HAL open science

Inégalités fonctionnelles liées aux formes de Dirichlet. De l'isopérimétrie aux inégalités de Sobolev.

Pierre Fougères

▶ To cite this version:

Pierre Fougères. Inégalités fonctionnelles liées aux formes de Dirichlet. De l'isopérimétrie aux inégalités de Sobolev.. Mathématiques [math]. Université Paul Sabatier - Toulouse III, 2002. Français. NNT : . tel-00002624

HAL Id: tel-00002624

https://theses.hal.science/tel-00002624v1

Submitted on 25 Mar 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

présentée en vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ PAUL SABATIER TOULOUSE III

Discipline : Mathématiques
Spécialité : Probabilités

par **Pierre Fougères**

Sous la direction de **Dominique Bakry**

Inégalités fonctionnelles liées aux formes de Dirichlet. De l'isopérimétrie aux inégalités de Sobolev.

Soutenue le 18 Octobre 2002 devant le jury composé de
Messieurs les Professeurs :

D.	Bakry	Université Paul Sabatier	Directeur
F.	Barthe	Université de Marne-La-Vallée	Examinateur
M.	Ledoux	Université Paul Sabatier	Examinateur
P.	Mathieu	Université de Provence	Rapporteur
L.	Miclo	Université Paul Sabatier	Examinateur
B.	Zegarlinski	Imperial College, London, UK	Rapporteur

Laboratoire de Statistique et Probabilités
UMR CNRS 5883, Université Paul Sabatier, Toulouse III

REMERCIEMENTS

Je ne sais comment exprimer ma reconnaissance envers mon directeur de thèse Dominique BAKRY pour le soutien qu'il m'a apporté durant ces années. Ses idées florissantes et sa facilité à poser des questions simples et intéressantes m'ont fortement guidé et m'ont fait aimer ce métier difficile qu'est la recherche. Je mesure la chance qui m'a été offerte de profiter de son expérience et de côtoyer son enthousiasme communicatif. En particulier, les derniers temps de ma thèse ont vraiment été passionnants.

Pierre MATHIEU et Boguslaw ZEGARLINSKI m'ont fait la joie d'être rapporteurs de ma thèse. Je tiens à les remercier d'avoir accepté cette tâche et d'avoir apporté leurs regards sur mon travail. J'ai apprécié les remarques sincères et amicales de Pierre MATHIEU auxquelles j'ai essayé de répondre au mieux. J'espère que ma collaboration présente avec Boguslaw ZEGARLINSKI sera fructueuse et le remercie de m'avoir invité à Londres.

Merci à tous les membres du jury d'avoir fait le voyage pour venir m'écouter présenter mes travaux et également à ceux dont le voyage était plus court, pour leur participation et le temps passé à lire ma thèse ; Franck BARTHE que j'ai connu il y a quelques années lors d'un colloque qu'il (co)organisait et Laurent MICLO dont le naturel et l'aisance à poser des questions font plaisir à voir...

Michel LEDOUX a toujours été d'une aide discrète et efficace. La clarté de ses écrits est fort agréable pour ses lecteurs, dont je fais partie. Je profite de l'occasion qui m'est donnée de lui exprimer ma gratitude. Avec Dominique BAKRY, il nous a offert la chance, à nous tous les *logsob*'s, d'entreprendre ce projet qui, bien que me dépassant au départ, m'a beaucoup appris, tant humainement que scientifiquement.

Cécile, Cyril, Djélil, Florent, Greg, Ivan et Sébastien, merci à vous pour cette belle expérience. C'était un plaisir de travailler avec vous. Florent, le souvenir de nos discussions dans le bureau 147 me sera toujours agréable. Djé, merci pour ta profondeur d'esprit et notre amitié tumultueuse, et plus techniquement pour tes conseils informatiques précieux.

J'ai une pensée également pour les doctorants du LSP (et des autres labos) que j'ai côtoyés et avec qui j'ai partagé de bons moments. En particulier, les différents bureaux 147.

Merci à Christophe Brouttelande pour ses explications sur l'Analyse sur les variétés qui m'ont été fort utiles.

Et puis, ces années de thèse ne seraient rien sans la vie qui va avec. Alors...
Merci à Anne-Laure et Ivan de leur soutien aux moments difficiles et de leur joie de vivre. Merci, Ivan, de m'avoir fait partager une autre passion si enrichissante. Ces moments de liberté m'ont été d'un grand secours.

Merci à mes amis, la clique bisontine et les autres, qui au long de ces années ont toujours été là pour entretenir mon enthousiasme ou me remonter le moral quand il le fallait.

Enfin, mille pensées pour mon Florent à moi dont la présence a ensoleillé ma vie et m'a aidé à continuer. J'aimerais qu'il sache un jour combien je lui en suis reconnaissant.

TABLE DES MATIÈRES

Remerciements 3

Introduction générale 7
0.1. Semi-groupes et générateurs de MARKOV 7
0.2. Inégalités fonctionnelles $\mathbf{L}^2$ 11
0.3. Inégalités fonctionnelles et contrôles du semi-groupe 15
0.4. Le critère de courbure-dimension 19
0.5. Inégalités fonctionnelles $\mathbf{L}^1$ 25
0.6. Organisation de la thèse 27

1. Hypercontractivité et isopérimétrie gaussienne 29
1.1. Introduction 29
1.2. Isopérimétrie et hypercontractivité 31
1.3. Inégalité isopérimétrique pour les spins continus 45
1.4. Spins discrets et isopérimétrie 53

2. Spectral gap for log-concave measures 65
2.1. Introduction 65
2.2. A simple criterion for POINCARÉ inequality 68
2.3. Variational proof of the main result 69
2.4. CHEEGER inequality and BOBKOV's argument 92

3. Liftings of a DIRICHLET form 99
3.1. Introduction 99
3.2. A simple example in spherical geometry 102
3.3. Weighted SOBOLEV inequality and lifting of a DIRICHLET form 108
3.4. Proof of the main result 116

A. Concentration de la mesure 123
A.1. Introduction 123
A.2. La concentration de la mesure 124
A.3. Concentration via le critère de courbure 130

A.4. Concentration et inégalité de SOBOLEV logarithmique 133
A.5. L'argument de HERBST inverse 138
A.6. Notes 145

Bibliographie 147

INTRODUCTION GÉNÉRALE

Certaines inégalités fonctionnelles jouent un rôle important dans l'étude des semi-groupes de MARKOV. Elles sont fortement liées aux propriétés ergodiques de ces semi-groupes et interviennent dans l'étude de la mesure invariante. Lorsque le semi-groupe est naturel et sous-jacent (comme le semi-groupe de la chaleur sur une variété riemannienne compacte), elles fournissent des informations géométriques et analytiques sur l'espace (bornes sur le diamètre, contrôle des solutions d'équations aux dérivées partielles, etc.).

Notre objectif dans cette introduction est, après avoir décrit le cadre général des semi-groupes de MARKOV dans lequel se situe ce travail, de présenter les inégalités fonctionnelles qui nous intéresserons dans la suite. Afin de souligner leur importance et les liens qui existent entre elles, nous décrivons des résultats maintenant classiques sur les contrôles qu'elles induisent sur le semi-groupe. Cela sera l'occasion en particulier d'introduire une famille plus vaste d'inégalités. Bien que ces dernières ne soient pas l'objet direct de cette thèse, leur importance est un motif suffisant à les évoquer. Puis, nous présentons le critère de courbure-dimension qui permet d'obtenir des inégalités fonctionnelles avec les bonnes constantes dans les modèles correspondants. Ce critère fournit également des preuves fonctionnelles de résultats de comparaison sur le volume et le diamètre (théorème de BISHOP et de MYERS) qui s'étendent à des théorèmes de comparaisons sur les valeurs propres des générateurs de MARKOV. Nous décrivons ensuite les inégalités $\mathbf{L}^1$ (ou isopérimétriques) correspondant aux inégalités $\mathbf{L}^2$ vues précédemment. Nous terminons par une brève description des résultats obtenus dans la thèse et de l'organisation des chapitres suivants.

0.1. Semi-groupes et générateurs de MARKOV

Nous allons maintenant décrire brièvement le cadre général dans lequel se situe ce travail. Soit $(\mathbb{E}, \mathcal{F}, \mu)$ un espace mesuré.

Définitions et construction. — Un *semi-groupe* de MARKOV est une famille $(\mathbf{P}_t)_{t\geqslant 0}$ d'opérateurs bornés de $\mathbf{L}^2(\mu)$ telle que

(i) $\mathbf{P}_0 = \mathrm{Id}$,

(ii) $\mathbf{P}_t \circ \mathbf{P}_s = \mathbf{P}_{t+s}$,

(iii) pour toute fonction $f \in \mathbf{L}^2(\mu)$, $\mathbf{P}_t f \to f$ dans $\mathbf{L}^2(\mu)$, quand t tend vers 0,

(iv) $\mathbf{P}_t 1 = 1$, pour tout $t \geqslant 0$ (propriété de MARKOV),

(v) $\mathbf{P}_t f \geqslant 0$ si f est une fonction positive.

Donnons dès maintenant deux exemples. Sur $\mathbb{R}^n$ muni de la mesure de LEBESGUE, le semi-groupe de la chaleur admet la représentation

$$\mathbf{P}_t f(x) = \int_{\mathbb{R}^n} f(y) e^{-\frac{|x-y|^2}{4t}} \frac{dy}{(4\pi t)^{n/2}}, \tag{1}$$

pour toute fonction f de classe $\mathcal{C}^\infty$ à dérivées bornées. Un second exemple fondamental est celui du semi-groupe d'ORNSTEIN-UHLENBECK. Il est défini sur $\mathbb{R}^n$ muni de la mesure gaussienne canonique $\gamma_n(dx) = \exp\left(-|x|^2/2\right) dx/(2\,\pi)^{n/2}$ par la représentation

$$\mathbf{P}_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2\,t}}y)\,\gamma_n(dy) = \int_{\mathbb{R}^n} f(z) p_t(x, dz),$$

où les noyaux gaussiens $p_t(x, dz)$ de moyenne $e^{-t}\,x$ et de covariance $(1-e^{-2\,t})\,I_n$ constituent un *chemin continu* entre les mesures δ_x et γ_n. Il est rare qu'une telle représentation à l'aide d'un noyau de probabilités soit explicite. En général, le semi-groupe est défini à l'aide de son *générateur infinitésimal*.

Sur l'ensemble $\mathcal{D}(\mathbf{L}) = \{f \in \mathbf{L}^2(\mu) \,:\, \lim_{t\to 0} (\mathbf{P}_t f - f)/t$ existe dans $\mathbf{L}^2(\mu)\}$, on définit l'opérateur non borné $\mathbf{L}f \stackrel{\text{def.}}{=} \lim_{t\to 0} (\mathbf{P}_t f - f)/t$. $\mathbf{L}$ est dit *générateur infinitésimal* du semi-groupe. $\mathcal{D}(\mathbf{L})$ s'appelle le domaine de $\mathbf{L}$. C'est un sous-espace dense de $\mathbf{L}^2(\mu)$ stable par tous les opérateurs $\mathbf{P}_t$. De plus, pour toute fonction f du domaine $\mathcal{D}(\mathbf{L})$, $\mathbf{P}_t f$ est dérivable dans $\mathbf{L}^2(\mu)$ et

$$\frac{\partial}{\partial t}\mathbf{P}_t f = \mathbf{P}_t \mathbf{L} f = \mathbf{L}\mathbf{P}_t f.$$

Le théorème de HILLE-YOSIDA permet de reconstruire le semi-groupe $(\mathbf{P}_t)$ à l'aide du générateur $\mathbf{L}$ (voir [Yos80]). Un opérateur linéaire $\mathbf{L}$ de domaine $\mathcal{D}(\mathbf{L})$ est le générateur infinitésimal d'un semi-groupe continu de contractions ssi $\mathcal{D}(\mathbf{L})$ est dense, L est fermé, $\lambda\, I - \mathbf{L}$ est inversible de $\mathcal{D}(\mathbf{L})$ dans $\mathbf{L}^2$ et la résolvante $(\lambda\, I - \mathbf{L})^{-1}$ est un opérateur borné de norme inférieure à λ^{-1}. Les conditions supplémentaires assurant que le semi-groupe est de MARKOV sont, d'une part, que la fonction $\mathbb{1}$ est dans le domaine et $\mathbf{L}\mathbb{1} = 0$ et, d'autre part, que la résolvante préserve la positivité (voir [GZ00]).

En pratique, l'opérateur, toujours noté $\mathbf{L}$, qui permet de définir le semi-groupe n'est pas exactement son générateur infinitésimal mais une restriction à un sous-ensemble $\mathcal{D}$ du domaine $\mathcal{D}(\mathbf{L})$. Un tel opérateur est nécessairement fermable. Précisément, un opérateur $\mathbf{L}$ de domaine $\mathcal{D}$ est fermable et sa fermeture est le générateur d'un semi-groupe de contractions ssi $\mathcal{D}$ est dense dans $\mathbf{L}^2$, $\mathbf{L}$ est dissipatif (i.e. $\|\lambda f - \mathbf{L}f\|_2 \geqslant \lambda \|f\|_2$, pour toute $f \in \mathcal{D}$) et l'image $(\lambda I - \mathbf{L})(\mathcal{D})$ est dense dans $\mathbf{L}^2$ pour un $\lambda > 0$ (voir [EK86]). Cette dernière propriété constitue la difficulté majeure pour définir rigoureusement un semi-groupe. Elle assure que le domaine considéré $\mathcal{D}$ est assez *gros* pour que la solution $u(t,x) = \mathbf{P}_t f(x)$ de l'équation d'évolution

$$\left\{ \begin{array}{rcl} \dfrac{\partial}{\partial t} u(t,x) & = & \mathbf{L}u(t,x) \quad \forall (t,x) \in \mathbb{R}_*^+ \times \mathbb{E} \\ u(0,x) & = & f(x), \end{array} \right. \tag{2}$$

où $f \in \mathcal{D}$, admette une unique extension aux fonctions de $\mathbf{L}^2(\mu)$.

L'exemple fondamental soulignant ce dernier point est celui des semi-groupes des mouvements browniens sur $[0,1]$ tué et réfléchi au bord. La donnée de l'opérateur $\mathbf{L} = d^2/dx^2$ sur les fonctions $\mathcal{C}^\infty$ à support compact dans $]0,1[$ ne suffit pas à déterminer un semi-groupe de Markov unique sur $\mathbf{L}^2([0,1], dx)$. Si l'on impose aux fonctions d'être nulles au bord (conditions de Dirichlet) , le semi-groupe est le mouvement brownien tué au bord. Les conditions de Neumann (fonctions à dérivées nulles au bord) mènent au semi-groupe du mouvement brownien réfléchi au bord (voir [RS78, Bak94]).

Nous nous placerons de plus dans la suite sous l'hypothèse d'*algèbre standard* décrite dans [Bak94] ou [ABC+00]. C'est une hypothèse de régularité portant sur le domaine de $\mathbf{L}$. On suppose que $\mathcal{D}(\mathbf{L})$ contient une algèbre de fonctions bornées $\mathcal{A}$ comme sous-espace dense (pour la norme $\|f\|_{\mathcal{D}(\mathbf{L})} = \|f\|_2 + \|\mathbf{L}f\|_2$), stable par $\mathbf{L}$ et par l'action des fonctions $\mathcal{C}^\infty$ nulles en 0. $\mathcal{A}$ est dense dans tous les espaces $\mathbf{L}^p(\mu)$, $p \geqslant 1$ et contient les constantes. $\mathcal{A}$ est de plus supposée satisfaire à une condition technique de convergence pour la topologie du domaine (voir [Bak94]). Ces conditions sont bien sûr remplies dans les exemples simples que nous allons présenter ci-dessous et qui constitueront l'essentiel des semi-groupes que nous aborderons.

Avant cela, définissons l'*opérateur carré du champ* associé à $\mathbf{L}$.

Définition 0.1.1. — *On appelle opérateur* carré du champ *la forme bilinéaire symétrique suivante : pour f et g dans $\mathcal{A}$,*

$$\mathbf{\Gamma}(f,g) \overset{\text{déf.}}{=} \frac{1}{2}(\mathbf{L}(fg) - f\mathbf{L}g - g\mathbf{L}f).$$

Nous noterons $\mathbf{\Gamma}(f)$ pour $\mathbf{\Gamma}(f,f)$.

Une caractérisation des semi-groupes markoviens parmi les semi-groupes continu de contractions a été donnée dans [Maz02] en termes de carré du champ. Si le semi-groupe admet une mesure invariante (voir plus loin), il est markovien (i.e. préserve la positivité et laisse $\mathbb{1}$ invariante) ssi la forme bilinéaire $\mathbf{\Gamma}$ est positive.

Les exemples de base sont les générateurs de Markov sur un ensemble fini ou les diffusions sur une variété compacte (ou tout au moins complète pour la distance riemannienne associée à l'opérateur). Les algèbres $\mathcal{A}$ sont alors respectivement l'algèbre de toutes les fonctions et l'algèbre des fonctions $\mathcal{C}_c^\infty$. Introduisons brièvement ces deux exemples (voir aussi [Bak94, Led00, ABC+00]).

Exemple 0.1.2. — *Soit $\mathbb{E}$ un ensemble fini. Le générateur $\mathbf{L}$ est alors donné par une matrice markovienne $(\mathbf{L}_{ij})_{i,j\in\mathbb{E}}$. Si f est une fonction sur $\mathbb{E}$,*

$$\mathbf{L}f(i) = \sum_{j\in\mathbb{E}} \mathbf{L}_{ij}\, f(j)$$

où, pour tout $j \neq i$, $\mathbf{L}_{ij} \geqslant 0$ et $\sum_{j\in\mathbb{E}} \mathbf{L}_{ij} = 0$. Le carré du champ s'écrit

$$\mathbf{\Gamma}(f,g)(i) = 1/2 \sum_{j\in\mathbb{E}} \mathbf{L}_{ij}\,(f(i)-f(j))\,(g(i)-g(j)).$$

Exemple 0.1.3. — *Soit M une variété compacte et connexe, et soit (g) une cométrique riemannienne, i.e. la donnée dans n'importe quel système de coordonnées et en chaque point $x \in M$, d'une matrice symétrique définie positive $(g^{ij})(x)$ dont les coefficients sont $\mathcal{C}^\infty$ en la variable x. On demande de plus des conditions de recollement pour que cet objet ne dépende pas des coordonnées choisies. $\mathbf{L}$ est un opérateur différentiel du second ordre sans terme constant*

$$\mathbf{L}f(x) = \sum_{ij} g^{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_i b^i(x)\frac{\partial f}{\partial x_i}(x).$$

Le carré du champ est alors

$$\mathbf{\Gamma}(f,g)(x) = \sum_{ij} g^{ij}(x)\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_j}(x).$$

Dans la métrique riemannienne (g_{ij}) donnée par l'inverse des matrices (g^{ij}), l'opérateur $\mathbf{L}$ s'écrit plus simplement

$$\mathbf{L} = \Delta + X,$$

où X est un champ de vecteurs. La terminologie carré du champ *s'explique par la formule*

$$\mathbf{\Gamma}(f) = |\nabla f|^2,$$

toujours dans la métrique associée à l'opérateur.

Les opérateurs de diffusion se distinguent des opérateurs sur un ensemble discret par la propriété suivante, dite de diffusion, qui joue un grand rôle dans de nombreuses méthodes pour montrer l'existence d'inégalités fonctionnelles, basées sur l'utilisation de semi-groupes.

Définition 0.1.4. — *On dit que* $\mathbf{L}$ *est un opérateur de* diffusion *si pour toute fonction* Φ, $\mathcal{C}^\infty$ *de* $\mathbb{R}^n$ *dans* $\mathbb{R}$, *et toute fonction* $F = (f_1, \ldots, f_n)$ *dans* $\mathcal{A}^n$,

$$\mathbf{L}(\Phi(F)) = \sum_i \frac{\partial \Phi(F)}{\partial x_i} \mathbf{L}(f_i) + \sum_{i,j} \frac{\partial^2 \Phi(F)}{\partial x_i \partial x_j} \mathbf{\Gamma}(f_i, f_j). \tag{3}$$

Nous utiliserons souvent le versant carré du champ de cette propriété pour une fonction d'une variable :

$$\mathbf{\Gamma}(\Phi(f), g) = \Phi'(f)\mathbf{\Gamma}(f, g).$$

Invariance et réversibilité. —

Définition 0.1.5. — *La mesure* μ *est dite* invariante *pour* $(\mathbf{P}_t)$ *si, pour tout* $t \geqslant 0$ *et toute* $f \in \mathcal{A}$,

$$\int \mathbf{P}_t f \, d\mu = \int f \, d\mu.$$

Cela revient à dire que

$$\int \mathbf{L} f \, d\mu = 0.$$

Si de plus le semi-groupe $\mathbf{P}_t$ *est symétrique dans* $\mathbf{L}^2(\mu)$, *ou de manière équivalente si* $\mathbf{L}$ *est autoadjoint, la mesure* μ *est dite* réversible.

Dans les exemples présentés précédemment, il existe toujours une mesure invariante. Dans le cas discret, le semi-groupe est réversible (par rapport à cette mesure μ) si, pour tout $i, j \in \mathbb{E}$,

$$\mu(i)\mathbf{L}_{ij} = \mu(j)\mathbf{L}_{ji}.$$

Pour un semi-groupe de diffusion, la condition de réversibilité porte sur le champ de vecteur X. $\mathbf{L}$ admet comme mesure réversible μ, de densité u par rapport à la mesure riemannienne, si (et seulement si)

$$X = \nabla \log(u).$$

En particulier, les semi-groupes de la chaleur et d'Ornstein-Uhlenbeck admettent respectivement les mesures de Lebesgue et gaussienne comme mesure réversible.

0.2. Inégalités fonctionnelles $\mathbf{L}^2$

Nous rencontrerons de nombreuses inégalités liées aux semi-groupes de Markov. En fait, ces inégalités ne font intervenir que la mesure invariante μ et le carré du champ $\mathbf{\Gamma}$. Nous parlerons indifféremment d'inégalité satisfaite par le semi-groupe ou par le couple $(\mu, \mathbf{\Gamma})$. Nous allons commencer par introduire la plus faible d'entre elles.

Inégalité de Poincaré. — On dit que $(\mu, \mathbf{\Gamma})$ satisfait à une inégalité de Poincaré s'il existe une constante C_2 telle que, pour toute fonction f de $\mathcal{A}$,

$$\int f^2\, d\mu - \left(\int f\, d\mu\right)^2 \leqslant C_2 \int \mathbf{\Gamma}(f)\, d\mu. \tag{4}$$

Le modèle associé à cette inégalité est la mesure exponentielle

$$\epsilon(dx) = \exp\left(-|x|\right) dx/2$$

sur $\mathbb{R}$, pour laquelle la meilleure constante C_2 vaut 4. Comme cette mesure est également le cas typique des mesures log-concaves sur $\mathbb{R}$, il n'est pas surprenant que ces mesures satisfassent une inégalité de Poincaré. En particulier, la mesure gaussienne $\gamma(dx) = \exp\left(-x^2/2\right) dx/\sqrt{2\,\pi}$ vérifie (4) avec $C_2 = 1$. L'inégalité (4) est encore appelée inégalité de *trou spectral* car la meilleure constante C_2 est l'inverse de la plus petite valeur propre non nulle de $-\mathbf{L}$. Ceci se reflète dans la propriété suivante décrivant la convergence exponentielle du semi-groupe vers sa mesure invariante.

Théorème 0.2.1. — *Les deux assertions suivantes sont équivalentes :*

(i) Il existe une constante $C_2 > 0$, telle que pour toute fonction $f \in \mathcal{A}$,

$$\|\mathbf{P}_t f - \mathbf{E}_\mu(f)\|_2 \leqslant e^{-t/C_2} \|f - \mathbf{E}_\mu(f)\|_2, \tag{5}$$

(ii) La mesure μ vérifie une inégalité de Poincaré *de constante $C_2 > 0$.*

Signalons que l'existence d'un trou spectral pour un semi-groupe $(\mathbf{P}_t)$ symétrique peut être caractérisée également à l'aide de la forme de Dirichlet associée $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$. Rappelons que $\mathcal{E}$ est la fermeture de la forme $\tilde{\mathcal{E}}(f, g) = -\int f\, \mathbf{L} g\, d\mu$, $f, g \in \mathcal{D}(\mathbf{L})$. La propriété équivalente au trou spectral est alors que, pour toute suite $(f_n)_n$ de fonctions centrées de $\mathcal{D}(\mathcal{E})$ dans la boule unité de $\mathbf{L}^2(\mu)$, si $\mathcal{E}(f_n, f_n) \to 0$ alors $f_n \to 0$ dans $\mathbf{L}^2(\mu)$. En n'exigeant qu'une convergence en probabilité de (f_n) vers 0, on aboutit à la notion de *trou spectral faible* et à la propriété UPI du semi-groupe (de l'anglais *uniformly positive improving*) qui lui est équivalente (voir [Aid98]). Cette propriété est liée à des études récentes concernant des caractérisations fines de l'existence d'un trou spectral et de l'ergodicité $\mathbf{L}^2$ du semi-groupe à une vitesse exponentielle (voir [GW00, Hin00] et les références incluses).

Inégalités de Sobolev. — On dit que le couple $(\mu, \mathbf{\Gamma})$ satisfait une inégalité de Sobolev de dimension $m > 2$ (ou d'exposant $p = 2m/(m-2) > 2$) de constantes A_2 et B_2 si, pour toute fonction $f \in \mathcal{A}$,

$$\left(\int f^{2m/(m-2)} \mu(dx)\right)^{(m-2)/m} \leqslant A_2 \int f^2 \mu(dx) + B_2 \int \mathbf{\Gamma}(f) \mu(dx). \tag{6}$$

Lorsque le couple $(\mu, \mathbf{\Gamma})$ est donné par le laplacien d'une variété riemmannienne, le problème de déterminer les meilleures constantes intervenant dans cette inégalité au moyen de paramètres géométriques est connu sous le nom de (A, B) *problem* (voir [Heb99] par exemple). Si la mesure μ est une probabilité, $A_2 \geqslant 1$.

De manière générale, lorsque dans une inégalité fonctionnelle il y a égalité pour les fonctions constantes, on dit que l'inégalité est *tendue*. Pour une inégalité de SOBOLEV, cela revient à dire que $A_2 = 1$. L'existence d'une inégalité tendue assure l'ergodicité du semi-groupe (au moins dans le cas symétrique). L'inégalité implique en effet que les seules fonctions telle que $\mathbf{\Gamma}(f) = 0$ sont les constantes. Alors les seules fonctions invariantes (i.e. telles que $\mathbf{L}f = 0$) sont les constantes (voir [Bak94], proposition 2.2). $\mathbf{P}_t(f)$ converge dans $\mathbf{L}^2$ vers la projection de f sur l'espace des fonctions invariantes et cette projection, étant constante, n'est autre que la moyenne de f.

En appliquant une inégalité de SOBOLEV tendue

$$\left(\int f^{2m/(m-2)}\mu(dx)\right)^{(m-2)/m} \leqslant \int f^2\mu(dx) + B_2 \int \mathbf{\Gamma}(f)\mu(dx)$$

aux fonctions $f = 1 + \varepsilon\, g$ avec $\varepsilon \to 0$, on obtient une inégalité de POINCARÉ de constante $C_2 = B_2/(p-2)$ où $p = 2m/(m-2)$. Inversement, s'il existe une inégalité de POINCARÉ, on peut *tendre* une inégalité de SOBOLEV (i.e. choisir $A_2 = 1$) quitte à remplacer B_2 par une constante plus grande $B_2' = (p-1)\,(A_2\, C_2 + B_2)$ (voir [Bak94]).

Le modèle de référence pour ce genre d'inégalités est la mesure uniforme (normalisée en une probabilité) sur la sphère de dimension m pour laquelle la constante B_2 optimale (pour $A_2 = 1$) est $B_2 = \frac{4}{m\,(m-2)}$. En projetant le laplacien sphérique sur un diamètre, on obtient un modèle unidimensionnel, l'opérateur de JACOBI symétrique sur $]-1,1[$ défini par

$$\mathbf{L}_n f(v) = (1-v^2)\, f''(v) - n\, v\, f'(v), \quad v \in]-1,1[.$$

Cet opérateur vérifiant la propriété de diffusion, on a

$$\mathbf{L}_n(\psi(u)) = \psi''(u)\, \mathbf{\Gamma}_n(u) + \psi'(u)\, \mathbf{L}_n(u)$$

et, en résolvant $\mathbf{\Gamma}_n(u) = 1$, on voit encore que l'opérateur défini sur $]-\pi/2, \pi/2[$ par

$$\mathbf{L}_{n-1,n}\psi(u) = \psi''(u) - (n-1)\, \tan(u)\, \psi'(u)$$

satisfait à la même inégalité de SOBOLEV.

