\mathrm{c_P}(\mu_1), \mathrm{c_P}(\mu_2)\}$. Mais on ne peut pas faire mieux. Prenons f_1 non-constant qui réalise (presque... prendre un $\varepsilon > 0$...) l'égalité dans Poincaré pour μ_1. Si on considère la fonction $f(x, y) = f_1(x)$, alors on trouve

$$\mathrm{Var}_{\mu_1\otimes\mu_2}(f) = \mathrm{Var}_{\mu_1}(f) \leq \mathrm{c_P}(\mu_1 \otimes \mu_2) \int_{\mathbb{R}^{n_1}} |\nabla_x f_1|^2\, d\mu_1(x),$$

et par le choix de f_1 (ou en prenant le sup sur les f_1) on a $\mathrm{c_P}(\mu_1) \leq \mathrm{c_P}(\mu_1 \otimes \mu_2)$. On a de même pour μ_2. □

Théorème. — *Soit μ_1 et μ_2 deux probabilités sur $\mathbb{R}^{n_1}$ et $\mathbb{R}^{n_2}$, respectivement. Si μ_1 vérifie* LSI(ρ_1) *et μ_2 vérifie* LSI(ρ_2), *alors $\mu_1 \otimes \mu_2$ vérifie* LSI$(\min(\rho_1, \rho_2))$.

En particulier, si μ vérifie LSI(ρ), alors $\mu^{\otimes N}$ vérifie aussi LSI(ρ) pour tout $N \geq 1$.

Ce qu'il y a d'important dans ce dernier énoncé, c'est que la constante est indépendante de la dimension N.

Démonstration. — Soit $G : \mathbb{R}^{n_1+n_2} \longrightarrow \mathbb{R}$ régulière. Par Fubini, on écrit (astucieusement)

$$\mathrm{Ent}_{\mu_1\otimes\mu_2}(G^2) = \int \left(\int G^2 \log\left(\frac{G^2}{\int G^2\, d\mu_1}\right) d\mu_1 \right) d\mu_2 + \int \left(\int G^2\, d\mu_1 \right) \log\left(\frac{\int G^2\, d\mu_1}{\int G^2\, d\mu_1 d\mu_2} \right) d\mu_2.$$

Dans la première intégrale, on applique LSI(ρ_1) à la fonction $x \to G(x, y)$, pour $y \in \mathbb{R}^{n_2}$ fixé. Dans la deuxième intégrale, on applique LSI(ρ_2) sous la forme (2) à la fonction $f(y) = \int G^2(x, y)\, d\mu_1(x)$. On obtient donc, en écrivant ∇_x pour le vecteur formé par les n_1 premières dérivées partielles (idem pour ∇_y),

$$\mathrm{Ent}_{\mu_1\otimes\mu_2}(G^2) \leq \frac{2}{\rho_1} \iint |\nabla_x G|^2\, d\mu_1 d\mu_2 + \frac{1}{2\rho_2} \int \frac{\left|\nabla_y \int G^2\, d\mu_1\right|^2}{\int G^2\, d\mu_1}\, d\mu_2.$$

On conclut en dérivant sous l'intégrale puis utilisant l'inégalité de Cauchy-Schwarz, pour $y \in \mathbb{R}^{n_2}$ fixé,

$$\left|\nabla_y \int G^2\, d\mu_1\right|^2 = 4\left|\int G\nabla_y G\, d\mu_1\right|^2 \leq \int G^2\, d\mu_1 \int |\nabla_y G|^2\, d\mu_1,$$

et enfin le fait que $|\nabla_x G|^2 + |\nabla_y G|^2 = |\nabla G|^2$. □

De manière plus systématique, on montre (voir exercice ci-dessous) une inégalité de tensorisation pour l'entropie, de laquelle la tensorisation de log-Sobolev découle immédiatement (en n'utilisant que la forme (1) de log-Sobolev, ce qui sera important dans le cas discret).

***Remarque* (Tensorisation de l'entropie).** — *Soit $(\Omega_1, \mu_1), \ldots, (\Omega_n, \mu_n)$ des espaces de probabilité. On note $\Omega = \Omega_1 \times \ldots \Omega_n$ et $P = \mu_1 \otimes \ldots \otimes \mu_n$. Pour $x = (x_1, \ldots, x_n) \in \Omega$ et $i \leq n$, on notera $\hat{x}_i = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$. Pour $f : \Omega \longrightarrow \mathbb{R}$ et $i \leq n$, on introduit $f_i : \Omega_i \longrightarrow \mathbb{R}$,*

$$f_i(x_i) := f_i^{\hat{x}_i}(x_i) := f(x_1, \ldots, x_n)$$

lorsque $\hat{x}_i$ est fixé. Alors, pour $f \geq 0$ on peut définir, pour $z \in \Omega$,

$$\begin{aligned}\mathrm{Ent}_{\mu_i}(f_i)(z) &:= \int f_i^{\hat{z}_i}(t) \log \frac{f_i^{\hat{z}_i}(t)}{\int f_i^{\hat{z}_i}\, d\mu_i}\, d\mu_i(t) \\ &= \int f(z_1, \ldots, z_{i-1}, z_i, z_{i+1}, z_n) \log \frac{f(z_1, \ldots, z_{i-1}, z_i, z_{i+1}, z_n)}{\int f(z_1, \ldots, z_{i-1}, s, z_{i+1}, z_n)\, d\mu_i(s)}\, d\mu_i(z_i),\end{aligned}$$

qui ne dépend donc que de $\hat{z}_i$ ($\mathrm{Ent}_{\mu_i}(f_i)(z) = \mathrm{Ent}_{\mu_i}(f_i)(\hat{z}_i)$ *: on a intégré $f(z)\log f(z)$ par rapport à $d\mu_i(z_i)$.*

L'inégalité de tensorisation de l'entropie dit que :

$$\mathrm{Ent}_P(f) \leq \sum_{i=1}^{n} \int \mathrm{Ent}_{\mu_i}(f_i)\, dP.$$

En utilisant la tensorisation de l'entropie, on obtient directement le résultat précédent sur la tensorisation de l'inégalité de log-Sobolev, puisque du coté gradient, on aura, sur un produit d'espaces euclidiens $\Omega_i = \mathbb{R}^{n_i}$, que $|\nabla f|^2 = \sum |\nabla_i f|^2$.

5.3. Lien avec la concentration

On commence par remarquer que l'inégalité de Poincaré permet d'établir une concentration exponentielle auteur de leur moyenne pour les fonctions lipschitziennes.

***Théorème* (Gromov-Milman).** — *Si μ satisfait une inégalité de Poincaré, alors pour toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ qui est 1-lipschitzienneon a*

$$\left\| f - \int f\, d\mu \right\|_{\psi_1(\mu)} \leq \sqrt{\mathrm{c_P}(\mu)}\, L.$$

Démonstration. — On peut supposer que $L = 1$ et on note $C = \mathrm{c_P}(\mu)$. On va établir que pour toute fonction L-lipschitzienne on a

$$\forall t \leq \frac{2}{\sqrt{C}}, \qquad \int e^{tf}\, d\mu \leq e^{t\int f\, d\mu} \psi(\sqrt{C}\, t/2), \qquad \text{où } \psi(s) := \prod_{k=0}^{\infty} \left(1 - \frac{s^2}{4^k}\right)^{-2^k}.$$

On remarque que pour établir cette inégalité, il suffit de considérer d'abord les fonctions lipschitziennes bornées (par convergence monotone et Fatou), de sorte que les quantités en question sont finies. On va bien sûr faire une méthode de Laplace. Posons

$$\alpha(t) = \int e^{tf}\, d\mu.$$

Si on applique l'inégalité de Poincaré à $g = e^{tf/2}$ on obtient

$$\alpha(t) - \alpha(t/2)^2 \leq C \cdot (t/2)^2 \cdot \alpha(t)$$

c'est à dire,

$$1 - \frac{C\,t^2}{4} > 0 \Longrightarrow \alpha(t) \leq \Big(1 - \frac{Ct^2}{4}\Big)^{-1} \alpha(t/2)^2.$$

On obtient l'inégalité voulue en remarquant que $\alpha(t/2^k)^{2^k}$ tend vers $e^{t\int f\, d\mu}$ lorsque $k \to +\infty$.