On peut montrer (voir proposition 4.6.2 dans [ABC$^+$00]) que si μ vérifie une inégalité de SOBOLEV tendue de constante B_2 et de dimension n alors elle vérifie une inégalité de SOBOLEV tendue pour toutes les dimensions $m \geqslant n$ de constante $B_2\, n/m$. On peut vérifier aisément que, pour $f \geqslant 0$,

$$\frac{\|f\|_q - \|f\|_1}{q-1} \to \mathbf{Ent}_\mu(f) \overset{\text{déf.}}{=} \int f \log f\, d\mu - \left(\int f\, d\mu\right) \log\left(\int f\, d\mu\right),$$

lorsque $q \to 1$. Ainsi, lorsque la dimension m tend vers l'infini, l'information contenue dans les inégalités de SOBOLEV dégénère-t-elle en un facteur logarithmique. On obtient

$$\mathbf{Ent}_\mu(f^2) \leqslant B_2\, n/2 \int |\nabla f|^2\, d\mu.$$

Inégalité de SOBOLEV logarithmique. — Le couple $(\mu, \mathbf{\Gamma})$ vérifie une inégalité de SOBOLEV logarithmique de constante A_2 et B_2 si, pour toute fonction f de $\mathcal{A}$,

$$(7) \qquad \int f^2 \log f^2 \, d\mu - \left(\int f^2 \, d\mu\right) \log \left(\int f^2 \, d\mu\right) \leqslant A_2 \int f^2 d\mu + B_2 \int \mathbf{\Gamma}(f) d\mu.$$

De la même manière que pour les inégalités de SOBOLEV, la tension d'une inégalité de SOBOLEV logarithmique équivaut à l'existence d'une inégalité de POINCARÉ. Cela signifie ici que l'on peut choisir $A_2 = 0$.

Une inégalité de SOBOLEV logarithmique tendue s'écrit

$$\mathbf{Ent}_\mu(f^2) \leqslant B_2 \int \mathbf{\Gamma}(f) d\mu.$$

Le modèle de référence pour cette inégalité est la mesure gaussienne $\gamma(dx) = \exp(-x^2/2)\, dx/\sqrt{2\pi}$ pour laquelle elle est vérifiée avec la constante $B_2 = 2$. Par la même méthode qu'en partant d'une inégalité de SOBOLEV, on montre qu'une inégalité de SOBOLEV logarithmique tendue de constante B_2 implique une inégalité de POINCARÉ de constante $C_2 = B_2/2$. Mais cette inégalité est strictement plus forte en général qu'une inégalité de POINCARÉ. Par exemple, la mesure exponentielle $\epsilon(dx)$ typique pour POINCARÉ ne vérifie pas d'inégalité de SOBOLEV logarithmique (voir [ABC+00] par exemple).

Pour souligner l'importance du modèle gaussien pour l'inégalité de SOBOLEV logarithmique, mentionnons ici qu'elle implique la propriété de concentration gaussienne de la mesure i.e. que, pour toute $f \in \mathcal{A}$ telle que $\mathbf{\Gamma}(f) \leqslant 1$,

$$\mu\{|f - \mathbf{E}_\mu(f)| > r\} \leqslant C\, e^{-r^2/c}.$$

Cette implication, due à HERBST et présentée en toute généralité dans [Led99], sera explicitée dans l'annexe A.

Le caractère infini-dimensionnel de l'inégalité de SOBOLEV logarithmique, déjà souligné précédemment, lui assure la propriété suivante, dite de *tensorisation* : si elle est vérifiée pour les couples $(\mu_1, \mathbf{\Gamma}_1)$ et $(\mu_2, \mathbf{\Gamma}_2)$, elle est vérifiée sur l'espace produit pour le couple $(\mu_1 \otimes \mu_2, \mathbf{\Gamma}_1 + \mathbf{\Gamma}_2)$ (voir le chapitre 3 de [ABC+00]). L'inégalité de POINCARÉ satisfait aussi cette condition. En revanche, une inégalité de SOBOLEV (d'exposant optimal $p = 2m/(m-2)$) ne peut être tensorisée en général puisque sur une variété riemannienne compacte la dimension critique m est exactement la dimension n de l'espace. La propriété de tensorisation est à la base de l'utilisation de ce genre d'inégalité dans des problèmes de dimension infinie, comme nous le verrons plus tard.

Remarquons que les trois types d'inégalités que nous venons de présenter sont des cas particuliers de la famille d'inégalités

$$\forall f \in \mathcal{A}, \frac{1}{p-2}\left[\left(\int f^p \, d\mu\right)^2 - \int f^2 \, d\mu\right] \leqslant D_2 \int \mathbf{\Gamma}(f)\, d\mu.$$

L'inégalité de POINCARÉ correspond à $p = 1$, l'inégalité de SOBOLEV de dimension $n > 2$ à $p = 2n/(n-2)$ tandis que l'inégalité de SOBOLEV logarithmique est obtenue lorsque $p \to 2$.

0.3. Inégalités fonctionnelles et contrôles du semi-groupe

L'un des grands attraits des inégalités de SOBOLEV est leur lien avec la propriété d'ultracontractivité du semi-groupe

$$\|\mathbf{P}_t\|_{1\to+\infty} \leqslant C/t^{\frac{n}{2}}$$

(pour $t \leqslant 1$ si μ est une probabilité, pour tout $t > 0$, si $A_2 = 0$ comme dans l'espace euclidien). Ce lien a été étudié par VAROPOULOS dans les années 80 (voir [Var84, Var85]).

Lorsque $\|\mathbf{P}_t\|_{1\to+\infty} < +\infty$, l'opérateur $\mathbf{P}_t$ est représentable grâce à un noyau de probabilités à densité par rapport à la mesure de référence μ (voir [Bak94], lemme 4.3). Ainsi,

$$\|\mathbf{P}_t\|_{1\to+\infty} = \sup_{x,y\in\mathbb{E}} p_t(x,y)$$

(ou tout au moins l'essentiel sup dans $\mathbf{L}^0(\mu\otimes\mu)$). Des résultats similaires de contrôle uniforme de la solution d'EDP avaient été obtenus précédemment par NASH et MOSER à l'aide d'inégalités qui portent maintenant leur noms (voir [Nas58, Mos64]). Ces dernières peuvent être vue comme des cas particuliers des inégalités de GAGLIARDO-NIRENBERG $GN(s,r,\theta)$

$$\forall f \in \mathcal{A}, \qquad \|f\|_r \leqslant \left(A\,\|f\|_2^2 + B\int \mathbf{\Gamma}(f)\,d\mu\right)^{\theta/2} \|f\|_s^{1-\theta}, \tag{8}$$

où $1 \leqslant s < r$ et $\theta \in [0,1]$. L'inégalité de NASH de dimension n correspond à $s = 1$, $r = 2$ et $\theta = n/(n+2)$. L'inégalité de MOSER à $s = 2$, $r = 2+4/n$ et $\theta = n/(n+2)$.

Une inégalité de SOBOLEV d'exposant $p > 2$ induit une inégalité de GAGLIARDO-NIRENBERG $GN(s,r,\theta(s,r,p))$ pour tous $1 \leqslant s < r < p$, où $\theta = \theta(s,r,p)$ est défini par

$$\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{s}.$$

Cela provient directement d'un résultat d'interpolation (qui se déduit de l'inégalité de HÖLDER) : si $1 \leqslant s < r < p$ et $f \in \mathbf{L}^s \cap \mathbf{L}^p$, alors $f \in \mathbf{L}^r$ et

$$\|f\|_r \leqslant \|f\|_p^{\theta}\,\|f\|_s^{1-\theta},$$

θ étant défini comme précédemment. Fixons $r = 2$ et $p = 2n/(n-2) > 2$. Alors $\theta = \theta(s)$ décroît de $n/(n+2)$ à 0 lorsque s parcourt $[1,2[$. L'inégalité d'interpolation précédente affirme exactement que la fonction $u \mapsto \log\|f\|_{1/u}$ est convexe. Cela permet de voir que les inégalités $GN(s,2,\theta(s))$ sont de plus en plus fortes (avec les

mêmes constantes) quand s croît (voir [ABC$^+$00]). A la limite, quand $s \to 2$, on obtient l'inégalité dite d'*entropie-énergie logarithmique*

$$\forall f \in \mathcal{A}, \qquad \mathbf{Ent}_\mu\big(f^2\big) \leqslant \frac{n}{2}\,\|f\|_2^2 \log\left(A + B\,\frac{\int \mathbf{\Gamma}(f)\,d\mu}{\|f\|_2^2}\right)$$

notée $EEL(A,B)$. D'après [Nas58], une inégalité de NASH de dimension n implique la propriété d'ultracontractivité qui elle-même induit une inégalité de SOBOLEV de dimension n d'après le théorème de VAROPOULOS. Ainsi, on a une chaîne d'inégalités de plus en plus faibles (avec les mêmes constantes et la même dimension) de l'inégalité de SOBOLEV à l'inégalité de NASH en passant par l'inégalité EEL et les inégalités $GN(s,2,\theta(s))$. Et on peut *remonter* à l'inégalité de SOBOLEV quitte à perdre sur les constantes.

Lorsque l'on n'a plus d'inégalité de SOBOLEV mais seulement une inégalité de SOBOLEV logarithmique, la propriété d'ultracontractivité précédente doit être remplacée par une propriété plus faible, l'hypercontractivité du semi-groupe : pour tous $t > 0$ et $p > 1$, $\mathbf{P}_t$ est une contraction de $\mathbf{L}^p(\mu)$ dans $\mathbf{L}^q(\mu)$ pour $q = q(t) > p$ défini par $q(t) = 1 + (p-1)\,e^{4\,t/B_2}$. Le théorème de GROSS (voir [Gro75]) établit l'équivalence entre une inégalité de SOBOLEV logarithmique tendue de constante $B_2 > 0$ et la propriété d'hypercontractivité du semi-groupe avec la même constante. Remarquons ici que cette propriété permet dans certains modèles (pour la dynamique de GLAUBER en mécanique statistique là où il y a un trou spectral) de renforcer l'ergodicité exponentielle dans $\mathbf{L}^2$ donnée par l'inégalité de POINCARÉ en une ergodicité exponentielle en norme uniforme.

En fait, le théorème de GROSS peut être étendu à une hypercontractivité similaire de $\mathbf{L}^p(\mu)$ dans $\mathbf{L}^q(\mu)$ (mais $\mathbf{P}_t$ n'est plus une contraction) pour des inégalités de SOBOLEV logarithmique non tendues. En se basant sur ce résultat, BAKRY a étendu, pour les semi-groupes de diffusion, les bornes uniformes sur le noyau à des inégalités plus générales que l'inégalité de SOBOLEV. Rappelons que l'inégalité $E.E.L(A,B)$ s'écrit

$$\forall f \in \mathcal{A}, \qquad \mathbf{Ent}_\mu\big(f^2\big) \leqslant \|f\|_2^2\,\Phi\left(\frac{\int \mathbf{\Gamma}(f)\,d\mu}{\|f\|_2^2}\right)$$

où $\Phi(x) = \frac{n}{2}\log\big(A + B\,x\big)$. Pour une fonction Φ concave, strictement croissante et de classe $\mathcal{C}^1$, l'inégalité précédente s'appelle inégalité d'*entropie-énergie* de fonction Φ (notée $E.E.\Phi$). La concavité de Φ montre que, pour tous $u \geqslant 0$ et $v > 0$,

$$\Phi(u) \leqslant \Phi(v) + \Phi'(v)\,(u - v)$$

et, par suite $E.E.\Phi$ équivaut à la famille d'inégalités de SOBOLEV logarithmique non tendues paramètrée par $v > 0$

$$\forall f \in \mathcal{A}, \qquad \mathbf{Ent}_\mu\big(f^2\big) \leqslant \Phi'(v)\int \mathbf{\Gamma}(f)\,d\mu + \Psi(v)\,\|f\|_2^2,$$

où $\Psi(v) \overset{\text{déf.}}{=} \Phi(v) - v\,\Phi'(v)$. D'après la propriété de diffusion, en appliquant l'inégalité précédente à $f^{p/2}$, pour $p \geqslant 1$, on obtient, pour tout $v > 0$,

$$\text{(9)} \qquad \forall f \in \mathcal{A}, \qquad \mathbf{Ent}_\mu(f^p) \leqslant -\Phi'(v)\frac{p^2}{4\,(p-1)} \int f^{p-1}\,\mathbf{L}f\,d\mu + \Psi(v)\,\|f\|_p^p.$$

Cette inégalité dans laquelle on choisit une fonction $v(p)$ adéquate est à l'origine du résultat suivant dû à Bakry.

Théorème 0.3.1 ([Bak94]). — *Supposons E.E.Φ vérifiée (et par conséquent* (9)*). Si $1 \leqslant p \leqslant q \leqslant +\infty$ et $v(s)$ est une fonction sur $[p,q]$ telle que les intégrales*

$$t = \int_p^q \Phi'(v(s))\frac{ds}{4\,(s-1)} \quad \textit{et} \quad m = \int_p^q \Psi(v(s))\frac{ds}{s^2}$$

sont finies, alors

$$\text{(10)} \qquad \|\mathbf{P}_t f\|_q \leqslant e^m\,\|f\|_p.$$

Décrivons brièvement la preuve de ce résultat. L'idée principale est, suivant Gross, de poser

$$V(u) = e^{-m(u)}\,\|\mathbf{P}_u f\|_{p(u)}$$

(où $p(u)$ est un difféomorphisme $\mathcal{C}^1$ croissant de $[0,t]$ dans $[p,q]$ et $m(u)$ est une fonction régulière) et d'étudier les variations de $\log V$. Quitte à changer f en $|f|+\varepsilon$, on peut supposer $f > 0$. Posons $K(u) = \|\mathbf{P}_u f\|_{p(u)}^{p(u)}\,p^2(u)/p'(u) > 0$. Alors

$$\begin{aligned} &K(u)\,(\log V)'(u) \\ &\quad = \mathbf{Ent}_\mu\Big((\mathbf{P}_u f)^{p(u)}\Big) + \frac{p^2(u)}{p'(u)}\,\mathbf{E}_\mu\Big((\mathbf{P}_u f)^{p(u)-1}\,\mathbf{L}\mathbf{P}_u f\Big) - \frac{p^2(u)m'(u)}{p'(u)}\,\|\mathbf{P}_u f\|_{p(u)}^{p(u)} \end{aligned}$$

que l'on majore grâce à (9) par

$$\begin{aligned} &\left[\Psi(v(p(u))) - \frac{p^2(u)m'(u)}{p'(u)}\right]\|\mathbf{P}_u f\|_{p(u)}^{p(u)} \\ &\qquad - \left[\Phi'(v(p(u)))\,\frac{p^2(u)}{4\,(p(u)-1)} - \frac{p^2(u)}{p'(u)}\right]\mathbf{E}_\mu\Big((\mathbf{P}_u f)^{p(u)-1}\,\mathbf{L}\mathbf{P}_u f\Big). \end{aligned}$$

Ainsi, V est décroissante sur $[0,t]$ pourvu que

$$\text{(11)} \qquad \text{pour tout } u \in [0,t], \quad \left|\begin{array}{l} m'(u) = \frac{p'(u)\,\Psi(v(p(u)))}{p^2(u)} \\ \Phi'(v(p(u)))\,\frac{p'(u)}{4\,(p(u)-1)} = 1 \end{array}\right. .$$

Il ne reste plus qu'à remarquer que si les intégrales définies plus haut sont finies, alors

$$t(r) \overset{\text{déf.}}{=} \int_p^r \Phi'(v(s))\frac{ds}{4\,(s-1)}$$

est un difféomorphisme croissant de $[p,q]$ dans $[0,t]$ dont l'inverse $p(u)$ vérifie (11) pour la fonction

$$m(u) = \int_p^{p(u)} \Psi(v(s)) \frac{ds}{s^2}.$$

La preuve est complète.

Le choix de la fonction v provient d'une optimisation dans les inégalités rencontrées, et ceci pour chaque p. Dans le cas de $E.E.L$, on a

$$v(s) = \frac{\lambda s^2}{s-1}, \tag{12}$$

avec un paramètre $\lambda > 0$.

Nous avons vu que les inégalités de type SOBOLEV sont toutes équivalentes aux constantes près. Tenir compte des meilleures constantes est néanmoins fondamental et leur connaissance permet d'obtenir des résultats de comparaison. Pour illustrer ceci, signalons un résultat dû à BAKRY, CONCORDET et LEDOUX concernant l'inégalité $E.E.L$ de type euclidien (voir [BCL97]). En utilisant l'inégalité de SOBOLEV logarithmique optimale pour la mesure gaussienne, ces derniers prouvent que si pour toute fonction régulière f sur $\mathbb{R}^n$ telle que $\int f^2\,dx = 1$,

$$\int f^2 \log f^2\,dx \leqslant \Phi\Big(\int |\nabla f|^2\,dx\Big), \tag{13}$$

alors

$$\Phi(u) \geqslant \frac{n}{2} \log\left(\frac{2}{n\pi e}\,u\right).$$

De plus, (13) a lieu pour cette dernière fonction. Le contrôle (10) avec les constantes adaptées et le choix de la fonction $v(s)$ mentionné en (12) montrent alors le théorème de comparaison suivant : si $\mathbf{L}$ est un générateur de MARKOV de diffusion admettant une mesure réversible μ qui vérifie l'inégalité $E.E.L$ avec la constante optimale de $\mathbb{R}^n$, i.e. si, pour toute $f \in \mathcal{A}$,

$$\mathbf{Ent}_\mu\big(f^2\big) \leqslant \frac{n}{2}\,\|f\|_2^2 \log\left(\frac{2}{n\pi e}\,\frac{\int \mathbf{\Gamma}(f)\,d\mu}{\|f\|_2^2}\right)$$

alors, pour tout $t > 0$,

$$\|\mathbf{P}_t\|_{1\to+\infty} = \sup_{x,y\in\mathbb{E}} p_t(x,y) \leqslant \frac{1}{(4\,\pi t)^{n/2}}.$$

Et cette borne est atteinte par le semi-groupe de la chaleur (1).

En fait, par une méthode due à DAVIES et développée dans [Bak94], on peut obtenir avec les mêmes arguments un contrôle non uniforme (dit *hors-diagonal*) des densités de transition. L'idée consiste à déduire de $E.E.\Phi$ une inégalité du type (9) avec des constantes modifiées, mais cette fois pour le semi-groupe

$$\mathbf{P}_t^{\lambda g}(f) = e^{\lambda g}\mathbf{P}_t\big(e^{-\lambda g} f\big),$$

ce qui donne un contrôle uniforme des densités $p_t^{\lambda\, g}(x,y) = e^{\lambda\,(g(x)-g(y))}\, p_t(x,y)$. En optimisant en λ et en g, on aboutit à une borne $p_t(x,y) \leqslant e^{K(t,d(x,y))}$ pour la distance (usuelle dans le cas des variétés riemanniennes)

$$d(x,y) = \sup_{\substack{g\in\mathcal{A}\\ \mathbf{\Gamma}(g)\leqslant 1}} (d_g).$$

Ici le sup est en fait un sup essentiel dans $\mathbf{L}^0(\mu\otimes\mu)$ et la fonction d_g est définie presque partout sur $\mathbb{E}\times\mathbb{E}$ par $d_g(x,y) = g(x) - g(y)$.

Malheureusement, cette borne n'est pas optimale dans le cas euclidien. Pour aboutir à une borne optimale, il faut changer la distance en une distance plus faible, la distance harmonique

$$d_H(x,y) = \sup_{\substack{\mathbf{\Gamma}(g)\leqslant 1\\ \mathbf{L}g=0}} (d_g).$$

Cette distance coïncide avec la distance usuelle dans le cas euclidien. Finissons en donnant le résultat annoncé : si $\mathbf{L}$ est réversible, de diffusion et satisfait $E.E.L$ euclidien comme précédemment, alors

$$p_t(x,y) \leqslant \frac{1}{(4\,\pi\,t)^{n/2}}\, \exp\left(-\frac{d_H(x,y)^2}{4\,t}\right).$$

Et cette borne est optimale pour le semi-groupe de la chaleur.

0.4. Le critère de courbure-dimension

Pour établir des inégalités fonctionnelles de type SOBOLEV, un outil performant est basé sur une analyse *géométrique* via une généralisation de la notion de *courbure* et de *dimension* des variétés aux générateurs de MARKOV. Abstraitement, ces paramètres sont définis à l'aide d'une inégalité (dite de courbure-dimension) comparant trois *formes* quadratiques : le carré du champ $\mathbf{\Gamma}(f)$, la forme $(\mathbf{L}f)^2$ et un nouvel opérateur quadratique sur $\mathcal{A}$, l'opérateur $\mathbf{\Gamma_2}$.

En découlent alors plusieurs critères, dits encore de courbure-dimension, donnant lieu à des inégalités de SOBOLEV et de SOBOLEV logarithmique pour la mesure invariante et surtout à un contrôle des constantes explicite en la courbure et la dimension qui s'avère optimal pour les modèles sphérique et gaussien. L'hypothèse de diffusion (et la symétrie) y jouent un grand rôle. Ces critères sont souvent appelés critères de BAKRY-EMERY du nom de leurs initiateurs (voir [BÉ85b, BÉ85a] pour l'idée fondatrice ou encore [Bak94] pour l'obtention de la constante de SOBOLEV optimale des sphères via cette méthode). On parle encore, plus simplement, de critère $\mathbf{\Gamma_2}$.

Une analyse plus fine montre que, pour établir des inégalités pour la mesure invariante μ, il suffit d'inégalités de courbure-dimension *intégrées* par rapport à μ (néanmoins le contrôle optimal de la constante de SOBOLEV nécessite une inégalité

point par point). Lorsque la dimension est infinie (comme par exemple pour le semi-groupe d'ORNSTEIN-UHLENBECK), l'inégalité de courbure ρ point par point fournit en fait des inégalités *locales* (i.e. pour les noyaux $p_t(x, dy)$).

Ces différents points ne seront ici qu'abordés. Nous renvoyons à [Bak94, Bak97, Led00, ABC$^+$00] pour de plus amples détails.

L'inégalité de courbure-dimension. —

Définition 0.4.1. — *On appelle* opérateur $\mathbf{\Gamma_2}$ *la forme bilinéaire symétrique suivante : pour f et g dans $\mathcal{A}$,*

$$\mathbf{\Gamma_2}(f,g) \overset{déf.}{=} \frac{1}{2}[\mathbf{L}\mathbf{\Gamma}(f,g) - \mathbf{\Gamma}(f,\mathbf{L}g) - \mathbf{\Gamma}(\mathbf{L}f,g)]. \tag{14}$$

On note $\mathbf{\Gamma_2}(f)$ pour $\mathbf{\Gamma_2}(f,f)$. On dit que le semi-groupe $(\mathbf{P}_t)$ vérifie une inégalité de courbure-dimension $\mathrm{CD}(\rho, m)$ si, pour toute fonction $f \in \mathcal{A}$,

$$\mathbf{\Gamma_2}(f) \geqslant \rho\,\mathbf{\Gamma}(f) + \frac{1}{m}(\mathbf{L}f)^2.$$

L'inégalité $\mathrm{CD}(\rho, \infty)$ affirme naturellement que

$$\mathbf{\Gamma_2}(f) \geqslant \rho\,\mathbf{\Gamma}(f).$$

Pour un semi-groupe de diffusion sur une variété riemannienne, réversible pour la mesure μ de densité $\exp(\Phi)$ par rapport à la mesure riemannienne, l'inégalité $\mathrm{CD}(\rho, m)$ a un contenu géométrique. Nécessairement, on a $m \geqslant n$, où n désigne la dimension de la variété. Dans le cas $m > n$, cela revient à dire que, en tant que formes quadratiques,

$$\mathrm{Ric} - \nabla\nabla\Phi \geqslant \rho g + \frac{1}{m-n}\nabla\Phi \otimes \nabla\Phi.$$

Le cas $m = n$ implique que μ est la mesure riemannienne, ou de manière équivalente que $\mathbf{L} = \Delta$, et c'est plus simplement

$$\mathrm{Ric} \geqslant \rho g.$$

Cela signifie que le tenseur de RICCI est borné inférieurement par ρ.

Le critère de courbure-dimension pour l'inégalité de SOBOLEV. — Comme nous l'avons dit précédemment, l'inégalité de courbure-dimension permet d'établir des inégalités fonctionnelles en donnant dans les modèles typiques de courbure strictement positive (sphères, mesure gaussienne) les meilleures constantes. Le critère le plus fin est celui s'appliquant aux inégalités de SOBOLEV.

Théorème 0.4.2 **(Bakry-Ledoux)**. — *Soit* $\mathbf{L}$ *un générateur de* MARKOV *ergodique, de diffusion et symétrique par rapport à μ. Si $CD(\rho, m)$ a lieu pour un $\rho > 0$ et un $m > 2$, alors la mesure μ est finie, et, normalisée en une mesure de probabilité, elle vérifie l'inégalité de* SOBOLEV (6) *avec $A_2 = 1$ et $B_2 = \frac{4}{m(m-2)}\frac{m-1}{\rho}$.*

En remarquant que le laplacien de la sphère de dimension n vérifie $CD(n-1,n)$, on obtient ainsi les meilleures constantes de [Aub82].

La preuve du théorème 0.4.2 se décompose en deux étapes distinctes. La première consiste à prouver l'existence d'une inégalité de SOBOLEV via celle d'une inégalité *entropie-énergie* logarithmique $E.E.L$ de dimension m de constante $A=1$ et $B=\frac{4}{\rho m}$. Cette étape est l'occasion d'illustrer les raisonnements génériques qui interviennent dans les critères de courbure-dimension et de bien comprendre d'où vient l'opérateur $\mathbf{\Gamma_2}$. Elle utilise la formule générale de composition pour le $\mathbf{\Gamma_2}$ des générateurs de diffusion à travers les deux formules suivantes

$$\mathbf{\Gamma_2}(\Psi(f)) = \Psi'^2(f)\mathbf{\Gamma_2}(f) + \Psi''(f)\,\Psi'(f)\,\mathbf{\Gamma}(f,\mathbf{\Gamma}(f)) + \Psi''^2(f)\mathbf{\Gamma}^2(f), \tag{15}$$

$$\mathbf{\Gamma_2}(f,\Psi(g)) = \Psi'(g)\,\mathbf{\Gamma_2}(f,g) + \Psi''(g)\left(\mathbf{\Gamma}(g,\mathbf{\Gamma}(f,g)) - \frac{1}{2}\mathbf{\Gamma}(f,\mathbf{\Gamma}(g))\right). \tag{16}$$

La seconde étape a pour but le contrôle (optimal pour les sphères) de la constante de SOBOLEV . Elle est basée sur des raisonnements issus de l'analyse non-linéaire que nous rappelerons brièvement. Nous ne ferons que l'esquisser et renvoyons à [Bak94] pour les détails.

Pour prouver l'inégalité $E.E.L$, il suffit, en changeant f^2 en f, de prouver que pour toute fonction $f\in\mathcal{A}$ positive et d'intégrale 1,

$$\mathbf{E}_\mu(f\log f) \leqslant \frac{m}{2}\log\left(1+\frac{1}{\rho m}\mathbf{E}_\mu\left(\frac{\mathbf{\Gamma}(f)}{f}\right)\right).$$

Pour une telle fonction, posons $\alpha(t)=\mathbf{E}_\mu(\mathbf{P}_t f\log\mathbf{P}_t f)$. En utilisant la symétrie et puisque $\mathbf{L}1=0$,

$$\alpha'(t) = \mathbf{E}_\mu((\log g+1)\mathbf{L}g) = -\mathbf{E}_\mu(\mathbf{\Gamma}(g,\log g)) = -\mathbf{E}_\mu(g\,\mathbf{\Gamma}(\log g))\,.$$

Remarquons que $\alpha'(t)\leqslant 0$. Ici, on a noté (après avoir dérivé) $g=\mathbf{P}_t f$. Puisque $\mathbf{\Gamma}$ est bilinéaire, on a

$$\alpha''(t) = -\mathbf{E}_\mu\left(\mathbf{\Gamma}(\mathbf{L}g,\log g) + \mathbf{\Gamma}\left(g,\frac{\mathbf{L}g}{g}\right)\right).$$

Le terme intégré est *presque* $-2\,\mathbf{\Gamma_2}(g,\log g)$ et une utilisation judicieuse de (16) et de la propriété de diffusion $\mathbf{L}\log g = \mathbf{L}(g)/g - \mathbf{\Gamma}(\log g)$ montre alors que

$$\mathbf{\Gamma}(\mathbf{L}g,\log g) + \mathbf{\Gamma}\left(g,\frac{\mathbf{L}g}{g}\right) = -2\,g\,\mathbf{\Gamma_2}(\log g) + 2\,\mathbf{L}\,\mathbf{\Gamma}(g,\log g)$$

ce qui donne en intégrant

$$\alpha''(t) = 2\,\mathbf{E}_\mu(g\,\mathbf{\Gamma_2}(\log g)).$$

En multipliant par $g\geqslant 0$ l'inégalité $\mathrm{CD}(\rho,m)$ appliquée à $\log g$ puis en intégrant, on obtient

$$\mathbf{E}_\mu(g\mathbf{\Gamma_2}(\log g)) \geqslant \rho\mathbf{E}_\mu(g\mathbf{\Gamma}(\log g)) + \frac{1}{m}\big(\mathbf{E}_\mu\big(g(\mathbf{L}\log g)^2\big)\big).$$

Remarquons à ce stade que seule cette inégalité intégrée est nécéssitée. En se souvenant que g est de moyenne 1 et en utilisant l'inégalité de CAUCHY-SCHWARZ, on aboutit à l'inégalité différentielle non linéaire en α' suivante

$$\alpha''(t) \geqslant -2\rho\alpha'(t) + \frac{2}{m}\alpha'(t)^2$$

que l'on linéarise en posant $\beta = \psi(\alpha')$. On obtient

$$\psi(u) = \frac{1}{2\rho}\log\left(\frac{2}{m} - \frac{2\rho}{u}\right)$$

et $\beta'(t) \geqslant 1$. Ceci s'écrit

$$-\alpha'(t) \leqslant \frac{-m\rho\alpha'(0)}{(m\rho - \alpha'(0))\exp(2\rho t) + \alpha'(0)}.$$

et, en intégrant, on obtient

$$\alpha(0) - \alpha(\infty) = \frac{m}{2}\log\left(1 - \frac{\alpha'(0)}{\rho m}\right).$$

C'est exactement le résultat souhaité puisque $\alpha(\infty) = 0$ par ergodicité, $\alpha(0) = \mathbf{E}_\mu(f\log f)$ et enfin $\alpha'(0) = -\mathbf{E}_\mu(\mathbf{\Gamma}(f)/f)$.

L'idée de la seconde étape provient de l'analyse non linéaire sur les variétés. Le théorème de RELLICH-KONDRAKOV affirme que, sur une variété riemannienne compacte de dimension m (sur laquelle il y a une inégalité de SOBOLEV) le plongement de H_1 dans $\mathbf{L}^p$ est compact pour tout $p < \tilde{p} = 2\,m/(m-2)$. Ce théorème est d'une grande utilité pour montrer l'existence de fonctions extrémales (non constantes) pour l'inégalité de SOBOLEV ou plus généralement de fonctions minimisantes de fonctionnelles liées à l'énergie $\mathcal{E}(f,f)$ et aux normes $\mathbf{L}^p$. Ceci sert entre autre à résoudre le problème de YAMABE sur l'existence d'une métrique conforme plate pour la courbure scalaire (voir [Heb97]).

En étendant le théorème de RELLICH-KONDRAKOV pour des générateurs satisfaisant une inégalité de SOBOLEV, BAKRY prouve l'existence d'une fonction extrémale strictement positive et non constante pour l'inégalité de SOBOLEV approchée

$$\|f\|_p^2 \leqslant (1+\varepsilon)\,\|f\|_2^2 + B_{opt}\,\mathcal{E}(f,f)$$

où B_{opt} est la constante optimale de l'inégalité de SOBOLEV de dimension m. C'est pour pouvoir appliquer le théorème de RELLICH-KONDRAKOV qu'on doit considérer $p < \tilde{p}$. $\varepsilon > 0$ n'est là que pour assurer que l'extrémale n'est pas constante. Cette extrémale est une fonction minimisante d'une fonctionnelle et par là vérifie une équation d'EULER-LAGRANGE que l'on peut écrire, en oubliant les détails techniques (p et ϵ) comme dans [Led00] et en notant p pour $\tilde{p}$,

$$-B_{opt}\,\mathbf{L}f = \left(f^{p-1} - f\right).$$

En posant $u = f^{1/r}$ pour $r > 0$, cette équation permet d'écrire (en utilisant la propriété de diffusion, une intégration par parties et après un calcul assez fin) que

$$\mathbf{E}_\mu\left(\frac{1}{u}\,\mathbf{L}u\,\mathbf{\Gamma}u\right) = B_{opt}^{-1}\,a\,\mathbf{E}_\mu(\mathbf{\Gamma}u) - b\,\mathbf{E}_\mu\big((\mathbf{L}u)^2\big) - c\,\mathbf{E}_\mu\left(\frac{1}{u^2}\,\mathbf{\Gamma}^2 u\right). \tag{17}$$

avec $a = \frac{p-2}{r\,(p-1)}$, $b = \frac{1}{r\,(p-1)}$ et $c = \frac{r-1}{r}\,\frac{r(p-2)+1}{p-1}$. L'intérêt de cette égalité devient visible si l'on applique l'inégalité $\mathbf{\Gamma_2} \geqslant \rho\,\mathbf{\Gamma} + \frac{1}{m}(\mathbf{L})^2$ à u^s, $s > 0$. On obtient (grâce à (15) et toujours à des intégrations par parties)

$$\rho\,\mathbf{E}_\mu(\mathbf{\Gamma}u) \leqslant b'\,\mathbf{E}_\mu\big((\mathbf{L}u)^2\big) - a'\,\mathbf{E}_\mu\left(\frac{1}{u}\,\mathbf{L}u\,\mathbf{\Gamma}u\right) + c'\,\mathbf{E}_\mu\left(\frac{1}{u^2}\,\mathbf{\Gamma}^2 u\right)$$

avec $a' = (s-1)\,(1+2/m)$, $b' = 1 - 1/m$ et $c' = (s-1)\,[1 + (s-1)\,(1-1/m)]$. Ainsi (17) permet d'éliminer le terme en $\mathbf{E}_\mu\big(\frac{1}{u}\,\mathbf{L}u\,\mathbf{\Gamma}u\big)$. On ajuste alors le choix de s pour que le coefficient de $\mathbf{E}_\mu((\mathbf{L}u)^2)$ soit nul. Il faut prendre

$$s - 1 = -(p-1)\,r\,\frac{m-1}{m+2}.$$

Finalement, on est amené à

$$\left(B_{opt}^{-1}\,\frac{(p-2)\,(m-1)}{m} - \rho\right)\mathbf{E}_\mu(\mathbf{\Gamma}u) \geqslant K(r)\,\mathbf{E}_\mu\left(\frac{1}{u^2}\,\mathbf{\Gamma}^2 u\right)$$

où $K(r)$ est un polynôme de degré 2 dont le discriminant est positif d'après la valeur de p. On peut donc trouver r tel que $K(r) \geqslant 0$. Ce qui prouve la borne annoncée sur B_{opt} puisque u n'est pas constante.

Autres critères de courbure-dimension. — Lorsque la dimension est infinie, une adaptation de la méthode de semi-groupe développée dans la première étape de la preuve précédente donne directement des inégalités de POINCARÉ et SOBOLEV logarithmique avec des constantes qui s'avèrent optimales dans le cas du modèle gaussien. Nous nous contentons d'énoncer ces résultats et renvoyons à [ABC+00] pour une exposition détaillée.

Proposition 0.4.3. — *Si $(\mathbf{P}_t)_{t\geqslant 0}$ admet μ pour mesure réversible, est ergodique et si $\rho > 0$, les assertions suivantes sont équivalentes :*

(i) $\forall f \in \mathcal{A}, \quad \mathbf{E}_\mu(\mathbf{\Gamma_2}(f)) \geqslant \rho\mathbf{E}_\mu(\mathbf{\Gamma}(f))$

(ii) $\forall f \in \mathcal{A}, \quad \mathbf{Var}_\mu(f) \leqslant \dfrac{1}{\rho}\mathbf{E}_\mu(\mathbf{\Gamma}(f))$.

Proposition 0.4.4. — *Soit $(\mathbf{P}_t)_{t\geqslant 0}$ un semi-groupe de diffusion réversible et ergodique par rapport à μ et soit $\rho > 0$, satisfaisant l'assertion suivante*

$$\forall f \in \mathcal{A}, \quad \mathbf{E}_\mu\big(e^f\mathbf{\Gamma_2}(f)\big) \geqslant \rho\mathbf{E}_\mu\big(e^f\mathbf{\Gamma}(f)\big)\,. \tag{18}$$

Alors la mesure μ vérifie l'inégalité de SOBOLEV *logarithmique*

$$\forall f \in \mathcal{A}, \quad \mathbf{Ent}_\mu\big(f^2\big) \leqslant \frac{2}{\rho}\mathbf{E}_\mu(\mathbf{\Gamma}(f))\,.$$

La réciproque de cette dernière proposition est fausse (voir [ABC$^+$00]).

La condition $\mathrm{CD}(\rho, m)$ donne également des inégalités de POINCARÉ et de SOBOLEV logarithmique avec des constantes qui sont optimales pour les sphères.

Inégalités locales. — L'inégalité $\mathrm{CD}(\rho, \infty)$ lorsqu'elle est vérifiée point par point contient en fait beaucoup plus d'information que les inégalités pour la mesure invariante comme le suggère la proposition 0.4.3. Les résultats suivants sont issus de [Bak97] et sont tirés de [ABC$^+$00]. Ils sont valables quel que soit le signe de la courbure.

Théorème 0.4.5. — *Les assertions suivantes sont équivalentes :*

(i) $\mathrm{CD}(\rho, \infty)$ *est vérifiée*

(ii) $\forall f \in \mathcal{A}, \forall t > 0, \quad \mathbf{\Gamma}(\mathbf{P}_t f) \leqslant e^{-2\rho t}\mathbf{P}_t(\mathbf{\Gamma}(f))$

(iii) $\forall f \in \mathcal{A}, \forall t > 0, \quad \mathbf{P}_t(f^2) - (\mathbf{P}_t f)^2 \leqslant \dfrac{1 - e^{-2\rho t}}{\rho}\mathbf{P}_t(\mathbf{\Gamma}(f))$.

Si de plus le semi-groupe est de diffusion, elles sont encore équivalentes à chacune des inégalités suivantes :

(iv) $\forall f \in \mathcal{A}, \quad 4\mathbf{\Gamma}(f)[\mathbf{\Gamma_2}(f) - \rho\mathbf{\Gamma}(f)] \geqslant \mathbf{\Gamma}(\mathbf{\Gamma}(f))$

(v) $\forall f \in \mathcal{A}, \forall t > 0, \sqrt{\mathbf{\Gamma}(\mathbf{P}_t f)} \leqslant \exp(-\rho t)\mathbf{P}_t(\sqrt{\mathbf{\Gamma} f})$

(vi) $\forall f \in \mathcal{A}, \forall t > 0, \mathbf{P}_t(f^2 \log f^2) - \mathbf{P}_t(f^2)\log \mathbf{P}_t(f^2) \leqslant \dfrac{2}{\rho}(1 - e^{-2\rho t})\mathbf{P}_t(\mathbf{\Gamma}(f))$.

Nous ne détaillerons pas la preuve de ces équivalences. Signalons toutefois que, en ce qui concerne (i) à (iii), il s'agit de différentier des fonctions de la forme $\varphi(s) = \mathbf{P}_s(\mathbf{\Gamma}(\mathbf{P}_{t-s}f))$ ou $\psi(s) = \mathbf{P}_s((\mathbf{P}_{t-s}f)^2)$ ce qui fait apparaître les opérateurs $\mathbf{\Gamma_2}$ et $\mathbf{\Gamma}$ respectivement. L'amélioration surprenante que fournit (iv) -et par suite (v)- est due à la formule de composition pour l'opérateur $\mathbf{\Gamma_2}$ des diffusions que nous avons évoquée partiellement.

Ces inégalités locales ont été utilisées récemment par F.MALRIEU pour approcher la solution d'un certain type EDP non linéaires par la loi d'une particule d'un système de particules en interaction en grande dimension (voir [Mal01]).

Lorsque la dimension entre en jeu, il est beaucoup plus difficile d'obtenir des inégalités locales. En cherchant à atteindre ce but pour des inégalités de SOBOLEV, G.SCHEFFER a exhibé une famille de noyaux markoviens à un paramètre (mais qui n'est plus un semi-groupe) qui vérifie une inégalité de POINCARÉ locale sous la condition $\mathrm{CD}(\rho, m)$ ([Sch01]).

0.5. Inégalités fonctionnelles $\mathbf{L}^1$

Chacune des inégalités introduites dans la section 0.2 admet une version dite $\mathbf{L}^1$ ou isopérimétrique. Ces inégalités sont plus fortes que les inégalités $\mathbf{L}^2$ en général mais leur sont équivalentes aux constantes près lorsque le semi-groupe est de courbure minorée.

Inégalité de Cheeger. — On dit que $(\mu, \mathbf{\Gamma})$ vérifie une inégalité de Cheeger de constante $C_1 > 0$, si pour toute fonction f de $\mathcal{A}$ dont la loi sous μ a pour médiane m (on dit que m est la médiane de f),

$$\int |f - m| \, d\mu \leqslant C_1 \int \sqrt{\mathbf{\Gamma} f} \, d\mu.$$

Tout comme l'inégalité de Poincaré, cette inégalité admet comme modèle typique la mesure exponentielle $\epsilon(dx)$ pour laquelle la meilleure constante est 1.

Inégalité de Sobolev $\mathbf{L}^1$ de type euclidien. — Sur $\mathbb{R}^n$ muni de la mesure de Lebesgue, on peut montrer l'existence d'une inégalité de Sobolev de constante $A_2 = 0$: pour toute fonction $\mathcal{C}^\infty$ à support compact,

$$\|f\|_{2n/(n-2)} \leqslant B_2 \, \| \, |\nabla f| \, \|_2.$$

Ici, la constante A_2 n'est pas forcément plus grande que 1. L'algèbre $\mathcal{A}$ est en effet l'ensemble des fonctions $\mathcal{C}^\infty$ dont les dérivées sont à décroissance rapide et ne contient donc pas les constantes.

Sur $\mathbb{R}^n$ toujours, une inégalité $\mathbf{L}^1$ a encore lieu : pour toute fonction $\mathcal{C}^\infty$ à support compact,

$$\|f\|_{n/(n-1)} \leqslant B_1 \| \, |\nabla f| \, \|_1. \tag{19}$$

Les meilleures constantes B_1 et B_2 sont connues et valent

$$B_1(\mathbb{R}^n) = \left(n^{1-\frac{1}{n}} \omega_{n-1}^{\frac{1}{n}} \right)^{-1}$$

et

$$B_2(\mathbb{R}^n) = 4 \left(n(n-2) {\omega_n}^{\frac{2}{n}} \right)^{-1}$$

où ω_n est le volume de la sphère unité de $\mathbb{R}^{n+1}$(voir [Aub82]).

On dira que le couple $(\mu, \mathbf{\Gamma})$ vérifie une inégalité de Sobolev $\mathbf{L}^1$ de type euclidien de dimension $m > 1$ si, pour toute fonction de $\mathcal{A}$,

$$\left(\int f^{m/(m-1)} \mu(dx) \right)^{(m-1)/m} \leqslant B_1 \int \sqrt{\mathbf{\Gamma}(f)} \mu(dx).$$

Inégalité isopérimétrique gaussienne de Bobkov. — Le couple $(\mu, \mathbf{\Gamma})$ vérifie une inégalité isopérimétrique gaussienne (tendue) de constante B_1, si, pour toute fonction $f \in \mathcal{A}$ à valeurs dans $[0,1]$,

$$\mathcal{U}\left(\int f d\mu\right) \leqslant \int \sqrt{\mathcal{U}^2(f) + B_1 \mathbf{\Gamma}(f)} d\mu.$$

Ici, $\mathcal{U} = \varphi' \circ \varphi^{-1}$ où $\varphi(s) = \gamma(]-\infty, s])$ est la fonction de répartition de la mesure gaussienne γ. Comme pour l'inégalité de Sobolev logarithmique, le modèle typique pour cette inégalité est, comme son nom l'indique, la mesure gaussienne pour laquelle la meilleure constante est $B_1 = 1$.

Les inégalités de Cheeger et de Bobkov vérifient la propriété de tensorisation. Elles présentent donc un caractère infini-dimensionnel comme leurs analogues $\mathbf{L}^2$.

Inégalités isopérimétriques. — En appliquant formellement à des indicatrices d'ensembles les trois dernières inégalités avec les meilleures constantes dans les modèles qui leur sont associés, on obtient l'inégalité isopérimétrique du modèle en question. Pour la mesure exponentielle, on a, pour tout borélien $A \subset \mathbb{R}$,

$$\min\left(\epsilon(A), 1 - \epsilon(A)\right) \leqslant \epsilon_s(\partial A)$$

et les ensembles extrémaux sont les demi-droites. Pour la mesure de Lebesgue sur $\mathbb{R}^n$, notant $|\cdot|_n$ le volume de dimension n, on a, pour tout borélien $A \subset \mathbb{R}^n$,

$$n^{1-\frac{1}{n}} \omega_{n-1}^{\frac{1}{n}} |A|_n^{1-\frac{1}{n}} \leqslant |\partial A|_{n-1}$$

et les ensembles extrémaux sont les boules. Pour la mesure gaussienne enfin, pour tout borélien $A \subset \mathbb{R}$,

$$\mathcal{U}(\gamma(A)) \leqslant \gamma_s(\partial A)$$

et les ensembles extrémaux sont les demi-droites.

Comparaison entre inégalités $\mathbf{L}^1$ et $\mathbf{L}^2$. — Pour un semi-groupe de diffusion, une inégalité $\mathbf{L}^1$ entraîne toujours une inégalité $\mathbf{L}^2$. La réciproque est fausse en général (voir [BL96a]). Néanmoins, si le semi-groupe vérifie une inégalité $\mathrm{CD}(\rho, \infty)$ pour une courbure ρ éventuellement négative, l'inégalité $\mathbf{L}^2$ implique une inégalité $\mathbf{L}^1$. Il est important de remarquer que l'on ne conserve pas les constantes optimales lors de ces implications.

La conséquence fondamentale de l'inégalité de courbure $\mathrm{CD}(\rho, \infty)$ pour obtenir ces implications est la propriété régularisante du semi-groupe qui a alors lieu : il existe une constante c ne dépendant que de la borne de la courbure telle que, pour toute fonction $f \in \mathcal{A}$ et tout $t \in [0,1]$,

$$\|\mathbf{P}_t f\|_{Lip} = \| \, |\nabla \mathbf{P}_t(f)| \, \|_\infty \leqslant \frac{c}{\sqrt{t}} \|f\|_\infty.$$

Remarquons pour finir que les liens entre inégalités $\mathbf{L}^1$ et $\mathbf{L}^2$ sont loin de rester évidents lorsque la propriété de diffusion est mise en défaut, comme par exemple

pour un semi-groupe sur un ensemble fini. On pourra se référer au chapitre 1 pour le cas des inégalités de SOBOLEV logarithmique et d'isopérimétrie gaussienne.

0.6. Organisation de la thèse

Le premier chapitre est la reproduction d'un article publié aux Annales de l'Institut Henri POINCARÉ (voir [Fou00]). Nous montrons que, pour un semi-groupe de diffusion à courbure minorée (éventuellement négative), une inégalité de SOBOLEV logarithmique induit une inégalité isopérimétrique gaussienne de BOBKOV. Ce résultat souligne les liens étroits qui unissent ces deux inégalités, essentiellement reliées à une concentration gaussienne de la mesure. En outre, nous établissons une forme fonctionnelle affaiblie de l'inégalité de BOBKOV pour tout semi-groupe hypercontractif à sauts sur un ensemble fini. Cette inégalité permet d'obtenir une inégalité isopérimétrique gaussienne ensembliste lorsqu'elle est appliquée formellement aux indicatrices d'ensembles. Enfin, nous donnons des critères sur le potentiel d'interaction d'un système de spins permettant de contrôler la courbure de la dynamique de GLAUBER. Nous obtenons ainsi les inégalités isopérimétriques correspondantes sous la condition de mélange de DOBRUSHIN et SCHLOSMAN.

La seconde partie est consacrée à la détermination de l'ordre de grandeur exact de la constante de POINCARÉ d'une mesure log-concave sur la droite réelle. Le bon paramètre décrivant cette constante est le carré de la distance moyenne à la médiane. La preuve que nous donnons est basée sur la caractérisation de la constante de POINCARÉ à l'aide de la fonctionnelle de MUCKENHOUPT (voir [Mic99a, Mic99b, BG99]). Nous menons à bien un calcul de variations dans l'ensemble des fonctions convexes. Nous montrons que les fonctions extrémales du problème de minimisation sont des fonctions linéaires en 3 morceaux et prouvons que la fonctionnelle étudiée est minorée sur ces fonctions. Après avoir obtenu ce résultat, nous avons découvert un article de S.BOBKOV dans lequel une caractérisation similaire est donnée en termes de la variance de la mesure [Bob99]. Son approche est très éloignée de la nôtre. Elle est basée sur l'étude de la constante de CHEEGER et la caractérisation des mesures log-concaves sur $\mathbb{R}$ par leur fonction *isopérimétrique*. Nous décrivons les résultats obtenus par BOBKOV et HOUDRÉ sur le calcul exact de la constante de CHEEGER pour des mesures log-concaves en dimension 1. Puis nous adaptons la méthode de BOBKOV pour retrouver notre résultat.

La dernière partie est une collaboration avec D.BAKRY. Nous y établissons de nouveaux critères conduisant aux inégalités de SOBOLEV lorsque le critère de courbure-dimension de BAKRY et EMERY est mis en défaut. Contrairement aux opérateurs de diffusion, il est toujours possible de projeter une structure de DIRICHLET. La remarque fondamentale à l'origine de ce travail est que certaines de ces structures vérifiant le critère de courbure-dimension se projettent sur des structures qui ne le vérifient pas. Après avoir décrit en détails un exemple simple en géométrie sphérique qui servira de guide, nous donnons un critère général assurant l'existence d'une

structure de DIRICHLET satisfaisant le critère de BAKRY et EMERY qui se projette sur la structure de départ.

L'annexe A est un chapitre tiré d'un livre [ABC$^+$00] sur les inégalités de SOBOLEV logarithmique. Ce livre, écrit en collaboration avec sept autres doctorants du laboratoire de Statistique et Probabilités de l'Université Paul SABATIER, a été publié dans la collection *Panoramas et Synthèses* en 2000. Ce chapitre ne présentant pas de résultats personnels, il a été placé en annexe. Le théorème de WANG, qui y est décrit, est néanmoins à la base du travail qui a été effectué dans le chapitre 2 puisqu'il suggère que la fonction distance est liée à la constante de SOBOLEV logarithmique. Le but du chapitre est d'étudier les liens entre les inégalités de SOBOLEV logarithmiques et la propriété de concentration gaussienne de la mesure due à TALAGRAND. On y décrit l'*équivalence* entre concentration gaussienne, intégrabilité exponentielle des contractions et existence d'une inégalité de SOBOLEV logarithmique pour des semi-groupes de diffusion de courbure minorée.

CHAPITRE 1

HYPERCONTRACTIVITÉ ET ISOPÉRIMÉTRIE GAUSSIENNE. APPLICATIONS AUX SYSTÈMES DE SPINS.

Ce chapitre est la reproduction d'un article [Fou00] paru aux Annales de l'Institut Henri POINCARÉ en 2000.

1.1. Introduction

Pour approcher et étudier certaines mesures, on introduit des semi-groupes de MARKOV ergodiques. Si la mesure à étudier est invariante (ou mieux réversible) pour le semi-groupe, on obtient des informations sur la vitesse de convergence vers la moyenne en s'assurant que le semi-groupe possède une régularité analytique traduite par des inégalités faisant intervenir l'énergie, l'entropie, ou les moyennes de la racine du carré du champ. Nous considérerons ici les inégalités de trou spectral, de SOBOLEV logarithmiques et enfin, celles dont nous établirons l'existence en mécanique statistique, les inégalités isopérimétriques. Les relations existant entre ces inégalités font l'objet de nombreuses études (cf. par exemple [Bak94] et [BL96a] dans un cadre général, [Lar95] en mécanique statistique).

Les inégalités de SOBOLEV logarithmiques ont largement prouvé leur efficacité en mécanique statistique (cf. [HS87, HS89]). Leur existence a été étudiée pour des modèles de mécanique statistique à fibre compacte au début des années 90 par D.W.STROOCK et B.ZEGARLINSKI (cf. [Zeg90, SZ92b, SZ92a, Zeg92]) et parallèlement par LU et YAU [LY93] (voir aussi [MO94a, MO94b]). Plus récemment l'étude s'est portée à des modèles à fibre non compacte avec les travaux de B.ZEGARLINSKI, N.YOSHIDA, T.BODINEAU, B.HELFFER et M.LEDOUX (cf. [Zeg96, BH99, Yos01, Led01]).

Une inégalité de SOBOLEV logarithmique renforce l'inégalité de trou dans le spectre. En effet, pour un potentiel à portée finie et invariant par translation, HOLLEY, STROOCK et ZEGARLINSKI ont obtenu la convergence exponentielle de la dynamique de GLAUBER pour toute condition extérieure ξ vers l'unique mesure de GIBBS $\boldsymbol{\mu}$, non plus seulement dans $\mathbf{L}^2$ mais en norme uniforme (cf. [HS89, SZ95]).

Dans un cadre plus général, l'hypercontractivité fournit des constantes de convergence plus fines que l'existence d'un trou spectral (cf. [DSC96]).

L'isopérimétrie décrit une relation entre la mesure d'un ensemble et celle de son bord : $\eta(\mu(A)) \leq \mu_s(\partial A)$. L'isopérimétrie gaussienne correspond à la fonction isopérimétrique $\eta = \mathcal{U} = \varphi' \circ \varphi^{-1}$ où φ est la fonction de répartition d'une variable gaussienne. Suivant S. Bobkov, on peut introduire une forme fonctionnelle de l'inégalité isopérimétrique (cf. [Bob96b] et [Bob97]) qui lui est équivalente

$$(20) \qquad \forall f \in \mathcal{C}^\infty(\mathbb{R}) \text{ à valeurs dans } [0,1], \quad \mathcal{U}\left(\int f d\gamma\right) \leq \int \sqrt{\mathcal{U}^2(f) + |\nabla f|^2} d\gamma$$

Dans leurs travaux sur l'isopérimétrie gaussienne pour un générateur de diffusion de dimension infinie (cf. [BL96a]), D. Bakry et M. Ledoux ont montré que, dans ce cadre, l'isopérimétrie fonctionnelle entraîne l'hypercontractivité. L'isopérimétrie gaussienne en diffusion apporte ceci de plus que la loi de toute contraction au sens du carré du champ, i.e telle que $\Gamma(f) \leq 1$, est alors une contraction de gaussienne.

L'importance du critère $\mathbf{\Gamma_2}$ pour assurer l'existence d'une inégalité de Sobolev logarithmique dans les modèles de mécanique statistique a été exhibée par E.A. Carlen et D.W. Stroock (cf. [CS86]). Le résultat général sur l'implication de l'hypercontractivité par la condition $\mathbf{\Gamma_2}$ (i.e l'inégalité de courbure-dimension $\mathrm{CD}(R,\infty)$ pour une constante $R > 0$) est dû à D. Bakry et M. Emery (cf. [BÉ84]). Carlen et Stroock ont souligné son adéquation avec les modèles de mécanique statistique la liant aux propriétés du potentiel d'interaction via une condition suffisante de courbure strictement positive : une décroissance l^1 assez forte des variations secondes du potentiel.

D'autre part, dans [BL96a], D. Bakry et M. Ledoux s'attachent à l'étude des liens entre hypercontractivité et isopérimétrie. En dimension infinie, sous l'hypothèse de courbure réelle (et non plus strictement positive) du semi-groupe, l'hypercontractivité donne une isopérimétrie gaussienne amoindrie à savoir, pour tout A borélien,

$$\mathcal{U}(\mu(A)) \leq c\,\mu_s(\partial A)$$

où c est une constante > 0 dite constante d'isopérimétrie.

Nous renforçons ici ce résultat en obtenant pour les diffusions hypercontractives sur les variétés compactes une inégalité isopérimétrique fonctionnelle du même type que celle de S. Bobkov.

Nous étendons la méthode utilisée aux générateurs hypercontractifs réversibles sur un ensemble fini, pour lesquels nous obtenons une inégalité fonctionnelle plus faible que (20) (voir le paragraphe 1.2.4 pour plus de détails).

Pour les modèles de mécanique statistique à fibre continue, s'appuyant sur la démarche de E.A. Carlen et D.W. Stroock, on peut exhiber une condition sur le potentiel engendrant le contrôle, même négatif cette fois, de la courbure indépendamment de la taille de la boîte finie considérée. Sous cette condition et dans le cas

d'une diffusion, le contrôle de la constante de SOBOLEV logarithmique est équivalent à celui de la constante d'isopérimétrie gaussienne.

On rejoint ici les travaux de E. LAROCHE (cf. [Lar95]) dans lesquels, étendant les résultats de D.W. STROOCK et B. ZEGARLINSKI à des potentiels de portée infinie, l'auteur lie sous une condition de décroissance exponentielle des variations secondes du potentiel (recoupant notre condition de décroissance l^1) le contrôle du trou spectral à celui de la constante de SOBOLEV logarithmique et à une condition de mélange faible : toutes ces propriétés du modèle sont alors équivalentes. Vu ce qui précède, le contrôle de la constante d'isopérimétrie gaussienne est encore équivalent, sous la condition de décroissance exponentielle des interactions, à l'une quelconque de ces propriétés.