Pour conclure, on pourra montrer que pour $t \in [0,1[$ on a $\psi(t) \leq \frac{1+t}{1-t}$, ou du moins que $\psi(\frac{1}{4}) \leq 2$. Mais si on ne veut pas faire de calcul, on peut aussi dire que $\Psi(t)$ tend vers 1 lorsque t tend vers zero, et donc passe à gauche de 2 pour une certaine constante numérique t_0.

□

Le résultat suivant indique qu'une inégalité de log-Sobolev entraîne de la concentration gaussienne .

Théorème. — *Soit μ une probabilité vérifiant* LSI(ρ_0). *Alors, pour toute fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ 1-lipschitzienne on a :*

$$\forall \lambda \geq 0, \qquad \int e^{\lambda\left[f - \int f\, d\mu\right]}\, d\mu \leq e^{\lambda^2/2\rho_0}.$$

On a donc, pour toute fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ 1-lipschitzienne

$$\forall r \geq 0, \qquad \mu\left(\left\{x \in \mathbb{R}^n \,;\, f(x) \geq \int f\, d\mu + r\right\}\right) \leq e^{-\rho_0 r^2/2},$$

et avec une constante 2 pour la version symétrique (pour $\left|f - \int f\, d\mu\right| \geq r$).

Démonstration. — **(Méthode de Herbst)** On introduit, pour $\lambda \geq 0$,

$$H(\lambda) = \int e^{\lambda\left[f - \int f\, d\mu\right]}\, d\mu.$$

En appliquant (1) à $g = e^{\lambda f/2}$ et en utilisant que $|\nabla f| \leq 1$ (presque partout si f est seulement lipschitzienne) on obtient,

$$\lambda H'(\lambda) - H(\lambda)\log(H(\lambda)) \leq \frac{1}{2\rho_0}\lambda^2 H(\lambda),$$

soit encore, pour $\lambda > 0$,

$$\frac{1}{\lambda} \cdot \frac{H'(\lambda)}{H(\lambda)} - \frac{1}{\lambda^2} \cdot \log(H(\lambda)) \leq \frac{1}{2\rho_0},$$

c'est-à-dire $\dfrac{d}{d\lambda}\dfrac{\log H(\lambda)}{\lambda} \leq \dfrac{1}{2\rho_0}$. Par ailleurs, un DL en $\lambda = 0$ donne que $\frac{1}{\lambda}\log H(\lambda) \to \int f\, d\mu$ lorsque $\lambda \to 0$. Par conséquent, en intégrant l'inégalité précédente, on obtient

$$\frac{1}{\lambda} \cdot \log H(\lambda) - \int f\, d\mu \leq \frac{1}{2\rho_0} \cdot \lambda,$$

ce qui donne le résultat annoncé. □

On voit donc que si μ est une probabilité sur $\mathbb{R}^n$ qui satisfait une $LSI(\rho_0)$ avec $\int y\, d\mu(y) = 0$, alors pour toute direction $\theta \in S^{n-1}$ on a

$$\mu(\{x \in \mathbb{R}^n \ ; \ x \cdot \theta \geq r\}) \leq e^{-\rho_0 r^2/2}$$

et donc les queues sont sous-gaussiennes dans toutes les directions.

Cela montre par exemplequ'en dimension 1 la mesure $d\mu(t) = \frac{1}{2} e^{-|t|}\, dt$ ne satisfait pas une inégalité de log-sob. On verra qu'elle satisfait une inégalité de Poincaré, en tant que mesure log-concave (mais dans ce cas, on peut aussi le démontrer facilement de manière directe).