Pour les modèles de mécanique statistique discrets, nous utilisons ce raisonnement pour obtenir l'inégalité isopérimétrique fonctionnelle introduite au paragraphe 1.2.4 . B. ZEGARLINSKI a récemment démontré que pour les modèles à fibre discrète et à portée finie, une inégalité isopérimétrique du type (20) a lieu pour la mesure de GIBBS sous la condition de mélange de DOBRUSHIN et SCHLOSMAN (cf. [Zeg01]). Notre résultat transposé dans ce cadre est plus faible, mais la preuve que nous en donnons est conceptuellement simple, étant basée sur une méthode de semi-groupes, et laisse entrevoir un espoir d'obtenir une inégalité isopérimétrique fonctionnelle du type BOBKOV pour tout générateur hypercontractif en courbure minorée de constante indépendante de la taille de l'ensemble fini sous-jacent.

1.2. Isopérimétrie fonctionnelle gaussienne pour les semi-groupes hypercontractifs

1.2.1. Le critère Γ_2. — Ce paragraphe expose brièvement certaines notions et notations utilisées dans l'étude des semi-groupes de MARKOV et tirées de [Bak94] où l'hypercontractivité est traitée dans un cadre général. Nous renvoyons à cette référence pour de plus amples détails.

Tout d'abord, quelques précisions sur les semi-groupes que nous étudierons. Nous imposons au générateur la régularité suivante : on suppose que le domaine de $\mathbf{L}$ dans $\mathbf{L}^2(\mu)$ contient une algèbre de fonctions bornées $\mathcal{A}$ comme sous-espace dense stable par $\mathbf{L}$ et par l'action des fonctions $\mathcal{C}^\infty$. $\mathcal{A}$ est de plus supposée satisfaire à une condition technique de convergence pour la topologie du domaine. Ces conditions sont bien sûr remplies lorsque $\mathbf{L}$ est un générateur de MARKOV sur un ensemble fini ou une diffusion sur une variété compacte. Les algèbres $\mathcal{A}$ sont alors respectivement l'algèbre de toutes les fonctions et l'algèbre des fonctions $\mathcal{C}^\infty$.

Soit donc $\mathbf{L}$ le générateur d'un semi-groupe markovien sur $\mathbf{L}^2(\mu)$ de domaine $\mathcal{D}$. On lui associe canoniquement un opérateur bilinéaire $\mathbf{\Gamma}(.,.)$ sur l'algèbre $\mathcal{A}$,

$$\mathbf{\Gamma}(f,g) = 1/2(\mathbf{L}(fg) - f\mathbf{L}(g) - g\mathbf{L}(f)).$$

$\mathbf{\Gamma}(.,.)$ est dit *carré du champ* de $\mathbf{L}$.

Lorsque l'espace ambiant est une variété M et que $\mathbf{L}$ est un opérateur différentiel du second ordre elliptique sans terme constant, on peut munir la variété d'une structure riemannienne $(M,(g_{ij}))$ définie à partir des coefficients d'ordre 2 de l'opérateur pour laquelle $\mathbf{L}$ s'écrit $\mathbf{L}=\Delta+X$ où X est un champ de vecteurs. Les opérateurs réversibles pour les mesures de densité $\exp\Phi$, $\Phi\in\mathcal{C}^\infty(M)$, par rapport à la mesure riemannienne sont tels que $X=\nabla\Phi$. Dans ce cadre, le carré du champ s'écrit $\mathbf{\Gamma}(f,g)=\langle\nabla f,\nabla g\rangle$ i.e. $\mathbf{\Gamma}(f)=\mathbf{\Gamma}(f,f)=|\nabla f|^2$.

De la même manière, on construit l'opérateur $\mathbf{\Gamma_2}(.,.)$ à l'aide du générateur $\mathbf{L}$ et du carré du champ $\mathbf{\Gamma_2}(f,g)=1/2(\mathbf{L\Gamma}(f,g)-\mathbf{\Gamma}(f,\mathbf{L}g)-\mathbf{\Gamma}(\mathbf{L}f,g))$. Cet opérateur apparaît de façon naturelle lorsque l'on s'intéresse aux variations de certaines quantités englobant le semi-groupe $\mathbf{P}_t$, son générateur et son carré du champ. Plus particulièrement, on peut être amené à désirer une relation de commutation entre $\mathbf{\Gamma}$ et $\mathbf{P}_t$. La relation

$$\mathbf{\Gamma}(\mathbf{P}_tf)\leq e^{-2\rho t}\mathbf{P}_t(\mathbf{\Gamma}(f)) \tag{21}$$

est équivalente à la croissance de la fonction $F(s)=e^{-2\rho s}\mathbf{P}_s(\mathbf{\Gamma}(\mathbf{P}_{t-s}f))$ entre 0 et t. La positivité de $F'(s)$ n'est autre que la relation

$$\mathbf{\Gamma_2}(\mathbf{P}_{t-s}f)\geq\rho\mathbf{\Gamma}(\mathbf{P}_{t-s}f).$$

La condition

$$\mathbf{\Gamma_2}\geq\rho\mathbf{\Gamma},$$

pour un $\rho\in\mathbb{R}$ non forcément positif, sera dite *inégalité de courbure ρ et de dimension infinie* $\mathrm{CD}(\rho,\infty)$.

Dans le cas générique où $\mathbf{L}=\Delta+\nabla\Phi$ sur une variété riemannienne, l'inégalité $\mathrm{CD}(\rho,\infty)$ se transcrit dans son aspect géométrique par la

Proposition 1.2.1.1. — *(cf.* [Bak94] *Sect. 6) Soient Φ une fonction $\mathcal{C}^\infty$ sur une variété riemannienne (M,g) et $\mathbf{L}=\Delta+\nabla\Phi$ le générateur de* Markov *associé. L'inégalité CD(ρ,∞) pour $\mathbf{L}$ est équivalente à la condition*

$$\mathrm{Ric}(\mathbf{L})=\mathrm{Ric}(M)-\nabla\nabla\Phi\geq\rho g$$

en tant que formes quadratiques. Ici $\mathrm{Ric}(M)$ *est la courbure de* Ricci *et $\nabla\nabla\Phi$ la hessienne de Φ.*

Cela signifie en fait que la plus petite valeur propre $\rho_0(x)$ de la forme quadratique $\mathrm{Ric}(\mathbf{L})(x)$ est supérieure à ρ.

1.2.2. L'inégalité pseudo-isopérimétrique de Bakry-Ledoux. —

Dans [BL96a], Bakry et Ledoux déduisent de l'inégalité de log-Sobolev, une inégalité fonctionnelle *pseudo-isopérimétrique* au sens où elle fait intervenir la racine

du carré du champ. Rappelons qu'une inégalité de SOBOLEV logarithmique tendue de constante B équivaut à l'hypercontractivité

$$\|\mathbf{P}_s f\|_{q(s)} \leqslant \|f\|_p$$

pour $q(s) = 1 + (p-1)\, e^{4\,s/B} > p$. Pour $s = t/2$, si on impose $q(s) = 2$, on a $p = p(t) = 1 + e^{-(2\,t)/B}$. $\beta \overset{\text{déf.}}{=} 2/B$ sera dite constante d'hypercontractivité dans la suite.

Théorème 1.2.1. — *(cf.* [BL96a], *Thm* 4.1*) Soit* $\mathbf{L}$ *un générateur de* MARKOV *symétrique par rapport à* μ *de constante d'hypercontractivité* $\beta > 0$ *et de courbure* $\rho \in \mathbb{R}$. *Soit* $\mathcal{A}$ *l'algèbre de fonctions introduite plus haut.*
Alors, pour toute f *dans* $\mathcal{A}$ *à valeurs dans* $[0,1]$ *et tout* $0 \leq t \leq 1$,

$$\|f\|_2^2 - \|f\|_{p(t)}^2 \leq C\sqrt{t} \int \sqrt{\mathbf{\Gamma}(f)}\, d\mu \tag{22}$$

où $C > 0$ *ne dépend que de* ρ *et* $p(t) = 1 + e^{-\beta t}$.

Expliquons d'où vient cette inégalité. Comme précisé dans l'introduction, le point clé est le caractère régularisant du semi-groupe, pour $0 \leqslant t \leqslant 1$,

$$\left\| \sqrt{\mathbf{\Gamma}(\mathbf{P}_t f)} \right\|_\infty \leqslant \frac{c}{\sqrt{t}} \|f\|_\infty \tag{23}$$

pour une constante $c < \infty$ ne dépendant que de ρ. Ce contrôle s'obtient directement grâce à (21) puisque

$$\|f\|_\infty^2 \geqslant \mathbf{P}_t(f^2) - (\mathbf{P}_t f)^2 = \int_0^t \mathbf{P}_s(\mathbf{\Gamma}(\mathbf{P}_{t-s} f))\, ds,$$

et qu'on a l'équivalence $\frac{2\,\rho}{e^{2\,\rho\,t}-1} \sim_0 \frac{1}{t}$.

Maintenant, on obtient aisément (22) grâce à l'inégalité de CAUCHY-SCHWARZ pour $\mathbf{\Gamma}(.,.)$:

$$\begin{aligned}\|f\|_2^2 - \left\|\mathbf{P}_{t/2} f\right\|_2^2 &= \int f\,(f - \mathbf{P}_t f)\, d\mu = -\int_0^t ds \int f\, \mathbf{L}\mathbf{P}_s f\, d\mu = \int_0^t ds \int \mathbf{\Gamma}(\mathbf{P}_s f, f)\, d\mu \\ &\leqslant \int_0^t ds \int \sqrt{\mathbf{\Gamma}(\mathbf{P}_s f)}\sqrt{\mathbf{\Gamma}(f)}\, d\mu \leqslant c\|f\|_\infty \int_0^t \frac{ds}{\sqrt{s}} \int \sqrt{\mathbf{\Gamma}(f)}\, d\mu.\end{aligned}$$

BAKRY et LEDOUX soulignent le caractère isopérimétrique du précédent théorème en en déduisant le

Corollaire 1.2.2. — *Soit* $\mathbf{L}$ *un générateur symétrique, de mesure réversible* $d\mu = \exp \Phi dm$, *de courbure* $\rho \in \mathbb{R}$ *et de constante de log-*SOBOLEV $B < \infty$. *Alors, il existe une constante* C *ne dépendant que de* ρ *et* B *telle que, pour tout* A *borélien,*

$$\mathcal{U}\left(\mu(A)\right) \leqslant C\, \mu_s(\partial A)$$

Remarque 1.2.3. —

– La preuve du corollaire 1.2.2 est basée sur l'utilisation du théorème 1.2.1 pour déduire une inégalité de la forme

$$\mu(A)[1-\mu(A)^{(2/p(t))-1}] \leqslant C\sqrt{t}\,\mu_s(\partial A), \tag{24}$$

pour tout borélien A. Nous obtiendrons au paragraphe 1.2.3 une inégalité semblable, mais fonctionnelle cette fois, qui nous fournira la démonstration du théorème 1.2.4.

– Une lecture attentive de la preuve du théorème 1.2.1 permet de se convaincre que la propriété de diffusion imposée au générateur par Bakry et Ledoux n'est pas nécessaire. La propriété de symétrie est en revanche cruciale.

D'après le récent travail de Barthe et Maurey (cf. [BM00]), l'inégalité isopérimétrique ensembliste entraîne, pour un générateur de diffusion, l'inégalité isopérimétrique fonctionnelle. Le corollaire 1.2.2 assure donc de l'existence d'une inégalité isopérimétrique fonctionnelle pour toute diffusion hypercontractive sur une variété riemannienne. Nous en donnons maintenant une preuve entièrement fonctionnelle (ne faisant pas appel à la formule de co-aire).

1.2.3. Hypercontractivité et isopérimétrie pour les diffusions. — Pour la mesure gaussienne, S.Bobkov a déduit de l'isopérimétrie ensembliste l'inégalité isopérimétrique fonctionnelle (cf. [Bob97]). Sa démonstration était fondée sur une tensorisation par une mesure gaussienne en dimension un et par l'application de l'inégalité ensembliste à un ensemble du produit judicieusement choisi. Cette idée est à la base de notre résultat principal :

Théorème 1.2.4. — *Soit* $\mathbf{L}$ *un générateur de diffusion hypercontractif de mesure réversible* μ *sur une variété riemannienne compacte. Soit* $\rho \in \mathbb{R}$ *la courbure et* B *la constante de log-*Sobolev *de* $\mathbf{L}$. *Alors, il existe une constante* c *ne dépendant que de* ρ *et* B, *telle que, pour toute* $f \in \mathcal{C}^\infty$ *à valeurs dans* $[0,1]$,

$$\mathcal{U}\left(\int f d\mu\right) \leq \int \sqrt{\mathcal{U}^2(f) + c^2\,|\nabla f|^2}\,d\mu, \tag{25}$$

Remarque 1.2.5. — L'hypothèse de courbure minorée est cruciale. On peut trouver dans [BL96a] un exemple de générateur hypercontractif qui ne vérifie pas (25).

La preuve de ce théorème consiste en deux étapes. Tout d'abord, nous montrerons le

Lemme 1.2.6. — *Sous les mêmes hypothèses que le théorème 1.2.4, il existe une constante* $c > 0$ *telle que*

$$\mathcal{U}\left(\int f d\mu\right) \leq c \int \sqrt{\mathcal{U}^2(f) + |\nabla f|^2}\,d\mu, \tag{26}$$

pour toute $f \in \mathcal{C}^\infty$ *à valeurs dans* $[0,1]$.

L'inégalité (26) présente un défaut majeur : ce n'est pas une inégalité tendue (i.e. les fonctions constantes ne réalisent pas l'égalité). Obtenir une inégalité tendue présente une grande importance pour l'étude des propriétés ergodiques du semi-groupe. En effet, pour un générateur satisfaisant à une inégalité tendue, les seules fonctions invariantes sont les constantes (cf. [Bak94], proposition 2.2) ce qui assure l'ergodicité du semi-groupe. Mais cette faiblesse du lemme 1.2.6 peut être corrigée facilement grâce à un argument déjà utilisé par Bakry et Ledoux dans [BL96a] pour déduire de l'inégalité isopérimétrique $\mathcal{U}\left(\int f d\mu\right) \leq \int \sqrt{\mathcal{U}^2(f) + c\,|\nabla f|^2} d\mu$ une inégalité de log-Sobolev et plus récemment par Barthe et Maurey (cf. [BM00]) pour établir l'équivalence des inégalités isopérimétriques ensembliste et fonctionnelle pour les mesures à densité sur $\mathbb{R}^n$. Ceci consiste à tirer de l'inégalité d'origine une régularité de la loi de f sous μ (un contrôle de la norme de Lipschitz de son transport canonique en une mesure gaussienne) et à en déduire l'inégalité voulue (log-Sobolev ou isopérimétrique) pour μ grâce à l'inégalité de même forme pour la mesure gaussienne appliquée à la variable aléatoire canonique sur $(\mathbb{R}, \gamma)$ de même loi que f. Ceci permet de remarquer un comportement intéressant des inégalités isopérimétriques gaussiennes fonctionnelles : l'inégalité non tendue entraîne l'inégalité tendue. C'est ce qu'exprime le

Lemme 1.2.7. — *Soit* $\mathbf{L}$ *un générateur de* Markov *sur une variété riemannienne compacte* M *et* μ *une mesure sur* M *telle que, pour toute fonction* $\mathcal{C}^\infty$ f *sur* M *à valeurs dans* $[0,1]$ *,*

$$\mathcal{U}\left(\int f d\mu\right) \leq c \int \sqrt{\mathcal{U}^2(f) + \mathbf{\Gamma}(f)} d\mu$$

alors l'inégalité tendue (pour la même constante c) a également lieu i.e.

$$\mathcal{U}\left(\int f d\mu\right) \leq \int \sqrt{\mathcal{U}^2(f) + c^2\,\mathbf{\Gamma}(f)} d\mu$$

Preuve du lemme 1.2.6. — L'idée est d'appliquer l'inégalité (22) sur $M \times \mathbb{R}$ muni de la mesure produit $\mu \otimes \gamma$ au générateur $\overline{\mathbf{L}} = \mathbf{L} + \mathbf{L}_{O.U.}$ produit de L et d'un générateur d'Ornstein-Uhlenbeck sur $\mathbb{R}$.

Rappelons que le semi-groupe d'Ornstein-Uhlenbeck (cf. [Bak94]) admet la mesure gaussienne γ pour mesure réversible, est de courbure 1 et de constante de log-Sobolev 2. $\overline{\mathbf{L}}$ est donc de courbure $\overline{\rho} \geqslant \min(1, \rho)$ et de constante de log-Sobolev $\overline{B} \leqslant \max(2, B)$. Le théorème 1.2.1 conduit alors à l'inégalité

$$\|g\|^2_{2,\mu\otimes\gamma} - \|g\|^2_{p(t),\mu\otimes\gamma} \leq \overline{C}\sqrt{t} \int \sqrt{\overline{\mathbf{\Gamma}}(f)}\, d\mu \otimes \gamma \tag{27}$$

que nous appliquons à la fonction $g(x,s) = \mathbb{1}_{\{\varphi(s) \leqslant f(x)\}}$, où φ désigne la fonction de répartition de la mesure gaussienne. g étant une indicatrice,

$$\|g\|_2^2 = \|g\|_p^p = \mu \otimes \gamma(\varphi(s) \leqslant f(x)) = \mu(f),$$

où $\mu(f)$ désigne l'intégrale de f par μ. Et d'autre part, $\overline{\mathbf{L}}$ étant un générateur de diffusion,

$$\sqrt{\overline{\mathbf{\Gamma}}(g)}(x,s) = \big|\overline{\nabla} g\big|(x,s) = \delta_{f(x)}(\varphi(s))\,\sqrt{(\varphi'(s))^2 + |\nabla f|^2(x)}, \tag{28}$$

d'après la règle de la chaîne. On en déduit aisément que

$$\int \gamma(ds)\big|\overline{\nabla} g(x,s)\big| = \sqrt{\mathcal{U}^2(f(x)) + |\nabla f|^2(x)}$$

et (27) se réduit finalement à

$$\mu(f)[1-\mu(f)^{(2/p(t))-1}] \leqslant \overline{C}\,\sqrt{t}\,\mu\left(\sqrt{|\nabla f|^2 + \mathcal{U}^2(f)}\right). \tag{29}$$

L'argument qui suit a été utilisé par Bakry et Ledoux dans ([BL96a]) pour la preuve du corollaire 1.2.2. On peut tout d'abord supposer que $\mu(f) > 0$, l'inégalité étant triviale sinon, puis que $0 < \mu(f) < 1/2$ par symétrie de $\mathcal{U}$ par rapport à $1/2$. On s'appuie alors sur le fait que, pour tout $t \in [0,1]$,

$$(2/p(t)) - 1 \geqslant 1/2\,(\beta e^{-\beta})\,t$$

(rappelons que $\beta = 2/\overline{B}$ et $p(t) = 1 + e^{-\beta t}$). Les deux termes s'annulant en 0, cela découle de l'inégalité sur les dérivées qui est elle-même assurée puisque, pour $t \leqslant 1$, $e^{-\beta t} \geqslant e^{-\beta}$ et, comme $t \geqslant 0$, $1 + e^{-\beta t} \leqslant 2$. Posons $\lambda = 1/2\,(\beta e^{-\beta})$. L'inégalité (29) conduit alors à

$$\mu(f)\,\frac{1-\exp\left(-\lambda\,\log\left(\frac{1}{\mu(f)}\right)t\right)}{\sqrt{t}} \leqslant \overline{C}\,\mu\left(\sqrt{|\nabla f|^2 + \mathcal{U}^2(f)}\right). \tag{30}$$

Mais maintenant

$$\frac{1-\exp\left(-\lambda\,\log\left(\frac{1}{\mu(f)}\right)t\right)}{\sqrt{t}} \geqslant \lambda\,\log\left(\frac{1}{\mu(f)}\right)\sqrt{t}\,\exp\left(-\lambda\,\log\left(\frac{1}{\mu(f)}\right)t\right)$$

et, optimisant ce dernier terme, on voit que le meilleur comportement déduit de (30) est en $\mu(f)\sqrt{\log\frac{1}{\mu(f)}}$. Ce comportement est atteint pour tout $t = (\alpha\,\log(1/\mu(f)))^{-1}$. Nous n'avons plus qu'à fixer α pour que t soit inférieur à 1 (on peut prendre $\alpha = 2$). D'où

$$\mu(f)\sqrt{\log\frac{1}{\mu(f)}} \leqslant C'\,\mu\left(\sqrt{|\nabla f|^2 + \mathcal{U}^2(f)}\right)$$

pour une constante C' ne dépendant que de la courbure et de la constante de Sobolev-logarithmique. Vue l'équivalence, en 0, $\mathcal{U}(x) \sim x\,\sqrt{2\log\frac{1}{x}}$, cette inégalité se résume à l'inegalité isopérimétrique (26) pour une constante c ne dépendant toujours que de $\overline{\rho}$ et $\overline{B}$. □

Remarque 1.2.8. — Il faudrait bien sûr en toute rigueur remplacer les indicatrices $\mathbb{1}_{\{\varphi(s)\leqslant f(x)\}}$ considérées par des approximations $\psi_\varepsilon(\varphi(s)-f(x))$ pour une famille $(\psi_\varepsilon)_\varepsilon$ de fonctions $\mathcal{C}^\infty(\mathbb{R})$ à valeurs dans $[0,1]$ convergeant presque sûrement vers $\mathbb{1}_{]-\infty,0]}$ et telle que $(|{\psi_\varepsilon}'|)_\varepsilon$ converge faiblement vers δ_0, la mesure de DIRAC en 0. On peut l'obtenir en considérant une fonction $\varphi \in \mathcal{C}^\infty(\mathbb{R})$ à support dans $[0,2]$ et d'intégrale 1 et en posant $\varphi_\varepsilon(x) = 1/\varepsilon\,\varphi(x/\varepsilon)$, puis $\psi_\varepsilon(x) = 1 - \int_{-\infty}^x \varphi_\varepsilon(u)du$.

Preuve du lemme 1.2.7. — Soit $f \in \mathcal{C}^\infty$ sur M à valeurs dans $[0,1]$. Quitte à se placer sur $M \times \mathbb{R}$, $\mathbb{R}$ étant muni de la mesure gaussienne, (l'inégalité non tendue se tensorisant avec la gaussienne !) et à considérer

$$f_\varepsilon(x,u) = (1-\varepsilon^2)^{\frac{1}{2}} f(x) + \varepsilon u,$$

on peut, comme le font remarquer BAKRY et LEDOUX, supposer que f est à densité strictement positive $F'(x)$, où F est la fonction de répartition de f (voir le paragraphe 1.3.3 pour de plus amples détails).

Nous appliquons l'inégalité (26) aux composées de f et des translations $\psi_{\varepsilon,r} = \psi_\varepsilon \circ \tau_r$ de la fonction ψ_ε introduite précédemment. Lorsque ε tend vers 0, vu que $|\psi'_{\varepsilon,r}|$ converge faiblement vers δ_r, on obtient

$$\mathcal{U}(F(r)) \leqslant c\,\theta(r)F'(r) \tag{31}$$

où $\theta(r) = \mathbb{E}(\sqrt{\mathbf{\Gamma}(f)}/f = r)$ est une version de l'espérance conditionnelle de $\sqrt{\mathbf{\Gamma}(f)}$ sachant f. Or, il est montré dans [BM00] comment déduire de cette inégalité une inégalité isopérimétrique tendue. Posant $k = F^{-1} \circ \varphi$, k est une variable aléatoire sur $(\mathbb{R},\gamma)$ de même loi que f et l'inégalité (31) nous assure que, pour tout $s \in \mathbb{R}$, $k'(s) \leqslant c\,\theta(k(s))$. Appliquant à k l'inégalité isopérimétrique (20) vérifiée pour la gaussienne et se souvenant que la loi de k n'est autre que la loi de f, on en déduit que

$$\mathcal{U}(\int f d\mu) \leqslant \int \sqrt{(\mathbb{E}(\mathcal{U}(f)/f))^2 + \left(\mathbb{E}(c\sqrt{\mathbf{\Gamma}(f)}/f)\right)^2}\, d\mu$$

Suivant BARTHE et MAUREY, nous concluons grâce à l'inégalité de MINKOWSKI pour les espérances conditionnelles

$$\sqrt{\mathbb{E}(X)^2 + \mathbb{E}(Y)^2} \leqslant \mathbb{E}(\sqrt{X^2+Y^2}).$$

□

1.2.4. Hypercontractivité et isopérimétrie fonctionnelle sur un ensemble fini. — La propriété de diffusion $\mathbf{\Gamma}(\psi(f)) = (\psi'(f))^2\,\mathbf{\Gamma}(f)$ utilisée précédemment n'est pas réalisée lorsque les processus étudiés ne prennent qu'un nombre discret de valeurs. C'est en particulier le cas lorsqu'on s'intéresse aux propriétés ergodiques d'une chaîne de MARKOV à valeurs dans un ensemble fini. Néanmoins, la méthode exposée dans le cas continu pour obtenir l'isopérimétrie gaussienne à partir de l'inégalité de log-SOBOLEV est en partie transposable au cas discret. Elle fournit une

inégalité fonctionnelle du type isopérimétrique d'une forme un peu différente, généralement plus faible que l'inégalité de BOBKOV mais suffisante pour engendrer une inégalité isopérimétrique ensembliste.

Soit F un ensemble fini, μ une mesure de probabilités sur F et soit $(\alpha^y)_{y\in F}$ une famille de fonctions positives sur F. Le générateur de MARKOV défini par

$$\mathbf{L}f(x) = \langle \alpha, \nabla f\rangle(x) := \sum_{y\in F} \alpha^y(x)\nabla^y f(x) \tag{32}$$

est supposé réversible par rapport à μ (cf. paragraphe 1.4.1 pour de plus amples détails et une condition de réversibilité). Le carré du champ de $\mathbf{L}$ n'est autre que

$$2\,\mathbf{\Gamma}(f)(x) = \langle \alpha, |\nabla f|^2\rangle(x) = \sum_{y\in F} \alpha^y(x)|f(x)-f(y)|^2.$$

1.2.4.1. *Inégalité du type isopérimétrie fonctionnelle.* — Introduisons quelques notations. Soient f une fonction sur F, m une valeur de f et $x \in F$. Nous considérerons la quantité $\beta(f,x,m)$ définie comme suit. Si $m < f(x)$, $\beta(f,x,m)$ est la différence des racines carrées de

$$\sum_{\leqslant m} \overset{\text{déf.}}{=} 1/2 \sum_{y\in F; f(y)\leqslant m} \alpha^y(x) \quad \text{et de} \quad \sum_{<m} \overset{\text{déf.}}{=} 1/2 \sum_{y\in F; f(y)<m} \alpha^y(x).$$

Symétriquement, si $m > f(x)$, $\beta(f,x,m)$ vaut $\left(\sqrt{\sum_{\geqslant m}} - \sqrt{\sum_{>m}}\right)$. Nous pouvons maintenant énoncer le

Théorème 1.2.9. — *Soit* $\mathbf{L}$ *un générateur sur* F *réversible par rapport à* μ *de courbure* ρ *vérifiant une inégalité de log-*SOBOLEV *de constante* $B < \infty$ *i.e. que, pour toute fonction* f *sur* F,

$$\mathbf{Ent}_\mu(f^2) \leqslant B \int \mathbf{\Gamma}(f)\, d\mu$$

alors $\mathbf{L}$ *vérifie une inégalité isopérimétrique gaussienne de la forme*

$$\mathcal{U}\left(\int f d\mu\right) \leq c \int \big(\mathcal{U}(f) + \Xi(f)\big) d\mu, \tag{33}$$

pour toute fonction f *sur* F *à valeurs dans* $[0,1]$, *où la constante* c *ne dépend que de* ρ *et* B, *et où*

$$\Xi\,(f)\,(x) = \sum_{m\in f(F)} |f(x)-m|\beta(f,x,m).$$

Remarque 1.2.10. — Rappelons que

$$\mathbf{Ent}_\mu(f^2) = \int f^2 \log(f^2) d\mu - \int f^2 d\mu \, \log\Big(\int f^2 d\mu\Big).$$

Ce résultat est un résultat intermédiaire duquel découlerons dans la suite des inégalités du type isopérimétrique, l'une ensembliste, l'autre fonctionnelle.

Preuve

◀ On considère comme dans le cas continu le générateur produit $\overline{\mathbf{L}}$ sur l'espace $(F \times \mathbb{R}, \mu \otimes \gamma)$ de $\mathbf{L}$ et d'un générateur d'Ornstein-Uhlenbeck. Comme dans le cas diffusif, $\overline{\mathbf{L}}$ est de constante de log-Sobolev $\overline{B} \leqslant \max(2, B)$ et de courbure $\overline{\rho} \geqslant \min(1, \rho)$ (pour se convaincre que la courbure se tensorise, on peut par exemple voir [Bak97] où il est montré que la condition $\mathbf{\Gamma_2}$ est équivalente à l'inégalité de trou-spectral pour les noyaux définis par le semi-groupe ; il est bien connu que cette inégalité se tensorise). On applique comme précédemment l'inégalité *pseudo-isopérimétrique* (22) du théorème 1.2.1 à la fonction $g(x,s) = \mathbb{1}_{\{\varphi(s) \leqslant f(x)\}}$. On a toujours $\|g\|_2^2 - \|g\|_{p(t)}^2 = \mu(f)[1 - \mu(f)^{(2/p(t))-1}]$. La difficulté nouvelle réside dans le terme $\int \sqrt{\overline{\mathbf{\Gamma}}(g)}\, d\mu \otimes \gamma$. Le point important est que la formule de la chaîne (28) n'est plus valable dans ce cadre. Les carrés du champ en s et x se comportent de manière différente. D'une part,

$$\mathbf{\Gamma}_s(g)(x,s) = \delta_{f(x)}^2(\varphi(s))\,(\varphi')^2(s)$$

et d'autre part

$$\mathbf{\Gamma}_x(g)(x,s) = \frac{1}{2} \sum_{y \neq x} \alpha^y(x)\, \mathbb{1}_{]f(x) \wedge f(y), f(x) \vee f(y)]} \circ \varphi(s).$$

La notation $\delta_{f(x)}^2$ est à comprendre comme le carré des dérivées $\psi_\varepsilon{}'$ des fonctions $\mathcal{C}^\infty$ approximant $\mathbb{1}_{]-\infty, f(x)]}$ comme dans la remarque 1.2.8. Nous resterons ici à ce niveau formel pour ne pas alourdir les notations. Nous affirmons que

$$\int \sqrt{\mathbf{\Gamma}_s(g) + \mathbf{\Gamma}_x(g)}\, d\gamma = \mathcal{U}(f(x)) + \Xi\,(f)\,(x). \tag{34}$$

Tout d'abord, l'intégrale de la racine du carré du champ produit n'est autre que

$$\int_0^1 \sqrt{\mathbf{\Gamma}_{cont}(u) + \mathbf{\Gamma}_{disc}(u)}\, du$$

pour, d'une part,

$$\mathbf{\Gamma}_{cont}(u) \overset{\text{déf.}}{=} \delta_{f(x)}^2(u)\, \mathcal{U}^2(u)$$

et d'autre part

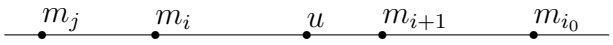
$$\mathbf{\Gamma}_{disc}(u) \overset{\text{déf.}}{=} 1/2 \sum_{y \neq x} \alpha^y(x)\, \mathbb{1}_{]f(x) \wedge f(y), f(x) \vee f(y)]}(u).$$

Nous scindons alors l'intégrale suivant les intervalles formés par les différentes valeurs de f, $m_1 < m_2 < \ldots < m_k$. Remarquons que $\delta_{f(x)}^2$ étant localisée en $f(x)$, $\mathbf{\Gamma}_{cont}(u)$

n'apporte de contribution qu'au voisinage de $m_{i_0} = f(x)$. Soit $i < i_0$. Sur $[m_i, m_{i+1}[$, l'intégrale se réduit alors à

$$\int_{m_i}^{m_{i+1}} \sqrt{1/2 \sum_{j<i_0} \alpha_j \, \mathbb{1}_{[m_j, m_{i_0}]}(u)} \, du = (m_{i+1} - m_i) \sqrt{1/2 \sum_{j \leqslant i} \alpha_j}$$
$$\overset{\text{déf.}}{=} (m_{i+1} - m_i) \sqrt{\sum_{\leqslant m_i}}$$

où l'on a posé $\alpha_j = \sum_{y \in F; f(y) = m_j} \alpha^y(x)$. Ceci découle du fait que, si $u \in [m_i, m_{i+1}[$, alors $u \in [m_j, m_{i_0}[$ ssi $1 \leqslant j \leqslant i$.

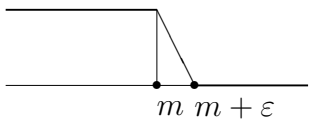


La contribution sur $]m_i, m_{i+1}]$ pour $i \geqslant i_0$ se traite de manière similaire (pour $]m_{i_0}, m_{i_0+1}]$ il faut plutôt considérer l'intervalle $]m_{i_0} + \varepsilon, m_{i_0+1}]$ et faire tendre ε vers 0).

La contribution en $f(x)$ (ou plutôt sur $[f(x), f(x) + \varepsilon]$) est différente. Il s'agit d'évaluer

$$\int_{\textit{Voisinage de } f(x)} \sqrt{\delta^2_{f(x)}(u) \mathcal{U}^2(u) + \textit{cte}} \, du.$$

Retranscrit pour des fonctions $\psi_\varepsilon(u) = \mathbb{1}_{\{u<m\}} + (-\frac{1}{\varepsilon}(u - m) + 1)\mathbb{1}_{]m, m+\varepsilon[}(u)$ approximant $\mathbb{1}_{\{u \leqslant m\}}$,



cela donne

$$\int_m^{m+\varepsilon} \sqrt{(\psi'_\varepsilon)^2(u) \mathcal{U}^2(u) + \textit{cte}} \, du = \frac{1}{\varepsilon} \int_m^{m+\varepsilon} \sqrt{\mathcal{U}^2(u) + \varepsilon^2 \, \textit{cte}} \, du \xrightarrow{\varepsilon \to 0} \mathcal{U}(m).$$

Regroupant toutes les contributions, on obtient

$$\int \sqrt{\overline{\mathbf{\Gamma}}(g)} \, d\mu \otimes \gamma = \mathcal{U}(f(x)) + \sum_{i<i_0} (m_{i+1} - m_i) \sqrt{\sum_{\leqslant m_i}} + \sum_{i \geqslant i_0} (m_{i+1} - m_i) \sqrt{\sum_{\geqslant m_i}}$$

Une intégration par partie d'Abel conduit alors à (34). On obtient finalement une inégalité de la forme

$$\mu(f)[1 - \mu(f)^{(2/p(t))-1}] \leqslant C \sqrt{t} \, \mu \left(\mathcal{U}(f) + \Xi(f) \right).$$

L'argument développé dans la preuve du théorème 1.2.4 pour déduire d'une inégalité de cette forme l'isopérimétrie fonctionnelle s'applique encore pour donner l'inégalité (33). ►

1.2.4.2. *Comparaison avec l'isopérimétrie gaussienne de* BOBKOV. — Le terme $\Xi(f)$ de l'inégalité (33) fait intervenir une somme sur les différentes valeurs de la fonction. On peut, en considérant plutôt une somme sur les points de F, obtenir une inégalité comparable avec le pendant discret de l'inégalité isopérimétrique (20). C'est ce à quoi nous nous attelons maintenant.

Choisissons une numérotation des éléments de F (i.e. $F = \{x_i / i = 1, \dots, q\}$) qui soit compatible avec la fonction f au sens où les $f_i := f(x_i)$ vérifient $f_i \leqslant f_{i+1}$ (les f_i ne sont pas forcément distinctes). Notons $\alpha_i^j = \alpha^{x_j}(x_i)$. Facilement, notant

$$h(j) = \begin{cases} \sum_{1 \leqslant l \leqslant j} \alpha_i^l & \text{si } 1 \leqslant j < i \\ \sum_{j+1 \leqslant l \leqslant q} \alpha_i^l & \text{si } q > j \geqslant i \end{cases}$$

et $h(0) = h(q) = 0$, on voit que

$$\Xi\,(f)\,(x_i) = \sum_{j=1,\dots,q} |f_i - f_j| \left| \sqrt{h(j)} - \sqrt{h(j-1)} \right|.$$

Remarque 1.2.11. — Pour une indicatrice (ou pour toute fonction à deux valeurs), $\Xi(f)$ coïncide avec $\sqrt{\mathbf{\Gamma}(f)}$ et donc l'inégalité fonctionnelle que nous avons obtenue permet de retrouver l'inégalité isopérimétrique ensembliste du corollaire 1.2.2.

Dans l'optique de la mécanique statistique, la question qui se pose est de savoir si $\Xi(f)$ se compare à $\sqrt{\mathbf{\Gamma}(f)}$ indépendamment du cardinal de F. On peut répondre par la négative à cette question. Cela provient de la remarque suivante. On peut tout d'abord se ramener à ne comparer ces quantités qu'en un point x_i où f atteint son maximum. Si l'on pose alors $u_k = |f_i - f_k| = f_i - f_k$ et $v_k = h(k)$, on est ramené à comparer, pour toute suite $(u_k)_{k \in \mathbb{N}} \geqslant 0$ décroissante et toute suite $(v_k)_{k \in \mathbb{N}} \geqslant 0$ telle que $v_0 = 0$, les deux quantités

$$\left(\sum_{k=1}^{n} u_k \left(\sqrt{v_k} - \sqrt{v_{k-1}} \right) \right)^2 \tag{35}$$

et

$$\sum_{k=1}^{n} u_k^2 \left(v_k - v_{k-1} \right), \tag{36}$$

indépendamment de n.

Tout d'abord, il n'est pas très difficile de montrer que (36) est inférieur à (35). En effet, développant le carré dans (35),

$$\left(\sum_{k=1}^{n} u_k\left(\sqrt{v_k}-\sqrt{v_{k-1}}\right)\right)^2 = \sum_{k=1}^{n} u_k^2\left(\sqrt{v_k}-\sqrt{v_{k-1}}\right)^2 + 2\sum_{k=1}^{n} u_k\left(\sqrt{v_k}-\sqrt{v_{k-1}}\right)\sum_{1\leqslant l<k} u_l\left(\sqrt{v_l}-\sqrt{v_{l-1}}\right)$$

que l'on peut minorer aisément par

$$\sum_{k=1}^{n} u_k^2\left\{\left(\sqrt{v_k}-\sqrt{v_{k-1}}\right)^2 + 2\left(\sqrt{v_k}-\sqrt{v_{k-1}}\right)\underbrace{\sum_{1\leqslant l<k}\left(\sqrt{v_l}-\sqrt{v_{l-1}}\right)}_{=\sqrt{v_{k-1}}}\right\}$$

puisque (u_k) est décroissante. Reste alors à remarquer que, vue l'égalité $a^2+2\,a\,b=(a+b)^2-b^2$, le terme entre crochets n'est autre que v_k-v_{k-1}.

En revanche, (35) ne peut être dominé par (36). Il suffit pour s'en assurer de considérer l'exemple $u_k=\frac{1}{k}$ et $v_k=k^2$ pour lequel (35) se comporte en $(\log n)^2$ tandis que (36) est en $\log n$.

Remarquons toutefois que l'on peut contrôler (35) par (36) pour peu que les dynamiques considérées aient un nombre uniformément borné de voisins. C'est le cas en particulier des laplaciens discrets sur un réseau $\frac{\mathbb{Z}^d}{n}$. Cela découle aisément de l'inégalité de Cauchy-Schwarz : posant

$$\Upsilon\left(f\right)(x)=\langle\sqrt{\alpha},|\nabla f|\rangle(x)=\sum_{y\in F}\sqrt{\alpha^y(x)}\,|f(x)-f(y)|,$$

on a

$$\Xi(f)(x)\leqslant\Upsilon\left(f\right)(x)=\sum_{y\in Vois(x)}\sqrt{\alpha^y(x)}\,|f(x)-f(y)| \leqslant\sqrt{\mathbf{Card}(Vois(x))}\,\sqrt{\mathbf{\Gamma}(f)}(x)$$

où $Vois(x)\overset{\text{déf.}}{=}\{y/\alpha^y(x)\neq 0\}$ est l'ensemble des voisins de x.

1.2.4.3. *Tension de l'inégalité.* — L'inégalité (33) conduit à

$$\mathcal{U}\left(\int f d\mu\right)\leq c\int\big(\mathcal{U}(f)+\Upsilon(f)\big)d\mu. \tag{37}$$

Afin de tendre cette inégalité, nous suivons la démarche que nous avons déjà utilisée à cette fin dans le cas des diffusions. Notre inégalité (37) induit une régularité de la variable $k=G^{-1}\circ\varphi$ de même loi que f sur $(\mathbb{R},\gamma)$ (G désigne la fonction de répartition

de la loi de f et $G^{-1}(u) = \inf\,(t \in \mathbb{R}/G(t) \geqslant u)$ est son inverse généralisé). L'inégalité isopérimétrique pour la gaussienne appliquée à k (ou plutôt à une approximation de k) jointe à cette régularité de k donne alors l'inégalité isopérimétrique pour la loi de f :

Lemme 1.2.12. — *Soit* $\mathbf{L} = \langle \alpha, \nabla f \rangle$ *un générateur de* MARKOV *sur un ensemble fini* F *et* μ *une mesure sur* F *telle que, pour toute fonction* f *sur* F *à valeurs dans* $[0,1]$,

$$\mathcal{U}\left(\int f d\mu\right) \leq c \int \big(\mathcal{U}(f) + \Upsilon(f)\big) d\mu,$$

où $\Upsilon\,(f) = \langle \sqrt{\alpha}, |\nabla f| \rangle$. *Alors l'inégalité tendue (pour la même constante c) a également lieu i.e.*

$$\mathcal{U}\left(\int f d\mu\right) \leq \int \big(\mathcal{U}(f) + c\,\Upsilon(f)\big) d\mu$$

Finalement, on obtient le

Théorème 1.2.13. — *Soit* $\mathbf{L}$ *un générateur hypercontractif sur* F *réversible par rapport à* μ *de courbure* ρ *et de constante de log-*SOBOLEV B *comme dans le théorème 1.2.9 alors il existe une constante* $c > 0$ *ne dépendant que de* ρ *et* B *telle que*

$$\mathcal{U}\left(\int f d\mu\right) \leq \int \big(\mathcal{U}(f) + c\,\Upsilon(f)\big) d\mu, \tag{38}$$

pour toute f *à valeurs dans* $[0,1]$.

Preuve du lemme 1.2.12. — Nous aurons besoin d'un lemme combinatoire qui met en lumière le rôle générique de l'inégalité isopérimétrique pour la gaussienne.

Soient $(f_i)_{i=1,\ldots,n}$ une suite croissante de réels compris entre 0 et 1 (non forcément distincts), et $(a_i)_{i=1,\ldots,n-1}$ une suite croissante de réels et soit

$$h(s) = \sum_{i=1,\ldots,n} f_i \mathbb{1}_{[a_{i-1},a_i[}(s)$$

la fonction en escalier sur $\mathbb{R}$ correspondante (on a posé $a_0 = -\infty$ et $a_n = +\infty$). On approxime h par la fonction

$$h_\varepsilon(s) = f_1 + \sum_{i=1,\ldots,n-1} (f_{i+1} - f_i)\tilde{\psi}_{a_i,\varepsilon}(s)$$

où $\tilde{\psi}_{a,\varepsilon}(s) = \left(\frac{1}{\varepsilon}\,(s-a)+1\right) \mathbb{1}_{]a-\varepsilon,a[}(s) + \mathbb{1}_{]a,+\infty]}(s)$ approxime la fonction indicatrice de $[a,+\infty]$. Appliquant l'inégalité (20) à h_ε, on obtient, faisant tendre ε vers 0 :

Lemme 1.2.14. —

$$\mathcal{U}\left(\sum_{i=1,\dots,n} f_i(\varphi(a_i)-\varphi(a_{i-1}))\right) \leqslant \sum_{i=1,\dots,n} \mathcal{U}(f_i)(\varphi(a_i)-\varphi(a_{i-1})) + \sum_{i=1,\dots,n-1} \varphi'(a_i)|f_{i+1}-f_i|$$

ou encore

$$\mathcal{U}\left(\int h d\gamma\right) - \int \mathcal{U}(h) d\gamma \leqslant \sum_{i=1,\dots,n-1} \varphi'(a_i)|f_{i+1}-f_i|.$$

Soit maintenant f une fonction définie sur F et soit G la fonction de répartition de sa loi. Il est bien connu que, si G^{-1} est son inverse généralisé, G^{-1} définit une variable aléatoire, sur $[0,1]$ muni de la mesure de LEBESGUE, de même loi que f. Il vient que $k := G^{-1}\circ\varphi$, définie sur $(\mathbb{R},\gamma)$ a même loi que f. Or, si $(f_i)_{i=1,\dots,n}$ désignent les valeurs de f (distinctes cette fois) rangées par ordre croissant, $p_i := \mu(f = f_i)$, $p_0 = 0$ et, pour $i = 0,\dots,n$, $\beta_i = \sum_{k=0}^{i} p_k$, pour $i = 0,\dots,n$,

$$k(s) = \sum_{i=1,\dots,n} f_i \mathbb{1}_{]\varphi^{-1}(\beta_{i-1}),\varphi^{-1}(\beta_i)]}(s)$$

Appliquant le lemme 1.2.14 à k, il vient

$$\mathcal{U}\left(\int f d\mu\right) - \int \mathcal{U}(f) d\mu \leqslant \sum_{i=1}^{n-1} \mathcal{U}(\beta_i)|f_{i+1}-f_i| \tag{39}$$

Il nous reste à exploiter l'information contenue dans l'inégalité (33). Pour ce faire, nous l'appliquons à la fonction $\psi_i \circ f$ où $\psi_i := \mathbb{1}_{]-\infty,f_i]}$.

On voit aisément que $\int \psi_i \circ f\, d\mu = \sum_{j=1}^{n} \psi_i(f_j)\mu(f = f_j) = \sum_{j\leqslant i} p_j = \beta_i$ et $\mathcal{U}(\psi_i(f)) = 0$. On obtient donc

$$\begin{aligned}
\mathcal{U}(\beta_i) &\leqslant c \int \Upsilon(\psi_i(f)) d\mu \\
&\leqslant c \sum_{x,y\in F} |\psi_i(f(x)) - \psi_i(f(y))| \sqrt{\alpha^y(x)}\mu(x) \\
&\leqslant c \sum_{\substack{x\in F \\ f(x)\leqslant f_i}} \mu(x) \sum_{\substack{y\in F \\ f(y)>f_i}} \sqrt{\alpha^y(x)} + c \sum_{\substack{x\in F \\ f(x)>f_i}} \mu(x) \sum_{\substack{y\in F \\ f(y)\leqslant f_i}} \sqrt{\alpha^y(x)} \\
&\leqslant c \sum_{\substack{x,y\in F \\ f(x)\leqslant f_i < f(y)}} \left(\mu(x)\sqrt{\alpha^y(x)} + \mu(y)\sqrt{\alpha^x(y)}\right)
\end{aligned}$$

Et insérant ceci dans (39),

$$\begin{aligned}
&\mathcal{U}\left(\int f d\mu\right) - \int \mathcal{U}(f) d\mu \\
&\quad \leqslant c \sum_{i=1}^{n-1} |f_{i+1} - f_i| \sum_{\substack{x,y\in F \\ f(x)\leqslant f_i < f(y)}} \left(\mu(x)\sqrt{\alpha^y(x)} + \mu(y)\sqrt{\alpha^x(y)}\,\right) \\
&\quad \leqslant c \sum_{\substack{x,y\in F \\ f(x)<f(y)}} \left(\mu(x)\sqrt{\alpha^y(x)} + \mu(y)\sqrt{\alpha^x(y)}\,\right) \sum_{\substack{1\leqslant i\leqslant n-1 \\ f(x)\leqslant f_i<f(y)}} (f_{i+1} - f_i) \\
&\quad \leqslant c \sum_{\substack{x,y\in F \\ f(x)<f(y)}} \left(\mu(x)\sqrt{\alpha^y(x)} + \mu(y)\sqrt{\alpha^x(y)}\,\right)(f(y) - f(x)) \\
&\quad \leqslant c \sum_{x,y\in F} \mu(x)\sqrt{\alpha^y(x)}|f(y) - f(x)| \leqslant c \int \Upsilon(f)\, d\mu,
\end{aligned}$$

ce qui termine la preuve. □

1.3. Inégalité isopérimétrique pour les modèles continus en mécanique statistique

Dans les modèles de mécanique statistique, un état d'équilibre est obtenu comme limite faible de mesures dites *mesures de* GIBBS *en volume fini à conditions extérieures fixées*. Génériquement, les courbures des dynamiques associées à ces mesures sont minorées uniformément, ce qui permet de déduire du théorème 1.2.4 le contrôle uniforme de la constante d'isopérimétrie gaussienne pour ces mesures dès que celui de la constante de log-SOBOLEV est assuré, ou encore d'après [Zeg92, SZ92a, Lar95] lorsque les corrélations décroissent exponentiellement. C'est l'objet du présent paragraphe.

1.3.1. Contrôle de la courbure. —

Contrôle des constantes pour les systèmes de spins

On considère un système de spins continus sur le réseau $\mathbb{Z}^\nu$. Les spins sont décrits par une variété riemannienne compacte M et l'espace des configurations est $M^{\mathbb{Z}^\nu}$. Les interactions sont décrites par un potentiel $(\mathcal{J}_\Lambda)_{\Lambda\in\mathcal{F}}$ où $\mathcal{F}$ désigne l'ensemble des parties finies de $\mathbb{Z}^\nu$. Étant donnée une partie finie Λ, $\mathcal{J}_\Lambda$ est une fonction $\mathcal{C}^\infty(M^\Lambda)$ considérée comme une fonction sur $M^{\mathbb{Z}^\nu}$ ne dépendant que des spins sur Λ. Sur la variété M^Λ décrivant un sous-système fini, la mesure d'équilibre de conditions extérieures ξ est $\mu_\Lambda^\xi(dx_\Lambda) = (Z_\Lambda^\xi)^{-1} \exp H_\Lambda^\xi(x_\Lambda)\; m^{\otimes\Lambda}(dx_\Lambda)$, où m est la mesure riemannienne

de M et Z_Λ^ξ est une constante de normalisation. $H_\Lambda^\xi(x_\Lambda) = \sum_{\substack{F\cap\Lambda\neq\emptyset \\ F\in\mathcal{F}}} \mathcal{J}_F(x_\Lambda.\xi_{\Lambda^c})$ est l'hamiltonien associé au potentiel. Le potentiel est supposé absolument sommable i.e. que, pour tout $i \in \mathbb{Z}^\nu$,

$$\|\mathcal{J}\|_i := \sum_{\substack{A\in\mathcal{F} \\ A\ni i}} \|\mathcal{J}_A\|_\infty < \infty$$

(cf. [Geo88] p. 29).

La dynamique de Glauber en volume fini associée est le semi-groupe de générateur $\mathbf{L}_\Lambda = \Delta_\Lambda + \nabla\Phi_\Lambda$, où $\Phi_\Lambda = H_\Lambda^\xi$, de mesure réversible μ_Λ^ξ. Δ_Λ désigne l'opérateur de Laplace-Beltrami sur M^Λ. Les propriétés d'ergodicité et de vitesse de convergence de ce semi-groupe vers sa mesure d'équilibre sont décrites par une étude analytique : l'existence d'un trou spectral fournit la convergence $\mathbf{L}^2\left(\mu_\Lambda^\xi\right)$ exponentielle, celle d'une inégalité de Sobolev logarithmique des informations plus précises sur la vitesse de convergence. Le problème fondamental auquel on est confronté pour appliquer ces résultats à l'étude des mesures de Gibbs est en un certain sens de pouvoir étendre ces inégalités au système infini. Pour ce faire, on doit contrôler les constantes intervenant dans les inégalités en volume fini Λ indépendamment de la taille de Λ.

Contrôle de la courbure pour les systèmes de spins continus

E.A. Carlen et D.W. Stroock (cf. [CS86]) ont fourni une condition sur le potentiel d'interaction $(\mathcal{J}_\Lambda)_{\Lambda\in\mathcal{F}}$ garantissant le contrôle positif de la courbure $\mathbf{\Gamma_2}^{\Lambda,\xi} \geq \rho_{\Lambda,\xi}\mathbf{\Gamma}^{\Lambda,\xi}$ i.e de la plus petite valeur propre du $(0,2)$ tenseur $\mathrm{Ric}(L_\Lambda^\xi)$ sur M^Λ. M étant compacte, celle-ci est minorée sur chaque M^Λ. La condition de Carlen et Stroock avait pour but d'obtenir un contrôle de la constante de Sobolev logarithmique via un contrôle strictement positif de la courbure. Elle contient donc certaines exigences techniques dont nous nous abstiendrons. Notre condition est quelque peu plus générale. Elle traite de la régularité d'une famille $(\Phi_\Lambda)_{\Lambda\in\mathcal{F}}$ de fonctions $\mathcal{C}^\infty$. Avant de l'énoncer, rappelons qu'une matrice $(\gamma(i,j))_{i,j\in\mathbb{Z}^\nu}$ définit naturellement un opérateur sur l'espace des fonctions à support fini sur $\mathbb{Z}^\nu$ par $(\gamma f)(i) = \sum_{j\in\mathbb{Z}^\nu} \gamma(i,j)f(j), \ \ f \in \mathcal{C}_0$.

Condition $(\partial^{\mathbf{2}}(\mathbf{\Phi_\Lambda})\mathbf{l^2})$ *Il existe un opérateur borné de $l^2(\mathbb{Z}^\nu)$, $(\gamma(i,j))$, tel que, pour toute famille $(X^{(i)})_{i\in\mathbb{Z}^\nu}$ de champs de vecteurs, toute partie finie Λ de $\mathbb{Z}^\nu$ et tout couple (i,j) de sites dans $\mathbb{Z}^\nu$,*

$$\sup_{x_\Lambda\in M^\Lambda} \nabla_{(i)}\nabla_{(j)}\Phi_\Lambda(x_\Lambda)\ \left(X^{(i)}, X^{(j)}\right) \leq \gamma(i,j)\,\|X^{(i)}\|\,\|X^{(j)}\|$$

$\nabla_{(i)}\nabla_{(j)}\Phi_\Lambda$ désigne le bloc (i,j) de la hessienne de Φ_Λ.
Sous cette condition, l'argument de CARLEN-STROOCK s'applique et on obtient un contrôle de la courbure sous la forme du

Théorème 1.3.1. — *Soient $(\Phi_\Lambda)_{\Lambda\in\mathcal{F}}$ une famille de densités $\mathcal{C}^\infty$ satisfaisant à la condition $(\partial^2(\Phi_\Lambda)l^2)$ et $\mathbf{P}_t^\Lambda$ le semi-groupe associé de générateur $\mathbf{L}_\Lambda = \Delta_\Lambda + \nabla\Phi_\Lambda$ et de mesure réversible*
$\mu_\Lambda(dx_\Lambda) = (Z_\Lambda)^{-1}\exp\Phi_\Lambda(x_\Lambda)\, m^{\otimes\Lambda}(dx_\Lambda)$. Il existe un $\rho\in\mathbb{R}$ tel que, pour chaque partie finie Λ de $\mathbb{Z}^\nu$,

$$\mathbf{L}_\Lambda \text{ vérifie } CD(\rho,\infty) \qquad i.e \qquad \mathrm{Ric}(M^\Lambda) - \nabla\nabla\Phi_\Lambda \geq \rho\, g_\Lambda\,.$$

ρ ne dépend que de la norme $\|\gamma\|_{2\to 2}$ et de la plus petite valeur propre β de $\mathrm{Ric}(M)$.

Preuve. ▼ Considérant l'expression explicite de $\nabla\nabla\Phi_\Lambda$, on est plongé dans $l^2(\mathbb{Z}^\nu)$. L'inégalité de CAUCHY-SCHWARZ permet alors de conclure :

$$\nabla\nabla\Phi_\Lambda(X^\Lambda,X^\Lambda) = \sum_{i,j\in\Lambda}\nabla_{(i)}\nabla_{(j)}\,\Phi_\Lambda\,(X^{(i)},X^{(j)}) \;\leq \sum_{i,j\in\mathbb{Z}^\nu}\gamma(i,j)f(i)f(j)$$

où $f\in l^2(\mathbb{Z}^\nu)$ est donnée par

$$f(i) = \begin{cases} \|X^{(i)}\| & \text{si } i\in\Lambda \\ 0 & \text{sinon} \end{cases}.$$

D'où, $\quad \nabla\nabla\Phi_\Lambda(X^\Lambda,X^\Lambda) \;\leq \sum_{i\in\mathbb{Z}^\nu}\gamma f(i)\,f(i) = \langle\gamma f, f\rangle_{l^2(\mathbb{Z}^\nu)}$

Par CAUCHY-SCHWARZ,

$\nabla\nabla\Phi_\Lambda(X^\Lambda,X^\Lambda) \;\leq \|\gamma f\|_2\,\|f\|_2 \leq \|\gamma\|_{2\to 2}\,\|f\|_2^2 = \|\gamma\|\; g_\Lambda(X^\Lambda,X^\Lambda).$

En conséquence, $\quad -\|\gamma\|\; g_\Lambda(X^\Lambda,X^\Lambda) \leq -\nabla\nabla\Phi_\Lambda(X^\Lambda,X^\Lambda) \quad$ et, par suite,

$$\mathrm{Ric}(M^\Lambda) - \nabla\nabla\Phi_\Lambda \geq (\beta - \|\gamma\|)\, g_\Lambda$$

où $\beta\in\mathbb{R}$ est la plus petite valeur propre de $\mathrm{Ric}(M)$ relativement à g. ▲

Conditions sur le potentiel d'interaction

Les densités $(\Phi_\Lambda)_{\Lambda\in\mathcal{F}}$ sont données en mécanique statistique par le potentiel par l'intermédiaire de l'hamiltonien $H_\Lambda^\xi(x_\Lambda) = \sum_{\substack{A\cap\Lambda\neq\emptyset\\ A\in\mathcal{F}}}\mathcal{J}_A(x_\Lambda.\xi_{\Lambda^c})$.
On peut introduire diverses conditions sur le potentiel ou sur l'hamiltonien assurant que les $\Phi_\Lambda^\xi = H_\Lambda^\xi$ vérifient la condition $(\partial^2(\Phi_\Lambda)l^2)$ indépendamment de ξ. La première est inspirée de [CS86], la seconde de [Lar95].