5.4. Inégalité isopérimétrique, constante de Cheeger et lien avec l'inégalité de Poincaré

Soit (M, d, μ) un espace métrique mesuré (sous-entendu μ est une mesure borélienne sur M). Pour un ensemble $A \subset M$ et $\varepsilon > 0$, on note A_ε l'ensemble des points à distance au plus ε de A. Ainsi, une notion raisonnable de "μ-mesure du bord de A" (avec A borélien) est donnée par

$$\mu^+(\partial A) := \liminf_{\varepsilon \to 0} \frac{\mu(A_\varepsilon) - \mu(A)}{\varepsilon}.$$

Le problème isopérimétrique consiste à étudier les ensembles qui, à mesure fixée, minimisent la mesure de bord. Sauf cas très particulier, on ne sait pas résoudre ce problème de manière exacte, mais on cherche à avoir une estimation du type

$$\mu^+(\partial A) \geq F(\mu(A))$$

pour une bonne fonction F sur $\mathbb{R}^+$. La meilleure fonction possible est par définition le "profil isopérimétrique" défini pour $a \geq 0$ par

$$I_\mu(a) = \inf\{\mu^+(\partial A) \ ; \ A \subset M, \ \mu(A) = a\}.$$

Dans la suite, on va étudier le cas où μ est une probabilité, auquel cas la mesure a d'un ensemble ensemble appartient à $[0, 1]$ et la fonction $F : [0, 1] \to \mathbb{R}^+$ sera prise symétrique par rapport à $1/2$ (tout ensemble raisonnable aura la même mesure de bord que son complémentaire).

Definition. — *Soit μ une probabilité borélienne sur un espace métrique (M, d). On dit que μ vérifie une inégalité de Cheeger de constante $h > 0$ si, pour tout $A \subset M$ on a*

$$\mu^+(\partial A) \geq h \min\{\mu(A), 1 - \mu(A)\},$$

ou de manière équivalente, si

$$I_\mu(a) \geq h \min\{a, 1 - a\}.$$

La meilleure constante h possible (i.e. la plus grande) s'appelle la constante de Cheeger et se note $h(\mu)$.

On remarque qu'il suffit de vérifier ces propriétés pour des ensembles de mesures appartenant à $[0, 1/2]$ auquel cas on souhaite avoir $\mu^+(\partial A) \geq h\mu(A)$. On souhaite donc que $\mu^+(\partial A)$ reste au moins comparable à la mesure de A (on ne veut pas qu'il soit trop petit, en $\mu(A)^2$ par exemple).

L'inégalité isopérimétrique de Cheeger peut s'exprimer par une inégalité de type "Poincaré L^1", comme suit :

***Proposition* (Forme fonctionnelle).** — *Soit μ une probabilité sur l'espace euclidien $\mathbb{R}^n$. Alors, pour toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ (localement) lipschitzienne μ-intégrable on a*

$$h(\mu) \int |f - m_f| \, d\mu \leq \int |\nabla f| \, d\mu,$$

où m_f désigne une médiane de f par rapport à μ.

Remarque. — *Il s'agit en fait d'une équivalence. La réciproque est facile et s'obtient en prenant $f = 1_A$. Plus précisément, si on a l'inégalité fonctionnelle ci-dessus, alors on peut approcher 1_A par des fonctions régulières ou lipschitziennes f_ε de sorte que $\int |\nabla var| \, d\mu \longrightarrow \mu^+(\partial A)$. Par exemple, on peut prendre la suite de fonction 1-lipschitzienne*

$$f_\varepsilon(x) = \big(1 - \frac{1}{\varepsilon} d(x, A)\big)_+$$

pour laquelle on a $|\nabla f_\varepsilon| \geq \frac{1}{\varepsilon} 1_{A_\varepsilon \setminus A}$ et donc l'inégalité $\int |\nabla f_\varepsilon| \, d\mu \geq \frac{1}{\varepsilon} \mu(A_\varepsilon \setminus A)$ qui nous suffit. Par ailleurs, si A est un ensemble de mesure inférieure à $\frac{1}{2}$, alors 0 est une médiane de 1_A , et on trouve bien que $h(\mu)\, \mu(A) \geq \mu^+(\partial A)$.