Soit $(\mathcal{J}_\Lambda)_{\Lambda\in\mathbb{Z}^\nu}$ un potentiel d'interaction. On supposera dans la suite qu'il satisfait à l'une quelconque des deux conditions suivantes :

Condition (Pot.1) *Pour tout couple de sites $k, l \in \mathbb{Z}^\nu$ et tous champs de vecteurs $X^{(k)} \in \chi(M^{(k)})$ et $X^{(l)} \in \chi(M^{(l)})$,*

$$\sup_{x \in M^{\mathbb{Z}^\nu}} \sum_{\substack{F \in \mathcal{F} \\ \{k,l\} \subseteq F}} \left| \nabla_{(k)} \nabla_{(l)} \mathcal{J}_F(x) \left(X^{(k)}, X^{(l)} \right) \right| \leq \gamma(k,l) \, \|X^{(k)}\| \, \|X^{(l)}\|$$

où $(\gamma(k,l))$ détermine un opérateur borné de $l^2(\mathbb{Z}^\nu)$.

Notons $\|\nabla_{(k)} \nabla_{(l)} H\|_u$ la norme uniforme sur $M^{\{k,l\}}$ de la norme de Hilbert-Schmidt de la forme bilinéaire $\nabla_{(k)} \nabla_{(l)} H_{\{k,l\}}$.

Condition (Pot.2) *Il existe un opérateur borné de $l^2(\mathbb{Z}^\nu)$ de matrice $(\gamma(k,l))$ tel que, pour tout couple de sites $k, l \in \mathbb{Z}^\nu$,*

$$\|\nabla_{(k)} \nabla_{(l)} H\|_u \leq \gamma(k,l)$$

où $(\gamma(k,l))$ détermine un opérateur borné de $l^2(\mathbb{Z}^\nu)$.

Montrons brièvement que la condition $(Pot.2)$ entraîne $(\partial^2 H_\Lambda^\xi l^2)$ uniformément en $\xi \in M^{\mathbb{Z}^\nu}$.

Sous $(Pot.2)$, soient $i, j \in \mathbb{Z}^\nu$, $(X^{(k)})_{k \in \mathbb{Z}^\nu}$ une famille de champs de vecteurs et Λ une partie finie du réseau. Vu que $H_\Lambda^\xi \in \mathcal{C}_\Lambda^\infty$ ne dépend que des spins sur Λ, on peut supposer que $i, j \in \Lambda$.

Alors,

$$\begin{aligned}
\nabla_{(i)} \nabla_{(j)} H_\Lambda^\xi(.) &= \nabla_{(i)} \nabla_{(j)} \sum_{\substack{F \cap \Lambda \neq \emptyset \\ F \in \mathcal{F}}} \mathcal{J}_F(. \, \xi_{\Lambda^c}) \\
&= \sum_{\substack{F \cap \Lambda \neq \emptyset \\ F \in \mathcal{F}}} \nabla_{(i)} \nabla_{(j)} \mathcal{J}_F(. \, \xi_{\Lambda^c}) \\
&= \sum_{\substack{F \supseteq \{i,j\} \\ F \in \mathcal{F}}} \nabla_{(i)} \nabla_{(j)} \mathcal{J}_F(. \, \xi_{\Lambda^c}) \qquad (\text{car } \mathcal{J}_F \in \mathcal{C}_F^\infty)
\end{aligned}$$

Et donc, $\nabla_{(i)} \nabla_{(j)} H_\Lambda^\xi(.) = \nabla_{(i)} \nabla_{(j)} \sum_{\substack{F \supseteq \{i,j\} \\ F \in \mathcal{F}}} \mathcal{J}_F(. \, \xi_{\Lambda^c}) = \nabla_{(i)} \nabla_{(j)} H_{\{i,j\}}^\xi(.)$.

Finalement,

$$\nabla_{(i)} \nabla_{(j)} H_\Lambda^\xi(.) \left(\frac{X^{(i)}}{\|X^{(i)}\|}, \frac{X^{(j)}}{\|X^{(j)}\|} \right) \leq \|\nabla_{(i)} \nabla_{(j)} H\|_u \leq \gamma(i,j)$$

ce qui n'est autre que $(\partial^2 H_\Lambda^\xi l^2)$.

Corollaire 1.3.2. — *Soit un potentiel d'interaction satisfaisant l'une des conditions $(Pot.1)$ ou $(Pot.2)$. Alors il existe $\rho \in \mathbb{R}$ tel que, pour tout $\Lambda \in \mathcal{F}$, toute condition extérieure ξ,*

$$\mathrm{Ric}(\mathbf{L}^{\Lambda,\xi}) \geqslant \rho g_\Lambda.$$

Remarque 1.3.3. — La matrice $(\gamma(k,l))$ décrit la dépendance par rapport au site l de la variation du potentiel au site k. Il est naturel de considérer que c'est une fonction de la distance entre les spins k et l dans $\mathbb{Z}^\nu$. Lorsque $\gamma(i,j) = \gamma(i-j)$ où $\gamma \in l^1(\mathbb{Z}^\nu)$, par l'inégalité de YOUNG pour la convolution, la matrice $(\gamma(i,j))$ définit un opérateur borné de $l^2(\mathbb{Z}^\nu)$.

1.3.2. Isopérimétrie gaussienne pour les systèmes de spins continus. — Dans la précédente section, nous avons abouti à des conditions raisonnables sur le potentiel permettant d'affirmer que pour de nombreux modèles continus la courbure de la dynamique de GLAUBER est bornée inférieurement dans $\mathbb{R}$. Or, cette condition renforce le contrôle de la constante de SOBOLEV logarithmique dans les cas où celui-ci est assuré : d'après le théorème 1.2.4, on a alors le contrôle de la constante d'isopérimétrie. C'est ce que traduit le

Théorème 1.3.4. — *Soit $(\mathcal{J}_\Lambda)_{\Lambda\in\mathbb{Z}^\nu}$ un potentiel d'interaction sur l'espace de configuration $M^{\mathbb{Z}^\nu}$ où M est une variété riemannienne compacte, potentiel satisfaisant à l'une des conditions $(Pot.1)$ ou $(Pot.2)$.*

Supposons de plus qu'on ait le contrôle de la constante de SOBOLEV *logarithmique pour les mesures $\mu_\Lambda^\xi(dx_\Lambda) = (Z_\Lambda^\xi)^{-1} \exp H_\Lambda^\xi(x_\Lambda)\, m^{\otimes\Lambda}(dx_\Lambda)$ à savoir l'existence d'un $B < \infty$ tel que, pour tout $\Lambda \in \mathcal{F}$, toute condition extérieure ξ et toute $f_\Lambda \in \mathcal{C}_\Lambda^\infty$, on ait*

$$\mathbf{Ent}_{\mu_\Lambda^\xi}(f_\Lambda^2) \leq B \int |\nabla f_\Lambda|^2 \, d\mu_\Lambda^\xi$$

Alors, il existe une constante c telle que, pour tout volume fini Λ, toutes conditions extérieures ξ et toute $f_\Lambda \in \mathcal{C}_\Lambda^\infty$ à valeurs dans $[0,1]$,

$$\mathcal{U}\Big(\int f_\Lambda \, d\mu_\Lambda^\xi\Big) \leq \int \sqrt{\mathcal{U}^2(f_\Lambda) + c^2\,|\nabla f_\Lambda|^2}\, d\mu_\Lambda^\xi.$$

On a identifié les mesures μ_Λ^ξ sur M^Λ et $\mu_\Lambda^\xi \otimes \delta_{\xi_{\Lambda^c}}$ sur $M^{\mathbb{Z}^\nu}$.
On rappelle que $\mathbf{Ent}_{\mu_\Lambda^\xi}(f^2) := \int f_\Lambda^2 \log f_\Lambda^2 \, d\mu_\Lambda^\xi - \int f_\Lambda^2 \, d\mu_\Lambda^\xi \log\Big(\int f_\Lambda^2 \, d\mu_\Lambda^\xi\Big)$.

Preuve. ▼ Sous l'une des deux conditions $(Pot.1)$ ou $(Pot.2)$, la famille $(H_\Lambda^\xi)_{\Lambda\in\mathcal{F}}$ des hamiltoniens vérifie $(\partial^2 H_\Lambda^\xi l^2)$ et le semi-groupe associé, qui n'est autre que la dynamique de GLAUBER, satisfait donc à une inégalité $\mathrm{CD}(\rho,\infty)$, $\rho \in \mathbb{R}$ ne dépendant ni de Λ ni de ξ. Le théorème 1.2.4 s'applique donc à chaque $\mathbf{L}_\Lambda^\xi$ avec les mêmes constantes. ▲

Ce résultat étend les travaux de E.LAROCHE (cf. [Lar95]) sur les modèles continus : il était déjà connu que pour les modèles de portée infinie à décroissance exponentielle des variations secondes de l'hamiltonien, **le contrôle du trou dans le spectre, la condition de mélange de DOBRUSHIN-SCHLOSMAN et le contrôle de la constante de SOBOLEV logarithmique sont tous équivalents**. La condition $(Pot.2)$ étant assurée dans ces modèles, on peut affirmer que **le contrôle de la**

constante d'isopérimétrie est encore équivalent à l'une quelconque de ces propriétés.

1.3.3. Déviations du spin moyen par rapport à sa moyenne et à une médiane. —

*Modèle d'*ISING *et convergence exponentielle du spin moyen*

La convergence exponentielle du spin moyen vers sa moyenne dans la région d'unicité de la phase pour le modèle d'ISING sur $\mathbb{Z}^\nu$ est étudiée en détail par R. ELLIS (cf. [Ell85]). Nous donnons ici un bref aperçu des résultats qui y sont présentés. Désignons par $\frac{S_\Lambda}{|\Lambda|} = \frac{1}{|\Lambda|}\sum_{k\in\Lambda} x_k$ le spin moyen sur la boîte $\Lambda \subset \mathbb{Z}^\nu$ ($x_k \in \{-1;1\}$). Dans la région d'unicité de la phase (i.e. pour un champ extérieur $h \neq 0$ ou, pour $h = 0$, quand l'inverse de la température est inférieur à sa valeur critique $0 < \beta < \beta_c$), nous noterons $\mu_{\beta,h}$ la mesure de GIBBS en volume infini. $\mu_{\beta,h}$ est l'unique limite faible quand Λ tend vers $\mathbb{Z}^\nu$ des mesures de GIBBS en volume fini $\mu_{\Lambda,\beta,h,\xi}(dx_\Lambda) = (Z^\xi_{\Lambda,\beta,h})^{-1}\exp\Big(-\beta\, H^\xi_{\Lambda,h}(x_\Lambda)\Big)dx_\Lambda$ pour l'hamiltonien

$$H^\xi_{\Lambda,h}(x_\Lambda) = \sum_{i,j\in\Lambda} J(i-j)x_ix_j + \sum_{\substack{i\in\Lambda\\ j\in\Lambda^c}} J(i-j)x_i\xi_j + h\sum_{i\in\Lambda} x_i.$$

h est le champ extérieur agissant sur les spins et ξ est la condition extérieure i.e. les spins hors de Λ interagissant avec la configuration x_Λ. Le terme hors interaction de l'hamiltonien est le produit de h par la magnétisation totale sur Λ. Cette forme très particulière de l'hamiltonien lie la transition de phase sur les mesures à un autre phénomène critique, la magnétisation spontanée i.e. l'existence d'une discontinuité en $h = 0$ de la magnétisation spécifique

$$m(\beta,h) = \lim_{\Lambda\to\mathbb{Z}^\nu}\int \frac{S_\Lambda}{|\Lambda|}\,\mu_{\Lambda,\beta,h,\xi}(dx_\Lambda),$$

(indépendante de la configuration extérieure ξ). Comme la mesure de GIBBS en volume infini est invariante par translation et ergodique, il vient

$$m(\beta,h) = \int_{\{-1,1\}^{\mathbb{Z}^\nu}} x_0\,\mu_{\beta,h}(dx).$$

Pour $h \neq 0$ ou $h = 0$ et $\beta < \beta_c$, cette magnétisation n'est autre que

$$m(\beta,h) = -\frac{\partial\Psi}{\partial h}(\beta,h)$$

où $\Psi(\beta,h) = \lim_{\Lambda\to\mathbb{Z}^\nu}\frac{1}{|\Lambda|}\,\Psi(\Lambda,\beta,h,\xi)$ est l'énergie libre spécifique, l'énergie libre sur Λ étant donnée par

$$\Psi(\Lambda,\beta,h,\xi) = -\,\beta^{-1}\,\log Z^\xi_{\Lambda,\beta,h}.$$

Pour $\beta > \beta_c$, la différentiabilité de Ψ en $h = 0$ est rompue et

$$-\frac{\partial \Psi}{\partial h}(\beta, 0_+) = m(\beta, 0_+) > m(\beta, 0) = 0 > -\frac{\partial \Psi}{\partial h}(\beta, 0_-) = m(\beta, 0_-).$$

L'énergie libre est encore la bonne quantité thermodynamique à considérer lorsqu'on s'intéresse aux propriétés de grandes déviations de la loi du spin moyen $\frac{S_\Lambda}{|\Lambda|}$ sous $\mu_{\beta,h}$. L'énergie libre des lois de $\frac{S_\Lambda}{|\Lambda|}$, i.e. la limite des log-LAPLACE spécifiques de S_Λ, à savoir $c_{\beta,h}(t) := \lim_{\Lambda \to \mathbb{Z}^\nu} \frac{1}{|\Lambda|} \log \left(\int \exp\left(t\, S_\Lambda\right) d\mu_{\beta,h} \right)$ n'est autre que

$$c_{\beta,h}(t) = \beta \left(\Psi(\beta, h + t/\beta) - \Psi(\beta, h) \right)$$

qui est différentiable en 0 dans la région d'unicité de la phase. Cela conduit à un principe de grandes déviations de fonction de taux $I_{\beta,h}(z) = \sup_{t \in \mathbb{R}} \left(tz - c_{\beta,h}(t) \right)$ et, par conséquent, à la convergence exponentielle du spin moyen sous $\mu_{\beta,h}$ vers $m(\beta, h)$. Cette convergence signifie que, pour tout $\varepsilon > 0$, il existe $A = A(\varepsilon) > 0$ telle que, pour toute boîte Λ de cardinal assez grand, on ait

$$\mu_{\beta,h} \left(|S_\Lambda / |\Lambda| - m(\beta, h)| > \varepsilon \right) \leqslant e^{-A\,|\Lambda|}. \tag{40}$$

La détermination de la meilleure constante A_∞ pour laquelle (40) est vérifiée pour toute $A < A_\infty$ est assurée par le principe de grandes déviations. La connaissance de cette vitesse optimale de convergence exponentielle est théoriquement donnée par celle de $\Psi(\beta, h)$:

$$A_\infty = \inf_{\{|z - m(\beta,h)| \geqslant \varepsilon\}} I_{\beta,h}(z).$$

Concentration de la mesure

Les inégalités de concentration de la mesure dérivées des inégalités de SOBOLEV logarithmique d'une part et de l'isopérimétrie gaussienne d'autre part donnent des résultats similaires de déviations par rapport à la moyenne et à la médiane. Les vitesses de convergence exponentielle obtenues, bien que non optimales en général, sont maintenant exprimées explicitement à l'aide des constantes de log-SOBOLEV et d'isopérimétrie. La convergence exponentielle y est de plus valable pour tout $|\Lambda|$ ce qui supprime l'inconvénient du caractère asymptotique de (40). Mais ces résultats s'expriment en volume fini la moyenne considérée dépendant de la boîte considérée, ne comportant donc pas en eux-mêmes de propriétés *ergodiques* de la loi du spin moyen.

L'argument par lequel une inégalité de concentration se déduit de l'inégalité de log-SOBOLEV est dû à HERBST. Il consiste en fait à obtenir une majoration quadratique du logarithme de la transformée de LAPLACE. Ce qui conduit à

$$\mu \left(\left| f - \int f\, d\mu \right| \geqslant r \right) \leqslant 2 \exp\left(-r^2/(C\,\kappa) \right). \tag{41}$$

Remarquons que ceci ne prend de réelle signification que pour r grand. Voir [Led97, Led99] pour plus de détails.

Supposons maintenant satisfaite l'inégalité d'isopérimétrie gaussienne fonctionnelle (25). Rappelons qu'alors ([BL96a], thm 3.2.) la loi ν d'une $\mathbf{\Gamma}$-contraction f est l'image par une contraction k de la gaussienne γ (i.e. $\nu = k[\gamma]$). On peut en déduire une inégalité de concentration par des arguments fonctionnels de la manière suivante.

Un cas simple est celui où la loi ν de f est à densité strictement positive. Dans ce cas, la fonction de répartition F_ν de ν est inversible et, par un procédé classique, la loi image $F_\nu[\nu]$ est la mesure de Lebesgue sur $[0,1]$. Comme on a de même $\varphi[\gamma] = \lambda_{|[0,1]}$, on voit facilement que $\nu = k[\gamma]$ où $k = F_\nu^{-1} \circ \varphi$.

L'inégalité isopérimétrique gaussienne fonctionnelle implique alors que, si $\mathbf{\Gamma}(f) \leqslant \kappa$, $\|k\|_{Lip} \leqslant c\sqrt{\kappa}$ où c est la constante d'isopérimétrie (cf. [BL96a]). Remarquons que l'unique médiane de f n'est autre que $m = k(0)$. En notant $u_r^+ = k^{-1}(m+r)$ et $u_r^- = k^{-1}(m-r)$, on obtient

$$\mu(|f-m| \geqslant r) = \nu(|x-m| \geqslant r) = \gamma(]-\infty, u_r^-]) + \gamma([u_r^+, +\infty[).$$

Et, comme $\|k\|_{Lip} \leqslant c\sqrt{\kappa}$ et $k(0) = m$, $u_r^+ \geqslant \frac{r}{c\sqrt{\kappa}}$ et $u_r^- \leqslant -\frac{r}{c\sqrt{\kappa}}$. Par suite,

$$\mu(|f-m| \geqslant r) \leqslant 2\,(1 - \varphi(\frac{r}{c\sqrt{\kappa}})) \leqslant \frac{2c\sqrt{\kappa}}{r}\, e^{-\frac{r^2}{2c^2\kappa}},$$

puisque $\gamma([r,+\infty[) \leqslant \frac{1}{r}\int_r^\infty x\,\gamma(dx) \leqslant \frac{1}{r}e^{-\frac{r^2}{2}}$. Finalement,

$$\mu(|f-m| \geqslant r) \leqslant \frac{2c\sqrt{\kappa}}{r}\, e^{-\frac{r^2}{2c^2\kappa}}. \tag{42}$$

Nous ne supposons plus maintenant la loi de f absolument continue. L'inégalité isopérimétrique gaussienne fonctionnelle étant stable par tensorisation, on l'obtient sur $(E \times \mathbb{R}, \mu \otimes \gamma)$ pour la constante $\tilde{c} = \max(c, 1)$. On applique ce qui vient d'être fait pour les lois ν_ε de $f_\varepsilon(x,u) = (1-\varepsilon^2)^{\frac{1}{2}} f(x) + \kappa\,\varepsilon u$. $(\nu_\varepsilon)_\varepsilon$ converge faiblement vers ν. Pour tout $\varepsilon > 0$, on obtient une $c\sqrt{\kappa}$-contraction k_ε telle que $\nu_\varepsilon = k_\varepsilon[\gamma]$. Vu que $k'_\varepsilon \leqslant \kappa$ p.s., k'_ε est dans la boule de $\mathbf{L}^2(dx)$ de rayon κ qui est faiblement compacte. Il existe donc une sous-suite $k'_{1/\tilde{n}}$ convergeant faiblement vers g telle que $\|g\|_2 \leqslant \kappa$. De plus, les médianes m_ε de ν_ε étant toutes dans un même compact de $\mathbb{R}$, on peut également en extraire une sous-suite convergente, disons vers m. m est alors une médiane de ν. Posant

$$k(x) = m + \int_{[0,x]} g(u)\,du,$$

on voit que k est une $c\sqrt{\kappa}$-contraction, que $\nu = k[\gamma]$ et que $k(0) = m$. C'est tout ce dont nous avons besoin pour obtenir (42).

En outre, on peut déduire de la concentration par rapport à une médiane (42) des constantes explicites pour la concentration par rapport à la moyenne.

Application au spin moyen

On considère maintenant un système de spins continus comme introduit page 1.3.1. Soit f une fonction $\mathcal{C}^\infty$ sur M de carré du champ borné, i.e. telle que $|\nabla f|$ soit borné, disons par 1. On considère le spin moyen associé $\frac{1}{|\Lambda|}S_\Lambda(x_\Lambda)$, où $S_\Lambda(x_\Lambda) = \sum_{k\in\Lambda} f(x_k)$. Un calcul facile montre que

$$\mathbf{\Gamma}_\Lambda\left(\frac{1}{|\Lambda|}\sum_{k\in\Lambda} f(x_k)\right) = \frac{1}{|\Lambda|^2}\sum_{k\in\Lambda}\mathbf{\Gamma}(f)(x_k) \leqslant \frac{1}{|\Lambda|}.$$

L'équation (42) montre alors que, si c désigne la constante d'isopérimétrie gaussienne,

$$\mu_\Lambda^\xi\left(\left|\frac{S_\Lambda}{|\Lambda|} - m_{\Lambda,\xi}\right| \geqslant r\right) \leqslant \frac{2c}{r\sqrt{|\Lambda|}}\, e^{-\frac{r^2|\Lambda|}{2c^2}}$$

pour une médiane $m_{\Lambda,\xi}$ de $\frac{S_\Lambda}{|\Lambda|}$. Et d'autre part, suivant (41), si C désigne la constante de log-SOBOLEV,

$$\mu_\Lambda^\xi\left(\left|\frac{S_\Lambda}{|\Lambda|} - \overline{m}_{\Lambda,\xi}\right| \geqslant r\right) \leqslant 2\,\exp\left(-r^2\,|\Lambda|/C\right)$$

où $\overline{m}_{\Lambda,\xi}$ est la moyenne de $\frac{S_\Lambda}{|\Lambda|}$ sous μ_Λ^ξ.

1.4. Spins discrets et isopérimétrie

1.4.1. Cadre général et définitions. —

Les générateurs

Soit F un ensemble fini de cardinal $|F| = q$. Un générateur de MARKOV $\mathbf{L}$ sur F est totalement déterminé par une matrice $(L_{ab})_{a,b\in F}$: $\mathbf{L}f(a) = \sum_{b\in F} L_{ab}\, f(b)$ où, pour tout $b \neq a$, $L_{ab} \geqslant 0$ et $\sum_{b\in F} L_{ab} = 0$. On dira dans la suite que deux points a et b sont *voisins* si $L_{ab} > 0$. Tout générateur peut être décrit de manière équivalente à l'aide du gradient discret de la façon suivante.

Toute famille $\alpha = (\alpha^b(.))_{b\in F}$ de fonctions positives sur F détermine un générateur

$$\mathbf{L}^\alpha f(a) = \langle \alpha, \nabla f\rangle_{\mathbb{R}^q}(a) = \sum_{b\in F} \alpha^b(a)\,\nabla^b f\,(a)$$

où le gradient discret est donné par le vecteur de $\mathbb{R}^q$

$$\nabla f(a) = (\nabla^b f\,(a))_{b\in F} \text{ avec } \nabla^b f\,(a) = f(b) - f(a)\,.$$

On a ainsi une correspondance univoque entre les classes de familles α, modulo coïncidence des $\alpha^a(.)$ sauf éventuellement en a, et les générateurs sur F. Il suffit de fixer

$$\alpha^b(a) = \left|\begin{array}{cc} L_{ab} & \text{si } b \neq a \\ \text{n'importe quoi} & \text{sinon} \end{array}\right.$$

Le générateur $\mathbf{L}^\alpha$ est donné par un noyau $K(.,.)$, $\mathbf{L}f(x) = \int f(y)\, K(x, dy) - f(x)$, si, pour tout $a \in F$, $\sum_{b \neq a} \alpha^b(a) \leq 1$. Dans ce cas, $K(a,b) = \alpha^b(a)$ si $b \neq a$.

Remarque 1.4.1. — La condition générale pour que $\mathbf{L}$ soit réversible pour une mesure μ chargeant tous les points (i.e $d\mu = \exp H\, dm$ où m est la mesure uniforme) est la relation de commutation

$$\forall a, b \in F, \qquad \alpha^b(a) = \exp \nabla^b H(a)\, \alpha^a(b).$$

Dans le cas où F est à deux éléments, $\langle \alpha, \nabla . \rangle = \beta \,.\, \overline{\nabla}$ avec $\overline{\nabla} f(.) = \nabla^{\tau(.)} f(.) = f \circ \tau - f$ et $\beta = \alpha^{\tau(.)}(.)$, τ échangeant les deux éléments. La condition se résume alors à

$$\beta = \beta \circ \tau \ \exp \overline{\nabla} H$$

ce qui, compte tenu que, ici, $\beta + \beta \circ \tau$ est constant, n'est autre que $\beta = c \exp \frac{\overline{\nabla} H}{2}$ (cf. [BM92]).

Construction sur un produit fini

Les processus sur les produits finis F^Λ, $\Lambda \in \mathcal{F}$, sur lesquels nous porterons notre attention sont construits en considérant deux configurations $\boldsymbol{i}$ et $\boldsymbol{j}$ sur F^Λ comme voisines si elles ne diffèrent qu'en un site. Le processus dans l'état $\boldsymbol{i}$ ne peut atteindre l'état $\boldsymbol{j}$ que si $\boldsymbol{i}$ et $\boldsymbol{j}$ sont voisins, le noyau de transition étant alors donné par un noyau sur F au site où les configurations diffèrent. Explicitement, ce semi-groupe admet pour générateur $\mathbf{L}_\Lambda f = \sum_{k \in \Lambda} \overline{\mathbf{L}}_k f$, avec $\overline{\mathbf{L}}_k f(x) = \mathbf{L}_k[f(x._k(.))](x_k)$, où $\mathbf{L}_k$ est un générateur sur F et où la configuration $x._k b$ est égale à x sauf au site k où elle prend la valeur b .
$\mathbf{L}_\Lambda$ ainsi obtenu sera dit *produit* des L_k.

Pour que ce semi-groupe soit réversible pour $m^{\otimes \Lambda}$, il suffit que, sur la fibre, les $\mathbf{L}_k$ soient réversibles pour m. Si chaque $\mathbf{L}_k$ s'obtient à partir du même noyau $K(.,.)$, cela revient à la condition $K(x,y)\, m(x) = K(y,x)\, m(y)$ ou encore, si m est la mesure uniforme, à la symétrie de K.

Etablissons à ce stade des notations utiles pour l'étude des générateurs précédents. Etant donnée une configuration x, nous noterons tout d'abord $\tau_k^b(x)$ la configuration $x._k b$. Si, sur chaque $k \in \Lambda$, est donné un générateur $\mathbf{L}_k$ associé à la famille α_k de fonctions sur F, le produit $\mathbf{L}$ sur F^Λ des $\mathbf{L}_k$ s'écrit

$$\mathbf{L}f = \sum_{(k,b) \in \Lambda \times F} \alpha_k^b \nabla_k^b f = \sum_{k \in \Lambda} \langle \alpha_k, \nabla_k f \rangle_k$$

où $\nabla_k^b f := f \circ \tau_k^b - f$ et $\langle \alpha_k, \nabla_k f \rangle_k(x) := \sum_{b \in F} \alpha_k^b(x_k) \nabla_k^b f(x)$.

En fait, les processus auxquels nous aurons affaire sont un peu plus généraux. Nous les appellerons des *quasi-produits*. Ils sont construits de la même manière

mais avec un semi-groupe sur F au site k pouvant dépendre de la configuration hors de k i.e $\mathbf{L}_k f(x) = \mathbf{L}_k^{x_{\Lambda-\{k\}}}[f(x_{\cdot k}(.))](x_k)$. $\mathbf{L}$ s'écrit comme précédemment $\mathbf{L}f = \sum_{k\in\Lambda}\langle \alpha_k, \nabla_k f\rangle_k$, α_k étant maintenant une famille $(\alpha_k^b(.))_{b\in F}$ de fonctions sur F^Λ.

1.4.2. Contrôle de la courbure pour les systèmes discrets. — Rappelons deux propriétés du gradient discret qui le distinguent du gradient usuel. Tout d'abord, la formule de Leibnitz discrète,

$$\nabla_k(fg) = \nabla_k f\,.\,g + \nabla_k g\,.\,f + \nabla_k f\,.\,\nabla_k g \tag{43}$$

$$= \nabla_k f\,.\,g + f\circ\tau_k\,.\,\nabla_k g \tag{44}$$

Enfin, la dérivation de la fonction exponentielle,

$$e^{-H}\,\nabla_k(e^H) = e^{(\nabla_k H)} - 1\ . \tag{45}$$

Ces égalités sont bien sûr entendues dans $\mathbb{R}^F$.

Calcul de l'opérateur $\mathbf{\Gamma_2}$ du générateur quasi-produit $\mathbf{L}f = \sum_{k\in\Lambda}\langle\alpha_k, \nabla_k f\rangle$

De par la forme de $\mathbf{L}$ ($\mathbf{L} = \sum_{k\in\Lambda}\langle\alpha_k, \nabla_k .\rangle$) et la définition de $\mathbf{\Gamma}$, la formule de Leibnitz pour le gradient discret assure que

$$2\,\mathbf{\Gamma}(f,g) = \sum_{k\in\Lambda}\langle\alpha_k, \nabla_k f\,.\,\nabla_k g\rangle_k$$

où $(\nabla_k f\,.\,\nabla_k g)^b := \nabla_k^b f\,.\,\nabla_k^b g$.
Rappelons qu'on note, pour deux vecteurs $X_k = (X_k^b)_{b\in F}$ et Y_k de $\mathbb{R}^q$ au site k, $\langle X_k, Y_k\rangle_k = \sum_{b\in F} X_k^b Y_k^b$. Ce produit scalaire jouera dans les calculs, en parallèle avec le cas continu, le rôle de la métrique riemannienne au site k.

Alors,

$$\begin{aligned} 4\,\mathbf{\Gamma_2}(f,f) &= \mathbf{L}(2\,\mathbf{\Gamma}(f,f)) - 2\,(2\,\mathbf{\Gamma}(f,\mathbf{L}f)) \\ &= \sum_{k,l\in\Lambda}\langle\alpha_k, \nabla_k\,\langle\alpha_l, (\nabla_l f)^2\rangle_l\rangle_k - \ 2\sum_{k,l\in\Lambda}\langle\alpha_l, \nabla_l f\,.\,\nabla_l\langle\alpha_k, \nabla_k f\rangle_k\rangle_l \end{aligned}$$

Appliquant la formule de Leibnitz, on obtient, par linéarité,

$$\nabla_k\left(\langle\alpha_l, \beta_l\rangle_l\right) = \langle\nabla_k\alpha_l, \beta_l\rangle_l + \langle\alpha_l\circ\tau_k, \nabla_k\beta_l\rangle_l$$

et, en outre,

$$\nabla_k\left[(\nabla_l f)^2\right] = \ 2\,\nabla_l f\,.\,\nabla_k\nabla_l f + (\nabla_k\nabla_l f)^2$$

D'où,

$$\begin{aligned}4\,\mathbf{\Gamma_2}(f,f) = \sum_{k\in\Lambda}\langle \mathbf{L}\alpha_k, (\nabla_k f)^2\rangle_k + \sum_{k,l\in\Lambda}\langle \alpha_k, \langle \alpha_l\circ\tau_k, (\nabla_k\nabla_l f)^2\rangle_l\rangle_k \\ +\ 2\sum_{k,l\in\Lambda}\langle \alpha_k, \langle \alpha_l\circ\tau_k, \nabla_l f\,.\,\nabla_k\nabla_l f\rangle_l\rangle_k \\ -\ 2\sum_{k,l\in\Lambda}\langle \alpha_l, \nabla_l f\,.\,\langle \alpha_k\circ\tau_l, \nabla_l\nabla_k f\rangle_k\rangle_l \\ -\ 2\sum_{k,l\in\Lambda}\langle \alpha_l, \nabla_l f\,.\,\langle \nabla_l\alpha_k, \nabla_k f\rangle_k\rangle_l\end{aligned}$$

Remarque 1.4.2. — La relation de commutation du gradient $\nabla_k^b\,\nabla_l^c f = \nabla_l^c\,\nabla_k^b f$ est fausse en toute généralité dans le cas discret. En effet, $\nabla_k^b\,\nabla_k^c f = -\nabla_k^b f$ diffère en général de $\nabla_l^c\,\nabla_k^b f$ si $b \neq c$.

Contrôle de la courbure

Afin de majorer, à une constante près indépendante de Λ, la forme quadratique $\mathbf{\Gamma_2^\Lambda}$ par $\mathbf{\Gamma}^\Lambda$, nous allons procéder en plusieurs étapes. Tout d'abord, comparer les deux formes $\sum_{k\in\Lambda}\langle \mathbf{L}\alpha_k, (\nabla_k f)^2\rangle_k$ et $\sum_{k,l\in\Lambda}\langle \alpha_l, \nabla_l f\,.\,\langle \nabla_l\alpha_k, \nabla_k f\rangle_k\rangle_l$ avec $\mathbf{\Gamma}^\Lambda(f,f)$. Ceci ne présente pas de difficultés particulières et impose simplement de pouvoir contrôler α_k uniformément et $\nabla_l\alpha_k$ par une fonction l^1 de la distance de k à l. L'étape suivante consiste à regrouper les autres membres de $\mathbf{\Gamma_2^\Lambda}$ à l'aide de la méthode de Gauss de réduction en somme de carrés. Cela nous permettra de conclure sous les mêmes conditions de contrôle.

En premier lieu, nous supposerons les α_k^Λ bornés uniformément indépendamment de Λ :

$$((\alpha^\Lambda)Unif.)\qquad \begin{array}{l}\textit{Il existe } R>0 \textit{ tel que, pour toute } \Lambda\in\mathcal{F},\\ \textit{tout } b\in F,\ \textit{tout } k\in\Lambda \textit{ et tout } x\in F^\Lambda,\\ \qquad\qquad |(\alpha^\Lambda)_k^b(x)| \leqslant R\end{array}$$

On se ramène ainsi à comparer les formes quadratiques avec la forme

$$\sum_{k\in\Lambda}\langle 1, (\nabla_k f)^2\rangle_k$$

Pour ce faire, nous supposerons qu'il existe une fonction $\gamma \in l^1(\mathbb{Z}^\nu)$ telle que la condition suivante soit satisfaite

$(\partial^1(\alpha^\Lambda)l^1)$ *Pour toute $\Lambda \in \mathcal{F}$, tous $b, c \in F$, tout $x \in F^\Lambda$, et tous $l, k \in \mathbb{Z}^\nu$,*

$$\left|\nabla_l^c \left[\left(\alpha^\Lambda\right)_k^b\right](x)\right| \leqslant \gamma(l-k)$$

Nous aborderons plus loin la version potentiel de cette condition.

Proposition 1.4.2.1. — *Pour toute famille de fonctions $(\alpha^\Lambda)_{\Lambda\in\mathcal{F}}$ satisfaisant aux conditions $((\alpha^\Lambda)Unif.)$ et $(\partial^1(\alpha^\Lambda)l^1)$, les courbures des générateurs induits $\mathbf{L}^{\alpha^\Lambda}$ sont minorées dans $\mathbb{R}$ indépendamment de Λ i.e.*

$$\exists\rho \in \mathbb{R},\ \forall\Lambda \in \mathcal{F},\ \mathbf{\Gamma_2}^{\alpha_\Lambda} \ \geqslant\ \rho\,\mathbf{\Gamma}^{\alpha_\Lambda}.$$

Preuve

◀ En ce qui concerne $\sum_{k\in\Lambda}\langle \mathbf{L}\alpha_k, (\nabla_k f)^2\rangle_k$, il suffit de vérifier que $\mathbf{L}\alpha_k$ est borné uniformément. On obtient, en omettant les indices Λ pour les α_k,

$$\begin{aligned}|\mathbf{L}\alpha_k| = |\sum_{l\in\Lambda}\langle\alpha_l, \nabla_l\alpha_k\rangle_l| \ &\leqslant\ \sum_{l\in\Lambda}\langle|\alpha_l|, |\nabla_l\alpha_k|\rangle_l \\ &\leqslant R\sum_{l\in\Lambda}\gamma(l-k)\,\langle 1, 1\rangle_l \ \leqslant\ R\,|F|\,\|\gamma\|_{l^1(\mathbb{Z}^\nu)}\end{aligned}$$

Étudions maintenant le cas de la seconde forme quadratique. Le problème revient à trouver une constante C telle que, pour toute Λ et toute famille $(Y_k)_{k\in\Lambda}$ de vecteurs de $\mathbb{R}^q$,

$$\left|\sum_{k,l\in\Lambda}\langle\alpha_l, Y_l\,.\,\langle\nabla_l\alpha_k, Y_k\rangle_k\rangle_l\right| \leqslant C\sum_{k\in\Lambda}\langle 1, (Y_k)^2\rangle_k$$

Or,

$$\begin{aligned}\left|\sum_{k,l\in\Lambda}\langle\alpha_l, Y_l\,.\,\langle\nabla_l\alpha_k, Y_k\rangle_k\rangle_l\right| &\leqslant \sum_{k,l\in\Lambda}\langle|\alpha_l|, |Y_l|\,.\,\langle|\nabla_l\alpha_k|, |Y_k|\rangle_k\rangle_l \\ &\leqslant R\sum_{k,l\in\Lambda}\gamma(k-l)\,\langle 1, |Y_l|\,.\,\langle 1, |Y_k|\rangle_k\rangle_l \\ &\leqslant R\sum_{k,l\in\Lambda}\gamma(k-l)\,\langle 1, |Y_l|\rangle_l\,.\,\langle 1, |Y_k|\rangle_k \\ &\leqslant R\,\langle\gamma * \langle 1, |Y_.|\rangle, \langle 1, |Y_.|\rangle\rangle_{l^2(\mathbb{Z}^\nu)} \leqslant R\,\|\gamma\|_1\,\|\langle 1, |Y_.|\rangle\|^2_{l^2(\Lambda)}\end{aligned}$$

Il suffit alors de remarquer que $\|\langle 1, |Y_.|\rangle\|^2_{l^2(\Lambda)} = \sum_{k\in\Lambda}\langle 1, |Y_k|\rangle_k^2$ et que, par CAUCHY-SCHWARZ, $\langle 1, |Y_k|\rangle_k^2 \leqslant |F|\langle 1, |Y_k|^2\rangle_k$.

Penchons nous sur les derniers termes à contrôler à savoir

$$\sum_{k,l\in\Lambda} \langle \alpha_k, \langle \alpha_l \circ \tau_k, (\nabla_k \nabla_l f)^2 \rangle_l \rangle_k$$
$$+\ 2 \sum_{k,l\in\Lambda} \langle \alpha_k, \langle \alpha_l \circ \tau_k, \nabla_l f \,.\, \nabla_k \nabla_l f \rangle_l \rangle_k$$
$$-\ 2 \sum_{k,l\in\Lambda} \langle \alpha_l, \nabla_l f \,.\, \langle \alpha_k \circ \tau_l, \nabla_l \nabla_k f \rangle_k \rangle_l$$

Pour des raisons techniques, nous isolons les termes diagonaux correspondants. On a vu que $\nabla_k^b \nabla_k^c f = -\nabla_k^b f$. En conséquence, ces termes se réduisent à

$$\sum_{k=\overline{k}} \langle \alpha_k, \langle \alpha_{\overline{k}} \circ \tau_k, (\nabla_k f)^2 \rangle_{\overline{k}} \rangle_k -\ 2 \sum_{k=\overline{k}} \langle \alpha_k, \langle \alpha_{\overline{k}} \circ \tau_k, \nabla_{\overline{k}} f \,.\, \nabla_k f \rangle_{\overline{k}} \rangle_k$$
$$+\ 2 \sum_{k=\overline{k}} \langle \alpha_{\overline{k}}, \nabla_{\overline{k}} f \,.\, \langle \alpha_k \circ \tau_{\overline{k}}, \nabla_{\overline{k}} f \rangle_k \rangle_{\overline{k}}$$
$$=\ 3 \sum_{k=\overline{k}} \langle \alpha_k, (\nabla_k f)^2 \, \langle \alpha_{\overline{k}} \circ \tau_k, 1 \rangle_{\overline{k}} \rangle_k -\ 2 \sum_{k=\overline{k}} \langle \alpha_k, \langle \alpha_{\overline{k}} \circ \tau_k, \nabla_{\overline{k}} f \rangle_{\overline{k}} \,.\, \nabla_k f \rangle_k$$

Le module du premier terme se majore aisément par $R^2 \, |F| \, \sum_{k\in\Lambda} \langle 1, (\nabla_k f)^2 \rangle_k$ grâce à $((\alpha^\Lambda) Unif.)$. Quant au second, il suffit de voir que

$$|\langle \alpha_k, \langle \alpha_{\overline{k}} \circ \tau_k, \nabla_{\overline{k}} f \rangle_{\overline{k}} \,.\, \nabla_k f \rangle_k| \leqslant \langle |\alpha_k|, \langle |\alpha_{\overline{k}} \circ \tau_k|, |\nabla_{\overline{k}} f| \rangle_{\overline{k}} \,.\, |\nabla_k f| \rangle_k$$
$$\leqslant R^2 \, \langle 1, |\nabla_{\overline{k}} f| \rangle_{\overline{k}} \, \langle 1, |\nabla_k f| \rangle_k \leqslant R^2 \, \langle 1, |\nabla_k f| \rangle_k^2 \leqslant R^2 \, |F| \, \langle 1, |\nabla_k f|^2 \rangle_k$$

Pour traiter les termes non diagonaux, on peut omettre les fonctions f et se ramener à des vecteurs X_{kl} symétriques et Y_l. Fixant x, on peut en effet déterminer une f telle que, pour tout $k \in \Lambda$, $\nabla_k f(x) = Y_k$ et, pour tout $k \neq l$, $\nabla_k \nabla_l f(x) = \nabla_l \nabla_k f(x) = X_{kl}$ i.e. une f dont les valeurs sont fixées sur les configurations différant de x en un ou deux sites. On supposera par commodité les X_{kl} définis pour chaque couple (k, l), imposant qu'ils s'annulent sur la diagonale.
Il vient,

$$\sum_{k,l\in\Lambda} \langle \alpha_k, \langle \alpha_l \circ \tau_k, \, 1/2 \, (X_{kl})^2 \rangle_l \rangle_k +\ 2 \sum_{k,l\in\Lambda} \langle \alpha_k, \langle \alpha_l \circ \tau_k, Y_l \,.\, X_{kl} \rangle_l \rangle_k$$
$$= \sum_{k,l\in\Lambda} \langle \alpha_k, \langle \alpha_l \circ \tau_k, \, 1/2 \, (X_{kl})^2 +\ 2 \, Y_l \,.\, X_{kl} \rangle_l \rangle_k$$
$$= \sum_{k,l\in\Lambda} \langle \alpha_k, \langle \alpha_l \circ \tau_k, (1/\sqrt{2} \, X_{kl} + \sqrt{2} \, Y_l)^2 -\ 2 \, Y_l^2 \rangle_l \rangle_k$$
$$\geqslant -2 \sum_{k,l\in\Lambda} \langle \alpha_k, \langle \alpha_l \circ \tau_k, Y_l^2 \rangle_l \rangle_k$$

Il s'agit donc de majorer le module de cette dernière forme par

$$C \sum_{k\in\Lambda} \langle 1, (\nabla_k f)^2 \rangle_k$$

ce qui est évident compte tenu de l'hypothèse $((\alpha^\Lambda)Unif.)$
Enfin,

$$\sum_{k,l\in\Lambda} \langle \alpha_k, \langle \alpha_l \circ \tau_k,\, 1/2\ (X_{kl})^2 \rangle_l \rangle_k = \sum_{k,l\in\Lambda} \langle \alpha_l, \langle \alpha_k \circ \tau_l,\, 1/2\ (X_{kl})^2 \rangle_k \rangle_l$$

par symétrie de X_{kl}. Et le terme restant se traite de façon similaire. ►

1.4.3. Dynamiques discrètes en mécanique statistique. — Nous avons vu précédemment des conditions suffisantes pour que la courbure des quasi-produits $\mathbf{L}^\Lambda = \sum_{k\in\Lambda} \langle \alpha_k, \nabla_k f \rangle_k$ soit minorée indépendamment de Λ. Nous considérons maintenant des quasi-produits particuliers utilisés en mécanique statistique et construits à l'aide du potentiel pour qu'ils soient réversibles par rapport à la mesure de Gibbs en volume fini. Nous donnons ensuite des conditions sur le potentiel d'interactions pour que ces quasi-produits aient une courbure minorée (dans $\mathbb{R}$).

La dynamique de Glauber

Il s'agit de la dynamique $\mathbf{L}f = \sum_{k\in\Lambda} \langle \alpha_k, \nabla_k f \rangle_k$ où

$$\alpha_k^b(x) = \mu_{\{k\}}^{x_\Lambda.\xi_{\Lambda^c}}(b) = \frac{\exp H_{\{k\}}^{x_\Lambda.\xi_{\Lambda^c}}(b)}{\sum_{c\in F} \exp H_{\{k\}}^{x_\Lambda.\xi_{\Lambda^c}}(c)}.$$

Posant $H(x_\Lambda) = H_{\{k\}}^{x_\Lambda.\xi_{\Lambda^c}}(x_k)$, il vient

$$\begin{aligned}
\alpha_k^b(x) &= \frac{\exp H \circ \tau_k^b(x)}{\exp H(x) \left(1 + \sum_{c\neq b} \exp\left(H \circ \tau_k^c(x) - H(x)\right)\right)} \\
&= \frac{\exp H \circ \tau_k^b(x)}{\exp H(x) \left(1 + \sum_{c\neq b} \exp \nabla_k^c H(x)\right)} \\
&= \frac{\exp \nabla_k^b H(x)}{1 + \sum_{c\neq b} \exp \nabla_k^c H(x)}
\end{aligned}$$

Finalement,

$$\alpha_k^b(x) = \frac{\exp \nabla_k^b H_\Lambda^\xi(x)}{1 + \sum_{c\neq b} \exp \nabla_k^c H_\Lambda^\xi(x)} \tag{46}$$

puisque, si $k \in \Lambda$, $\nabla_k^b H_{\{k\}}(x_\Lambda.\xi_{\Lambda^c}) = \nabla_k^b H_\Lambda^\xi$. Cette dynamique est très utilisée. On pourra par exemple consulter [Lar95] pour l'étude du contrôle de la constante de log-Sobolev associée.

Dynamique produit induite par un noyau

Un processus simple sur un ensemble fini est celui qui, d'un point, saute indifféremment sur un point distinct au bout d'un certain temps exponentiel. Le processus de générateur $\mathbf{L}f = \int f\,dm - f$, où m est la mesure uniforme sur la fibre, en fournit un exemple. Lorsque l'on s'intéresse à une étude plus fine de certains systèmes, on est amené à manipuler des semi-groupes qui englobent certaines propriétés de symétrie du modèle. Pour le rotateur plan $\mathbb{Z}/n\mathbb{Z}$, identifié aux racines n-ième de l'unité sur le cercle S^1, le semi-groupe n'est pas donné par la mesure uniforme mais par un noyau $K(x,y)$: $\mathbf{L}f = \int f(y)\,K(.,dy) - f$. Le processus saute de x à $y \neq x$ au bout d'un temps exponentiel de paramètre $\lambda_x = \sum_{y\neq x} K(x,y)$ avec la probabilité $K(x,y)/\sum_{y\neq x} K(x,y)$. Cette probabilité peut être proportionnelle au cosinus de l'angle entre x et y. Le choix d'un noyau pour lequel, pour chaque x, la mesure $K(x,dy)$ est portée par $F - \{x\}$ présente l'avantage de normaliser les paramètres exponentiels à l'unité et d'exprimer les probabilités de transitions directement par le noyau.

Sur un ensemble fini, contrairement au cas des variétés compactes où toute diffusion réversible par rapport à $d\mu = \exp\Phi\ dm$ s'écrit canoniquement $\Delta + \nabla\Phi$ dans une métrique riemannienne associée, une certaine liberté est autorisée dans le choix du générateur. Ce qui est donné pour l'étude du système est principalement un carré du champ $\mathbf{\Gamma}$, généralement réversible pour la mesure uniforme m sur la fibre finie F, et qui exprime l'étalement des valeurs d'une fonction autour de l'une d'entre elles. Cet opérateur induit un carré du champ $\mathbf{\Gamma}_\Lambda$ sur F^Λ. Les quantités effectivement significatives sont ses moyennes par rapport à une mesure de Gibbs en volume fini μ_Λ^ξ.

Une remarque simple permet toutefois de contourner le problème de la non-réversiblité par rapport à μ. $\mathbf{\Gamma}$, réversible pour m, étant fixé, il existe un unique générateur de Markov $\mathbf{L}'$ réversible pour $d\mu = Z^{-1}\exp\Phi\ dm$ tel que

$$\int \mathbf{\Gamma}(f,g)\,d\mu = -\int \mathbf{L}'f\,g\,d\mu = \int \mathbf{\Gamma}'(f,g)\,d\mu \tag{47}$$

Écrivant $\mathbf{\Gamma}$ en fonction de $\mathbf{L}$ et utilisant la réversibilité pour m de ce dernier, on obtient l'expression de $\mathbf{L}'$ en fonction de $\mathbf{L}$:

$$\mathbf{L}'f = 1/2[\mathbf{L}f + \exp(-\Phi)(\mathbf{L}(f\ \exp\Phi) - f\mathbf{L}(\exp\Phi))] \tag{48}$$

Notre carré du champ de référence sera fourni par un noyau de transition $K(.,.)$ *symétrique*. Soit $\mathbf{L}f(x) = \int f(y)\,K(x,dy) - f(x)$ le générateur associé sur F et $\mathbf{L}_\Lambda$ le générateur produit sur F^Λ.

L'opérateur symétrique pour la mesure $\exp H_\Lambda^\xi\ dm^{\otimes\Lambda}$ est alors, en omettant les indices, $\mathbf{L}'f(x) = \sum_{k\in\Lambda} \mathbf{L}'_k f(x)$, produit des $\mathbf{L}'_k$ associés aux $\mathbf{L}_k$ induits par $\mathbf{L}$ au site k. Ici

$$\mathbf{L}'_k f(x) = 1/2[\mathbf{L}_k f + \exp(-H)(\mathbf{L}_k(f\ \exp H) - f\mathbf{L}_k(\exp H))]$$

avec $\mathbf{L}_k f(x) = \langle K(x_k, .), \nabla_k f(x)\rangle$. Posant $\alpha_k^b(x) = K(x_k, b)$, $\mathbf{L}_k$ n'est autre que le générateur associé à la famille de fonctions α_k.
Ceci étant, la formule de LEIBNITZ discrète (43) et la formule (45) de dérivation discrète de l'exponentielle nous assurent que

$$\begin{aligned}\mathbf{L}'_k f &= 1/2\, \langle \alpha_k, (1 + e^{\nabla_k H})\, (\nabla_k f)\rangle_k \\ &= \langle \alpha_k\, \tfrac{1+e^{\nabla_k H}}{2}, \nabla_k f\rangle_k \;=\; \langle \beta_k, \nabla_k f\rangle_k\end{aligned}$$

où $\beta_k^b(x) = K(x_k, b)\, \frac{1+e^{\nabla_k^b H(x)}}{2}$.

Finalement, $\mathbf{L}'$ n'est autre que le quasi-produit des générateurs associés, en chacun des sites $k \in \Lambda$, à la famille β_k.

Remarque 1.4.3. — La validité de ce calcul pour tout $\mathbf{L}^{\alpha_k}$ initial ne doit pas faire oublier que $\mathbf{L}'_k$ n'a de signification réelle (i.e. les moyennes $\int \mathbf{\Gamma}(f)\, d\mu$ et $\int \mathbf{\Gamma}'(f)\, d\mu$ coïncident) que lorsque $\mathbf{L}^{\alpha_k}$ est réversible pour la mesure uniforme ou encore lorsque la matrice donnée par α_k est symétrique ($\alpha_k^b(a) = \alpha_k^a(b)$).

Conditions sur le potentiel et contrôle de la courbure

Nous allons maintenant examiner une condition suffisante sur le hamiltonien du système pour que les familles de fonctions $(\alpha^{\Lambda,\xi})_{\Lambda\in\mathcal{F}}$ et $(\beta^{\Lambda,\xi})_{\Lambda\in\mathcal{F}}$ définissant respectivement la dynamique de GLAUBER et la dynamique $(\mathbf{L}')^\xi_\Lambda$ associée à une dynamique produit $\mathbf{L}_\Lambda$ donnée par un noyau satisfassent aux conditions $(\alpha^\Lambda Unif.)$ et $(\partial^1(\alpha^\Lambda)l^1)$ pour l'une et $(\beta^\Lambda Unif.)$ et $(\partial^1(\beta^\Lambda)l^1)$ pour l'autre.

Intéressons-nous tout d'abord au cas de $(\mathbf{L}')^\xi_\Lambda$. Rappelons que $\left[\left(\beta^{\Lambda,\xi}\right)^b_k\right](x) = K(x_k, b)\, \frac{1+e^{\nabla_k^b H^\xi_\Lambda(x)}}{2}$.

$K(x_k, b)$ étant majoré par 1, la condition $(\beta^\Lambda Unif.)$ est donnée dès que $\|\nabla_k^b H^\xi_\Lambda\|_u$ est bornée indépendamment de $\Lambda \in \mathcal{F}, \xi \in F^{\mathbb{Z}^\nu}, k \in \Lambda$ et $b \in F$.

Nous supposerons dorénavant la condition suivante satisfaite :

$$(\partial^2\, H\, l^1) \qquad \begin{array}{l}\textit{Il existe une fonction } \gamma \in l^1(\mathbb{Z}^\nu) \textit{ telle que,} \\ \textit{pour tous } b, c \in F \textit{ et tous } k, l \in \Lambda\ , \\ \qquad \left\|\nabla_l^c \nabla_k^b H_{\{k,l\}}\right\|_u \leqslant \gamma(k-l)\end{array}$$

Cette condition couvre celle que nous venons de mentionner sur $\|\nabla_k^b H^\xi_\Lambda\|_u$ (considérer $k = l$ dans $\mathbb{Z}^\nu$).

D'autre part, elle engendre $(\partial^1(\beta^\Lambda)l^1)$. En effet, si $l \neq k$,

$$\begin{aligned} 2\,\nabla_l^c \left(\beta^{\Lambda,\xi}\right)_k^b (x) &= \nabla_l^c \left(K(.{}_k, b)\,\left(1 + e^{\nabla_k^b H_\Lambda^\xi(.)}\right)\right)(x) \\ &= K(x_k, b)\,\nabla_l^c \left(e^{\nabla_k^b H_\Lambda^\xi(.)}\right)(x) \\ &= K(x_k, b)\, e^{\nabla_k^b H_\Lambda^\xi(x)}\,\left(e^{\nabla_l^c \nabla_k^b H_\Lambda^\xi(x)} - 1\right) \end{aligned}$$

D'où, sous la condition $(\partial^2\, H\, l^1)$, vu que $K(.,.) \leqslant 1$ et que $\left\|\nabla_l^c \nabla_k^b H_\Lambda^\xi\right\|_u$ est borné par $\|\gamma\|_u$,

$$\begin{aligned} 2\left|\nabla_l^c \left(\beta^{\Lambda,\xi}\right)_k^b (x)\right| &\leqslant e^{\gamma(0)}\left|e^{\nabla_l^c \nabla_k^b H_\Lambda^\xi(x)} - 1\right| \\ &\leqslant e^{\gamma(0)+\|\gamma\|_u}\left|\nabla_l^c \nabla_k^b H_\Lambda^\xi(x)\right| \\ &\leqslant C\gamma(k-l) \end{aligned}$$

Le cas $l = k$ s'obtient de la même manière.

Quant à la dynamique de Glauber, un calcul similaire utilisant de plus la formule discrète

$$\nabla_l^d \left(\frac{1}{f}\right) = \frac{-\nabla_l^d f}{f\left(f \circ \tau_l^d\right)}\,.$$

nécessitée par la présence de la fonction de partition permet partant de l'expression (46) des α_k^b de prouver les conditions $(\alpha^\Lambda Unif.)$ et $(\partial^1(\alpha^\Lambda)l^1)$ à partir de $(\partial^2\, H\, l^1)$.