En toute généralité, ce genre de déduction repose sur la formule de la co-aire. De manière à ne pas entrer dans des questions trop techniques, nous nous contenterons du résultat partiel suivant, qui est vrai dans un cadre général (y compris discret) à condition d'avoir une bonne notion de "norme de gardien".

Lemme. — *Pour f une fonction lipschitzienne sur $\mathbb{R}^n$ et μ une probabilité borélienne, on a*

$$\int |\nabla f| \, d\mu \geq \int_{-\infty}^{+\infty} \mu^+\big(\partial \{f > t\}\big) \, dt.$$

Démonstration. — On commence par supposer que f est minorée. Dans ce cas, quitte à rajouter une constante, on peut supposer que f est positive. Introduisons, pour $h > 0$, la fonction

$$f_h(x) := \sup_{|y-x| \leq h} f(y)$$

et posons $A[t] := \{x \in \mathbb{R}^n \; ; \; f(x) > t\}$. On a alors

$$\{f_h > t\} = A[t]_h$$

et par conséquent

$$\int \frac{1}{h}(f_h - f) \, d\mu = \int_0^\infty \frac{1}{h}\big(\mu(A[t]_h) - \mu(A[t])\big) \, dt.$$

La fonction dans l'intégrale de gauche est bornée (f lipschitzienne) et tend presque partout vers $|\nabla f|$ lorsque $h \to 0$, puisque, en un point x_0 où f est dérivable on a

$$f_h(x_0) = \sup_{|u| \leq 1} f(x_0 + hu) = \sup_{|u| \leq 1} \big[f(x_0) + h \nabla f(x_0) \cdot u + o(h)\big],$$

où le o est uniforme en u. Donc par convergence dominée, le terme de gauche tend vers $\int |\nabla f| \, d\mu$. Pour le terme de droite, on utilise le lemme de Fatou :

$$\liminf_{h \to 0} \int_0^\infty \frac{1}{h}\big(\mu(A[t]_h) - \mu(A[t])\big) \, dt \geq \int_0^\infty \liminf_{h \to 0} \frac{1}{h}\big(\mu(A[t]_h) - \mu(A[t])\big) \, dt = \int_0^{+\infty} \mu^+\big(\partial \{A[t]\}\big) \, dt.$$

Pour f générale, on approche par $f_N(x) = \max\{f(x), -N\}$ (exercice). □

Démonstration de la Proposition. — On peut supposer que f a 0 comme médiane. On a, par le lemme précédent,

$$\begin{aligned}\int |\nabla f|\, d\mu \geq \int_{-\infty}^{\infty} \mu^+(\partial\{f > t\})\, dt \geq h(\mu) \int_{-\infty}^{\infty} \min\{\mu(\{f > t\}), \mu(\{f \leq t\})\}\, dt \\ = h(\mu) \int_{-\infty}^{0} \mu(\{f \leq t\})\, dt + h(\mu) \int_{0}^{+\infty} \mu(\{f > t\})\, dt \\ = h(\mu) \int_{0}^{+\infty} \big[\mu(\{f \leq -t\}) + \mu(\{f > t\})\big]\, dt = h(\mu) \int_{\mathbb{R}^n} |f|\, d\mu.\end{aligned}$$

□

On peut remonter de l'inégalité $L^1 - L^1$ à une inégalité $L^2 - L^2$.

***Théorème* (Inégalité de Cheeger).** — *Soit μ une probabilité borélienne sur $\mathbb{R}^n$ satisfaisant une inégalité isopérimétrique de Cheeger. Alors μ satisfait une inégalité de Poincaré, avec*

$$\lambda_1(\mu) = \mathrm{c_P}(\mu)^{-1} \geq \frac{1}{h(\mu)^2}.$$

En d'autres termes, pour toute fonction f (localement) lipschitzienne on a

$$\mathrm{Var}_\mu(f) \leq \frac{4}{h(\mu)^2} \int |\nabla f|^2\, d\mu.$$

Démonstration. — Par définition de la variance, il suffit de montrer que pour une fonction f de classe C^1 de μ-médiane nulle, on a

$$\int f^2\, d\mu \leq \frac{4}{h(\mu)^2} \int |\nabla f|^2\, d\mu.$$

Introduisons la fonction

$$g(x) = \varepsilon(f(x)) f(x)^2$$

où $\varepsilon(a) = \pm 1$ désigne le signe du réel a (on peut convenir que $\varepsilon(0) = 1$). La fonction g est aussi de classe C^1 et de μ-médiane nulle. De plus, on a

$$|\nabla g| = 2|f||\nabla f|.$$

Ceci est évident si $f(x) \neq 0$. Si $f(x_0) = 0$, on remarque que $\nabla(f^2)(x_0) = 0$ de sorte que $|g(x_0 + h)| = f^2(x_0 + h) = o(h)$ ce qui entraine que le gradient de g est nul en x_0.

En appliquant l'inégalité fonctionnelle à g qui est de μ-médiane nulle on trouve

$$\int f^2\, d\mu = \int |g|\, d\mu \leq \frac{1}{h(\mu)} \int |\nabla g|\, d\mu = \frac{2}{h(\mu)} \int |f||\nabla f|\, d\mu.$$

On conclut en appliquant l'inégalité de Cauchy-Schwarz au membre de droite. □

5.5. Inégalités de Cheeger et de Poincaré pour les mesures log-concaves

Les résultats suivants sont dûs à S. Bobkov.

Proposition. — *Soit μ une probabilité log-concave sur $\mathbb{R}^n$ et F une fonction localement lipschitzienne à valeurs dans $[0,1]$. Alors, pour tout $r \geq 1$ et $x_0 \in \mathbb{R}^n$ on a*

$$2r \int |\nabla F| \, d\mu \geq \mathrm{Ent}_\mu(F) + \mathrm{Ent}_\mu(1-F) + \log \mu\big(B(x_0, r)\big)$$

où $B(x_0, r)$ désigne la boule euclidienne centrée en x_0 de rayon r.

Démonstration. — On peut supposer par approximation que F est de classe C^2 et à support compact. On va appliquer l'inégalité de Prékopa-Leindler avec la mesure log-concave μ et les fonctions

$$f(x) = F(x)^{1/(1-t)}, \quad g(y) = 1_{B(x_0,r)}(y), \quad , \quad h(z) = F_t(z)$$

avec

$$F_t(z) = \sup_{y \in B(x_0,r)} F(\frac{1}{1-t} z - \frac{t}{1-t} y) = \sup_{|u| \leq r} F(z - \frac{t}{1-t}(z - u - x_0)),$$

de sorte que $F_t((1-t)x + ty) \geq f(x)^{1-t} 1_{B(x_0,r)}(y)^t$ et donc

$$\int F_t \, d\mu \geq \Big(\int F^{1/(1-t)} \, d\mu \Big)^{1-t} \mu(B(x_0, r))^t.$$

On remarque qu'on a, lorsque $t \to 0$,

$$F_t(z) = F(z) + t\big[r|\nabla F(z)| + \nabla F(z) \cdot (z - x_0)\big] + O(t^2)$$

uniformément en z, et donc

$$\int F_t \, d\mu = \int F \, d\mu + tr \int |\nabla f| \, d\mu + t \int \nabla F(z) \cdot (z - x_0) \, d\mu(z) + O(t^2).$$

Par ailleurs,

$$\Big(\int F^{1/(1-t)} \, d\mu \Big)^{1-t} = \exp\Big[(1-t) \log \int \exp[\frac{1}{1-t} \log F]\Big] = \int F \, d\mu + t \, \mathrm{Ent}_\mu(F) + O(t^2).$$

On trouve donc, en comparant les termes d'ordre t, que

$$r \int |\nabla f| \, d\mu + \int \nabla F(z) \cdot (z - x_0) \, d\mu(z) \geq \mathrm{Ent}_\mu(F) + \Big(\int F \, d\mu \Big) \log \mu\big(B(x_0, r)\big).$$

En appliquant l'inégalité à $1 - F$ puis en sommant, on obtient le résultat voulu. □

Théorème. — *Soit μ une probabilité log-concave sur $\mathbb{R}^n$ ayant son barycentre en zero, i.e. $\int y \, d\mu(y) = 0$. Alors μ satisfait une inégalité isopérimétrique de Cheeger avec*

$$h(\mu) \geq \frac{c}{\sqrt{\int_{\mathbb{R}^n} |x|^2 \, d\mu(x)}}$$

pour une constante numérique universelle $c > 0$. En particulier, μ satisfait une inégalité de Poincaré avec $\mathrm{c_P}(\mu) \leq \tilde{c} \int |x|^2 \, d\mu(x)$.

Sans hypothèse sur le barycentre, on peut écrire $h(\mu) \geq \dfrac{c}{\sqrt{\int_{\mathbb{R}^n} \left| x - \int y \, d\mu(y) \right|^2 d\mu(x)}}$.

Démonstration. — On se donne un ensemble A avec $p := \mu(A) \leq \frac{1}{2}$. En approchant 1_A par des fonctions lipschitziennes à valeurs dans $[0,1]$ (avec par exemple $f_\varepsilon(x) = \big(1 - \frac{1}{\varepsilon} d(x,A)\big)_+$), on trouve, en appliquant l'inégalité précédente avec $x_0 = 0$,

$$2r\mu^+(\partial A) \geq p\log(1/p) + (1-p)\log(1/(1-p)) + \log\mu(rB_2^n).$$

Fixons $r_0 > 0$ de sorte que $\mu(r_0 B_2^n) = \frac{2}{3}$, par exemple. En appliquant l'inégalité ci-dessus avec tr_0, on obtient que pour tout $t > 0$,

$$2r_0 t\, \mu^+(\partial A) \geq p\log(1/p) + (1-p)\log(1/(1-p)) + \log\mu(t(r_0B_2^n)).$$

Mais par l'inégalité de Borell pour la mesure log-concave μ et le convexe symétrique $r_0B_2^n$, on a que pour $t \geq 1$,

$$\mu(t(r_0B_2^n)) \geq 1 - e^{-t/3}$$

On veut choisir t pas trop grand mais satisfaisant

$$(1-p)\log(1/(1-p)) + \log\mu(t(r_0B_2^n)) \geq 0.$$

Pour cela, on remarque que d'une part,

$$(1-p)\log(1/(1-p)) \geq \frac{1}{2}\log(1/(1-p)) \geq \log(1/(1-p/2))$$

et d'autre part que si $t \geq 4$, alors $1 - e^{-t/3} \geq 1 - \frac{1}{2}e^{-t/8}$. Ainsi, si on choisit t tel que $e^{-t/8} = p$, c'est-à-dire

$$t := 8\log(1/p)$$

qui est plus grand que $8\log(2) \geq 4$, on a l'inégalité voulue et donc

$$2r_0 t\, \mu^+(\partial A) \geq p\log(1/p),$$

soit encore

$$16 r_0\, \mu^+(A) \geq p.$$

On a donc établi une inégalité de Cheeger, avec

$$h(\mu) \geq \frac{1}{16r_0}.$$

Reste à estimer r_0. Par l'inégalité de Markov on a que

$$\mu(\{x \in \mathbb{R}^n\,;|x| \geq \sqrt{3}M_2\}) \leq \frac{1}{3}$$

où $M_2 = \|\,|x|\,\|_{L^2(\mu)} = \sqrt{\int |x|^2\,d\mu(x)}$, et donc on a $r_0 \leq \sqrt{3}M_2$. □

5.6. Inégalité de Sobolev logarithmique de type gaussien

Théorème. — *Soit μ une probabilité sur $\mathbb{R}^n$ de densité e^{-V} avec V uniformément convexe : il existe $\lambda > 0$ tel que $D^2V(x) \geq \lambda\mathrm{Id}$ pour tout $x \in \mathbb{R}^n$. Alors μ satisfait $LSI(\lambda)$.*

L'exemple principal est la mesure gaussienne.

Démonstration. — Sans perte de généralité, on peut supposer que $\lambda = 1$ (OK ?). On va établir l'inégalité de log-sob pour une fonction strictement positive que l'on écrit sous la forme e^f, avec $f : \mathbb{R}^n \to \mathbb{R}$.

On applique la version à poids de Prékopa-Leindler. Pour notre f, si g vérifie, pour tout $x, y \in \mathbb{R}^n$

$$e^{f((1-t)x+ty)} \geq e^{-t(1-t)|y-x|^2/2}\big(e^{g(x)/(1-t)}\big)^{1-t} \times 1^t$$

alors $\int e^f\, d\mu \geq \big(\int e^{g/(1-t)}\, d\mu\big)^{1-t}$.

On va introduire la régularisation par le "semi-groupe" d'Hamilton-Jacobi. Pour $\varphi : \mathbb{R}^n \to \mathbb{R}$ et $s > 0$, on définit

$$Q_s(\varphi)(x) = \inf_{y \in \mathbb{R}^n} \big\{\varphi(z) + \frac{1}{2s}\,|z - x|^2\big\}.$$

On remarque que le meilleur choix possible de g ci-dessus correspond à

$$g(x) = \inf_y \big\{f((1-t)x + ty) + t(1-t)|y - x|^2/2\big\} = Q_{t/(1-t)}(f).$$

On a donc

$$\int e^f\, d\mu \geq \Big(\int e^{\frac{1}{1-t}Q_{t/(1-t)}(f)}\, d\mu\Big)^{1/(1-t)}.$$

On va supposer que f est C^2 et qu'elle est bornée, ainsi que ses dérivées d'ordre ≤ 2. On voit alors que $Q_s(f) \to f$ lorsque $s \to 0$. On peut dire mieux :

$$Q_s(f) = f - \frac{s}{2}|\nabla f|^2 + O(s^2)$$

uniformément sur $\mathbb{R}^n$. On va se contenter d'établir une inégalité, qui suffit à notre propos. On a

$$Q_s(f)(x) - f(x) = s \inf_{u \in \mathbb{R}^n} \big\{\frac{1}{s}\big(f(x + su) - f(x)\big) + \frac{1}{2}|u|^2\big\}.$$

Pour x fixé, prenons $u_0 = u_0(x, s)$ où l'inf est atteint. On a $u_0 = -\nabla f(x + su_0)$ et donc $|u_0(x, s)|$ reste borné. Il existe donc une boule B telle que, disons pour $s \geq 1$,

$$Q_s(f)(x) - f(x) = s \inf_{u \in B} \big\{\frac{1}{s}\big(f(x + su) - f(x)\big) + \frac{1}{2}|u|^2\big\}.$$

Par ailleurs, il existe une constante M telle que pour tout x, tout $u \in B$ et tout $s \leq 1$,

$$f(x + su) \geq f(x) + s\nabla f(x) \cdot u - s^2 M$$

et par conséquent

$$Q_s(f)(x) - f(x) \geq s \inf_{u \in B} \Big\{\nabla f(x) \cdot u + \frac{1}{2}|u|^2\Big\} + O(s^2) \geq -s\frac{1}{2}|\nabla f(x)|^2 + O(s^2),$$

comme souhaité.

On a donc

$$\Big(\int e^f\, d\mu\Big)^{1-t} \geq \int e^{(1+t)\left[f - \frac{t}{2}|\nabla f|^2 + O(t^2)\right]}\, d\mu,$$

ce qui donne, à l'ordre 1,

$$\Big(\int e^f\, d\mu\Big) \log\Big(\int e^f\, d\mu\Big) \geq \int e^f\,[f - \frac{1}{2}|\nabla f|^2]\, d\mu,$$

soit encore

$$\mathrm{Ent}_\mu(e^f) \leq \frac{1}{2}\int |\nabla f|^2 e^f\, d\mu = \frac{1}{2}\int \frac{|\nabla e^f|^2}{e^f}\, d\mu.$$

Il reste à étendre le résultat à des fonctions plus générales. $\square$