Finalement, nous avons prouvé le

Théorème 1.4.4. — *Sous la condition d'existence d'une fonction $\gamma \in l^1(\mathbb{Z}^\nu)$ telle que, pour tous $b, c \in F$ et tous $k, l \in \mathbb{Z}^\nu$,*

$$\left\|\nabla_l^c \nabla_k^b H_{\{k,l\}}\right\|_u \leqslant \gamma(k-l)$$

la dynamique de Glauber *et le générateur réversible associé à une dynamique produit $\mathbf{L}_\Lambda$ donnée par un noyau ont leur courbure minorée indépendamment de la boîte finie Λ et de la condition extérieure ξ considérées.*

1.4.4. Isopérimétrie pour les systèmes de spins discrets. — Le résultat obtenu dans la précédente section va nous permettre d'obtenir des inégalités isopérimétriques pour la mesure de Gibbs en volume infini sous deux conditions :

- le contrôle de la constante de Sobolev-logarithmique de la dynamique de Glauber
- la condition $(\partial^2\, H\, l^1)$

Le théorème 1.4.4 nous permet de contrôler sous la condition $(\partial^2 H\, l^1)$ la courbure des dynamiques réversibles en volume fini $\mathbf{L}'_\Lambda$ indépendamment de la taille de Λ. Aussi obtient-on le pendant discret du théorème 1.3.4.

Théorème 1.4.5. — *Soit $(\mathcal{J}_\Lambda)_{\Lambda\in\mathbb{Z}^\nu}$ un potentiel d'interaction sur l'espace $F^{\mathbb{Z}^\nu}$ des configurations où F désigne un ensemble fini, potentiel satisfaisant à la condition $(\partial^2 H\, l^1)$.*

Supposons de plus qu'on ait le contrôle de la constante de SOBOLEV *logarithmique pour les mesures $\mu_\Lambda(dx_\Lambda) = \left(Z^\xi_\Lambda\right)^{-1} exp\, H^\xi_\Lambda(x_\Lambda)\ m^{\otimes\Lambda}(dx_\Lambda)$ et les générateurs $\mathbf{L}^\xi_\Lambda$ définissant la dynamique de* GLAUBER *ou la dynamique réversible associée à une dynamique donnée par un noyau, à savoir l'existence d'un $c' > 0$ tel que, pour tout $\Lambda \in \mathcal{F}$, toute condition extérieure ξ et toute f définie sur F^Λ, on ait*

$$\mathbf{Ent}_{\mu^\xi_\Lambda}(f^2) \leq c' \int \mathbf{\Gamma}^\xi_\Lambda(f)\, d\mu^\xi_\Lambda.$$

Alors, il existe une constante c telle que, pour tout volume fini Λ, toutes conditions extérieures ξ et toute fonction f sur F^Λ,

$$\mathcal{U}\left(\int f d\mu^\xi_\Lambda\right) \leq c \int \big(\mathcal{U}(f) + \Xi^\xi_\Lambda(f)\big) d\mu^\xi_\Lambda,$$

où $\Xi^\xi_\Lambda(f)$ est l'opérateur introduit dans le théorème 1.2.9 associé à la dynamique $\mathbf{L}^\xi_\Lambda$. En conséquence, on a les inégalités ensembliste et fonctionnelle suivante : pour toute partie A_Λ de F^Λ,

$$\mathcal{U}(\mu^\xi_\Lambda(A_\Lambda)) \leq c\ (\mu^\xi_\Lambda)_s(\partial A_\Lambda).$$

et, pour toute fonction f sur F^Λ,

$$\mathcal{U}\left(\int f d\mu^\xi_\Lambda\right) \leq \int \big(\mathcal{U}(f) + c\,\Upsilon^\xi_\Lambda(f)\big) d\mu^\xi_\Lambda$$

où l'on rappelle que

$$\Upsilon^\xi_\Lambda(f) = \sum_{k\in\Lambda} \left\langle \sqrt{\left(\alpha^\xi_\Lambda\right)_k}, |\nabla_k f| \right\rangle_k.$$

Les contrôles de la constante de SOBOLEV logarithmique pour la dynamique de GLAUBER ou la dynamique réversible associée à une dynamique donnée par un noyau sont équivalents. Aussi les travaux de LAROCHE s'appliquent-ils à notre étude. En effet, il y est établi que sous la condition de décroissance exponentielle de $\left\|\nabla^c_l \nabla^b_k H^\xi_\Lambda\right\|_u$, recoupant notre condition, le contrôle de la constante de log-SOBOLEV pour la dynamique de GLAUBER est équivalent à la condition de mélange de DOBRUSHIN et SCHLOSMAN ou encore de mélange faible introduite dans [Lar95].

Par conséquent, **sous la condition de décroissance exponentielle des variations secondes du potentiel, la condition de mélange pour le modèle**

implique le contrôle de la constante d'isopérimétrie ensembliste. Pour un potentiel de portée finie sur un ensemble discret, B.ZEGARLINSKI a récemment montré (cf. [Zeg01]) que l'on peut obtenir mieux : la condition de mélange fort de DOBRUSHIN et SCHLOSMAN donne en fait dans ce cadre l'existence d'une inégalité isopérimétrique gaussienne fonctionnelle du même type que celle que nous avons obtenue pour des spins continus dans la section 1.3.

Remerciements

Je tiens à exprimer envers D.BAKRY mes très sincères remerciements pour son aide constante et les longues discussions qui ont engendré cet article. Nombre d'idées lui en sont dues. Je remercie grandement M.LEDOUX pour son aide, sa disponibilité et son intérêt pour ce travail. Merci à F.MALRIEU et G.SCHEFFER pour leur écoute ainsi qu'à tous les doctorants du LSP pour l'agréable dynamique qui règne entre nous.

CHAPITRE 2

SPECTRAL GAP FOR LOG-CONCAVE PROBABILITY MEASURES ON THE REAL LINE

2.1. Introduction

A log-concave measure $\mu(dx)$ on the real line is an absolutely continuous measure with respect to the LEBESGUE measure, with density $\exp(-\Phi)$, where Φ is a convex function. We assume that μ is a finite measure, or equivalently that Φ goes to infinity as $|x| \to +\infty$, and normalize μ to a probability. As we shall see later, such a measure always satisfies a POINCARÉ inequality, that is, there exists a constant $C_2 > 0$ such that, for any smooth function with compact support $f : \mathbb{R} \to \mathbb{R}$,

$$\mathbf{Var}_\mu(f) = \int f^2 \, d\mu - \left(\int f \, d\mu \right)^2 \leqslant C_2 \int f'^2 \, d\mu. \tag{49}$$

When Φ is smooth enough, this well known inequality is strongly related to the natural operator associated with μ which acts on a smooth function f by

$$\mathbf{L} f = f'' - \Phi' f'.$$

One may easily check that $\mathbf{L}$ satisfies the fundamental integration by part formula, for any smooth compactly supported function f,

$$\forall g, \quad \int \mathbf{L} f \, g \, d\mu = - \int f' \, g' \, d\mu.$$

Hence, this unbounded self-adjoint operator on $\mathbf{L}^2(\mu)$ is negative. Moreover, the first non-zero eigenvalue λ_1 of $-\mathbf{L}$, so called *spectral gap*, is given by the inverse of the best constant C_2 in (49). From now on, C_2 will stand for this best constant. As a consequence, the constant C_2 determines the ergodic behaviour of the semigroup generated by $\mathbf{L}$. The latter is a Markovian ergodic semigroup of operators $(\mathbf{P}_t)_{t \geqslant 0}$ on $\mathbf{L}^2(\mu)$ and inequality (49) is equivalent to the exponential rate convergence

$$\|\mathbf{P}_t f - \mu(f)\|_{2,\mu} \leqslant e^{-\frac{2}{C_2} t} \, \mathbf{Var}_\mu(f),$$

where we note $\mu(f) = \int f \, d\mu$.

These well known results are valid in a much more general setting and can be found in a large literature (see for instance [Bak94, Led99, Roy99, ABC+00]). They

emphasize the importance of the constant C_2 and justify the following question : how does this constant depend on the measure μ ?

A complete answer, based on MUCKENHOUPT's paper on HARDY inequalities [Muc72], has been recently given for measures on a one-dimensional space. This is due to MICLO for measures on $\mathbb{Z}$ ([Mic99a, Mic99b]) and to BOBKOV and GÖTZE for general measures on the real line (see [BG99]). They get a functional of the measure μ which describes the exact order of constant C_2. Let's now briefly present this result in our framework.

Let Φ be a function on $\mathbb{R}$ (not necessarily convex) and

$$\mu(dx) = Z^{-1} \exp\left(-\Phi(x)\right) dx$$

the associated BOLTZMANN measure supposed to be finite and normalized to a probability. Let m be the (unique) median of μ. Define

$$B_+ = \sup_{x \geqslant m} \left(\int_m^x e^{\Phi(t)}\, dt \int_x^{+\infty} e^{-\Phi(t)}\, dt \right)$$

and

$$B_- = \sup_{x \leqslant m} \left(\int_x^m e^{\Phi(t)}\, dt \int_{-\infty}^x e^{-\Phi(t)}\, dt \right).$$

Then we call $B(\Phi) \stackrel{\text{def.}}{=} \max\left(B_+, B_-\right)$ the MUCKENHOUPT functional. Now, there exists a POINCARÉ inequality for μ if and only if $B(\Phi)$ is finite, and then we have

$$\frac{1}{2}\, B(\Phi) \leqslant C_2 \leqslant 4\, B(\Phi). \tag{50}$$

See [ABC+00] for instance for more details. Notice that BOBKOV and GÖTZE get similar result and functional to characterize the logarithmic SOBOLEV constant of the measure μ.

Let's now suppose again that Φ is convex. It is not so easy to be convinced that $B(\Phi)$ is finite, even if it's true. The aim of this paper is to find a simpler and clearer functional of Φ which still describes the order of the POINCARÉ constant of the measure μ. The following theorem gives an answer to this question.

Theorem 2.1.1. — *Let Φ be a convex function on $\mathbb{R}$. Then the probability measure*

$$\mu(dx) = Z^{-1} \exp\left(-\Phi(x)\right) dx$$

satisfies a POINCARÉ *inequality*

$$\mathbf{Var}_\mu(f) = \int f^2\, d\mu - \left(\int f\, d\mu \right)^2 \leqslant C_2 \int f'^2\, d\mu \tag{51}$$

and there exists a universal constant D (not depending on Φ) such that, if $C_2(\Phi)$ stands for the best constant in (51)*, then*

$$\left(\int |x-m|\, \mu(dx) \right)^2 \leqslant C_2(\Phi) \leqslant D \left(\int |x-m|\, \mu(dx) \right)^2. \tag{52}$$

In other words, up to universal constants, the POINCARÉ constant of a log-concave measure is nothing but the square of the mean value of the distance to the median.

Remark 2.1.2. — *We shall see that the constant D could be chosen equal to* 16. *This precise value will not be deduced from our computations but will derive from a closer analysis of the method described in* [Bob99].

Remark 2.1.3. — *The first result which could indicate the importance of the distance function for the estimate of the* POINCARÉ *(or rather log-*SOBOLEV *) constant was* WANG*'s theorem (see* [Wan97]*). This theorem is described in Appendix A in a high generality. Applied to log-concave measures on Euclidean spaces, it says that such a measure satisfies a log-*SOBOLEV *inequality if and only if the square of the distance function is exponentially integrable. The log-*SOBOLEV *constant obtained by* WANG *depends in a non trivial manner on this exponential integral.*

A simple lemma is hidden behind the theorem 2.1.1. It concerns a universal bound of a convex function Φ in terms of the median of the associated log-concave measure. We now point out this simple result.

Lemma 2.1.4. — *Let Φ be a convex function on $\mathbb{R}$ which goes to infinity at infinity, and let m be the median of the associated log-concave probability measure. Then*

$$\Phi \geqslant \Phi(m) - \log 2.$$

As we shall see later, S. BOBKOV proved that the best constant in the CHEEGER inequality satisfied by a one-dimensional log-concave probability measure is $e^{\Phi(m)}/2$. Hence, the previous lemma says that the inverse of the CHEEGER constant is an upper bound of the density of such a measure.

Lemma 2.1.4 will be useful in section 2.3.5 (and will be proved there) to solve explicitly the variational problem on a simple enough class of functions. Its proof is a direct consequence of the earlier lemma 2.3.2. The latter characterizes the convex functions on a half-line $[m, \infty[$ with a convex extension on the whole line the associated log-concave measure of which has median m.

After having obtained the theorem 2.1.1, we discovered a paper of S.BOBKOV (see [Bob99] for details) in which the same question is in particular approached. The functional which BOBKOV considers is the variance of the measure and he get

$$\mathbf{Var}(\mu) \leqslant C_2(\Phi) \leqslant 12\mathbf{Var}(\mu),$$

where $\mathbf{Var}(\mu)$ is defined by

$$\mathbf{Var}(\mu) = \int x^2\, \mu(dx) - \left(\int x\, \mu(dx)\right)^2.$$

His approach seems to be more complete than our's. Namely, he gets a similar control for the log-SOBOLEV constant of a log-concave measure on the real line in terms of

the ORLICZ norm

$$\|x - \mathbf{E}_\mu(x)\|_\Psi^2$$

associated with the function $\Psi(s) = \exp(s^2) - 1$. Moreover, he extends his result to log-concave measures on the Euclidean space of dimension n : he gets an upper control of the constant C_2 in terms of $\|x\|_{2,\mu}^2$ but which doesn't remain of the good order in general.

Nevertheless, we think that our approach, based on totally different tools, is worth being explained. Whereas BOBKOV point of view make use of some control of the best constant in the CHEEGER inequality, we carry a variation calculus in the set of convex functions through, by the means of the MUCKENHOUPT functional.

The first section is devoted to a direct argument due to D.BAKRY to get a POINCARÉ inequality in particular for log-concave probability measures on $\mathbb{R}$. Afterwards, we turn to the variational proof of theorem 2.1.1. We end with a brief adaptation of BOBKOV's argument to find again our result.

2.2. A simple criterion for POINCARÉ inequality

Before entering into the variational proof of the theorem 2.1.1 and its technical computations, we would like to present a short and formal argument due to D.BAKRY in order to convince the reader from now on that convex functions do satisfy a POINCARÉ inequality.

In this section and only here, we will use the general setting of a diffusion operator. We don't want here to introduce properly this setting and refer to the general introduction.

Let $\mathbf{L}$ be a Markovian diffusion generator reversible w.r.t. the measure μ on a measurable space $(\mathbb{E}, \mathcal{F})$. Let $\mathbf{\Gamma}$ be the associated square field operator. The diffusion hypothesis amounts to

$$\mathbf{\Gamma}(\Psi(f)) = (\Psi'(f))^2\, \mathbf{\Gamma}(f).$$

These two operators are essentially related by the already introduced integration by part formula

$$\forall g, \quad \int \mathbf{L}f\, g\, d\mu = -\int \mathbf{\Gamma}(f, g)\, d\mu.$$

A function u is called a 1-LIPSCHITZ function (or a contraction) if $\mathbf{\Gamma}(u) \leqslant 1$. We are now able to state the

Proposition 2.2.1. — [1] *Let $\mathbf{L}$, $\mathbf{\Gamma}$ and μ as before. Suppose there exists a contraction u such that, for some real numbers $\alpha > 0$ and β, and some point $x_0 \in \mathbb{E}$,*

$$(\mathbf{L}u + \alpha)\, d\mu \leqslant \beta\, \delta_{x_0} \tag{53}$$

[1] I give D.BAKRY my sincere thanks for allowing me to write this argument here.

as measures. Then the measure μ satisfies the following POINCARÉ *inequality, for all f,*

$$\mathbf{Var}_\mu(f) \leqslant \frac{4}{\alpha^2} \int \mathbf{\Gamma}(f)\, d\mu. \tag{54}$$

Proof

◀ As $d\mu \leqslant \alpha^{-1}\,(\delta_{x_0} - \mathbf{L}u\, d\mu)$, we have

$$\begin{aligned}\int (f - f(x_0))^2\, d\mu &\leqslant -\alpha^{-1} \int (f - f(x_0))^2\, \mathbf{L}u\, d\mu \leqslant \alpha^{-1} \int \mathbf{\Gamma}\big((f - f(x_0))^2, u\big) \\ &\leqslant 2\,\alpha^{-1} \int (f - f(x_0)) \mathbf{\Gamma}(f, u)\, d\mu \\ &\leqslant 2\,\alpha^{-1} \int (f - f(x_0))\, \sqrt{\mathbf{\Gamma}(f)}\, d\mu \\ &\leqslant 2\,\alpha^{-1} \bigg(\int (f - f(x_0))^2\, d\mu\bigg)^{\frac{1}{2}} \bigg(\int \mathbf{\Gamma}(f)\, d\mu\bigg)^{\frac{1}{2}}.\end{aligned}$$

The two last inequalities are due to CAUCHY-SCHWARZ inequality first for the bilinear form $\mathbf{\Gamma}$ and then in $\mathbf{L}^2(\mu)$. Hence,

$$\mathbf{Var}_\mu(f) \leqslant \int (f - f(x_0))^2\, d\mu \leqslant \frac{4}{\alpha^2} \int \mathbf{\Gamma}(f)\, d\mu.$$

▶

Let us come back to BOLTZMANN measures on $\mathbb{R}$. Suppose that, for any $x \in \mathbb{R}$,

$$\Phi'(x)\, \mathrm{sign}(x - x_0) \geqslant \alpha. \tag{55}$$

Then one easily check that $|x - x_0|$ satisfies (53) and inequality (54) occurs. Now, if Φ is convex, one may always perturb Φ by a bounded function in order to obtain a function satisfying (55). And we get a POINCARÉ inequality for Φ thanks to the bounded perturbation property of this type of inequalities (see [ABC+00] for instance).

2.3. Variational proof of the main result

From now on, we call the median m of the measure μ the median of Φ as well. Proving theorem 2.1.1 amounts to find a constant D such that, for any convex function Φ with median m,

$$B(\Phi) \leqslant D \bigg(\int |x - m|\, \mu(dx)\bigg)^2. \tag{56}$$

The other inequality is immediate indeed since

$$\int |x - m|\, \mu(dx) \leqslant \int |x - \mathbf{E}_\mu(x)|\, \mu(dx) \leqslant \sqrt{\mathbf{Var}_\mu(x)} \leqslant \sqrt{C_2},$$

by the $\mathbf{L}^1$ variational characterization of the median, the CAUCHY-SCHWARZ inequality and the LIPSCHITZ property of the function x.

2.3.1. First reductions of the problem. — One may easily reduce the number of parameters by using simple transformations on convex functions. First, it is enough to show that

$$B_+(\Phi) \leqslant D \left(\int_m^\infty |x-m| \, \mu(dx) \right)^2 .$$

By homogeneity, we may assume that $\Phi(m) = 0$. By changing Φ into $\Phi(\cdot - m)$, one may suppose that $m = 0$ and the condition becomes, for any convex function Φ with median 0 which vanishes at 0 and any $x > 0$,

$$\int_0^x e^{\Phi(t)} \, dt \int_x^\infty e^{-\Phi(t)} \, dt \preccurlyeq \left(\int_0^\infty t \, e^{-\Phi(t)} \, dt \right)^2 \left(\int_0^\infty e^{-\Phi(t)} \, dt \right)^{-2},$$

where $\preccurlyeq$ stands for $\leqslant$ up to a universal constant. Now, remark that changing Φ into $\Phi(\cdot/x)$ allows us to consider $x = 1$. Hence, the problem is reduced to minimize the functional

$$\frac{1}{Z_2(\Phi) \, E_1(\Phi)} \left(\frac{M_1(\Phi) + M_2(\Phi)}{Z_1(\Phi) + Z_2(\Phi)} \right)^2$$

on the set of convex functions with median 0 which vanish at 0. Here, we put

$$Z_1(\Phi) = \int_0^1 e^{-\Phi(t)} \, dt, \qquad M_1(\Phi) = \int_0^1 t \, e^{-\Phi(t)} \, dt, \quad E_1(\Phi) = \int_0^1 e^{\Phi(t)} \, dt,$$
$$Z_2(\Phi) = \int_1^\infty e^{-\Phi(t)} \, dt, \qquad M_2(\Phi) = \int_1^\infty t \, e^{-\Phi(t)} \, dt.$$

We split the integrals in order to dissociate the two parts of Φ on $[0,1]$ and $[1,\infty[$ respectively.

Definition 2.3.1. — *The Z functionals will be called the* mass *functionals, the M one's the* mean *functionals and E_1 the* inverse mass *functional.*

As we are only concerned with the restriction of Φ to the half line $[0,\infty[$, it would be interesting to characterize the convex functions on this interval which have a convex extension with median 0 on the whole line. This question is settled by the following lemma.

Lemma 2.3.2. — *Let Φ be a convex function on $\mathbb{R}^+$ which goes to infinity at $+\infty$. Then Φ is the restriction of a convex function on $\mathbb{R}$ which goes to infinity at infinity as well and with median 0, if and only if,*

$$\Phi'(0) \geqslant \frac{-1}{Z_1(\Phi) + Z_2(\Phi)}.$$

Proof
◀ This is a simple argument. The case of interest is when $\Phi'(0_+) \leqslant 0$. If Φ has a convex extension, say $\overline{\Phi}$, then $\overline{\Phi}'(0_-) \leqslant \Phi'(0_+)$ and $\overline{\Phi}(x) \geqslant L(x)$, $x \in \mathbb{R}$, where L is the linear function vanishing at 0 and with slope $\overline{\Phi}(0_-)$. Denote by Z_- the mass functional of a function on $\mathbb{R}^-$. We have $Z_-\big(\overline{\Phi}\big) = Z_1(\Phi) + Z_2(\Phi)$ as the median is 0. And $Z_-\big(\overline{\Phi}\big) \leqslant Z_-(L)$. Conversely, the condition make sure that the linear function L with slope $\Phi'(0_+)$ at 0 is a convenient extension. ▶

As a consequence, let's define the class $\mathcal{C}$ of convex functions on $\mathbb{R}^+$ going to infinity at $+\infty$, vanishing at 0 and such that

$$\Phi'(0) \geqslant \frac{-1}{Z_1(\Phi) + Z_2(\Phi)}.$$

Let's now fix some denominations.

Definition 2.3.3. — *We will consider continuous piecewise linear functions on* $\mathbb{R}^+$. *Among them, the* k-piecewise linear *functions are those with* k *linear pieces. The* $k-1$ knots *of such a function are the points where the derivative is broken.*

The proof of the theorem 2.1.1 will derive from the two following propositions.

Proposition 2.3.4 **(Restriction to a simple class of convex functions)**
Let

$$K(\Phi) = \frac{1}{Z_2(\Phi)\, E_1(\Phi)} \left(\frac{M_1(\Phi) + M_2(\Phi)}{Z_1(\Phi) + Z_2(\Phi)} \right)^2$$

be the functional of interest on the class $\mathcal{C}$ *and* $\mathcal{T} \subset \mathcal{C}$ *be the class of* 3*-piecewise linear convex functions on a compact interval* $[0, \beta]$, $\beta \geqslant 1$, *with a knot at point* 1 *and the other inside the interval* $[0, 1]$. *Then, we have*

$$\min_{\Phi \in \mathcal{C}} K(\Phi) \geqslant \frac{1}{2} \min_{\Phi \in \mathcal{T}} K(\Phi).$$

Proposition 2.3.5. — *The functional* K *on the class* $\mathcal{T}$ *introduced before is bounded from below by a positive constant.*

Remark 2.3.6. — *We shall sometimes call the function of* $\mathcal{T}$ minimizing functions *of the variational problem even if they are not exactly.*

Subsections 2.3.2 to 2.3.4 are devoted to the proof of the proposition 2.3.4 while the proposition 2.3.5 will be proved in subsection 2.3.5.

2.3.2. Existence conditions for convex functions with fixed parameters. — As previously, we dissociate the two parts of Φ, Φ_1 on $[0,1]$ and Φ_2 on $[1,\infty[$ and note

$$\mathcal{C}^{\otimes} = \{(\Phi_1, \Phi_2)/\Phi_1 \mathbb{1}_{[0,1]} + \Phi_2 \mathbb{1}_{]1,\infty[} \in \mathcal{C}\}.$$

The pair (Φ_1, Φ_2) belongs to $\mathcal{C}^{\otimes}$ if

(i) Φ_1 is convex on $[0,1]$,

(ii) Φ_2 is convex on $[1,\infty[$ and goes to infinity at $+\infty$,

(iii) $\Phi_1(0) = 0$ et $\Phi_1(1) = \Phi_2(1)$,

(iv) $\Phi_1'(0) \geqslant -\frac{1}{Z_1(\Phi_1)+Z_2(\Phi_2)}$,

(v) $\Phi_1'(1_-) \leqslant \Phi_2'(1_+)$.

We now split the contributions of Φ_1 and Φ_2 when we fix some parameters as the value at point 1 or the masses Z_1 or Z_2. So we consider the first and second projections $\pi_1(\mathcal{C}^{\otimes})$ and $\pi_2(\mathcal{C}^{\otimes})$ of the class $\mathcal{C}^{\otimes}$.

2.3.2.1. *Second projection.* — Let x and b be two real numbers. Note $\pi_2(x,b)$ the set of functions $\Phi_2 \in \pi_2(\mathcal{C}^{\otimes})$ such that $\Phi_2(1) = x$ and $\Phi_2'(1) = b$.

Definition 2.3.7. — *In the rest of the paper, $\overline{x}$ will stand for the linear function on $[0,1]$ such that $\overline{x}(0) = 0$ and $\overline{x}(1) = x$. Its* mass *will be noted*

$$F(x) \stackrel{def.}{=} Z_1(\overline{x}) = \frac{1-e^{-x}}{x}.$$

We will need later to consider the following function

$$F_x(u) \stackrel{def.}{=} \frac{e^{-u}-e^{-x}}{x-u},$$

which is nothing but the slope of the secant line of the exponential function between $-u$ and $-x$.

Remark that F and F_x are both strictly decreasing. We now can state the two following simple results.

Lemma 2.3.8. — *The following conditions are equivalent.*

(i) $\pi_2(x,b) \neq \emptyset$

(ii) $b \geqslant x > -\log(2)$

In that case, if $\Phi_2 \in \pi_2(x,b)$ then $(\overline{x}, \Phi_2) \in \mathcal{C}^{\otimes}$.

Corollary 2.3.9. — *Let x be such that $b \geqslant x > -\log(2)$. Then*

$$\pi_2(x,b) = \left\{\Phi_2 \text{ convex on } [1,\infty[\quad : \quad x \geqslant -\frac{1}{F(x)+Z_2(\Phi_2)}\right\} \cap \{\Phi_2'(1) = b\}$$

The proof of lemma 2.3.8 is omitted now but the reader can check it easily. A very similar one will be given later in proposition 2.3.19. The main idea is that as soon as Φ_2 is in $\pi_2(x,b)$, we can extend Φ_2 by the linear function $\overline{x}$ to obtain a function in $\mathcal{C}$. It remains to check that $x > -1/Z_1(\overline{x})$ is exactly $x > -\log 2$.

Subsequently, we will fix the mass $Z_2(\Phi_2)$ on $[1,\infty[$ of the functions we are interested with. With this aim in view, let us consider its possible values

$$Z_2(\pi_2(x,b)) =]0, \mathcal{Z}_2(x,b)].$$

The explicit value of $\mathcal{Z}_2(x,b)$ may be computed as

$$\mathcal{Z}_2(x,b) = \begin{cases} e^{-x}/b & \text{si } x \geqslant 0 \text{ ou si } x < 0 \text{ et } b > e^{-x}/\eta(x) \\ \eta(x) & \text{si } x < 0 \text{ et } x \leqslant b \leqslant e^{-x}/\eta(x) \end{cases} .$$

where

$$\eta(x) = -F(x) - 1/x = \frac{e^{-x}-2}{x} \tag{57}$$

is an increasing homeomorphism from $]-\log(2), 0[$ to $]0,\infty[$. This notation will be used throughout the paper.

The two following remarks make the problem easier to handle.

Remark 2.3.10. — *Let's consider $b \geqslant x \geqslant -\log(2)$ and $0 < Z_2 \leqslant \mathcal{Z}_2(x,b)$. Note $\pi_2(x,b,Z_2)$ the set of functions Φ_2 in $\pi_2(x,b)$ with fixed mass $Z_2(\Phi_2) = Z_2$. On $\pi_2(x,b,Z_2)$, the functional $M_2(\Phi_2)$ is minimal as the mass Z_2 is located the nearest from 1 as possible i.e. when Φ_2 is the linear function*

$$l_{x,b,Z_2} = (x + b(\cdot - 1))\mathbb{1}_{[1,\alpha[} + \infty \mathbb{1}_{[\alpha,\infty[}.$$

The threshold $\alpha = 1 - \frac{1}{b}\log(1 - b\,Z_2\,e^x)$ is fixed in order that the mass is Z_2. In short,

$$\min_{\Phi_2 \in \pi_2(x,b,Z_2)} M_2(\Phi_2) = M_2(l_{x,b,Z_2}) \overset{def.}{=} M_2(x,b,Z_2).$$

Remark 2.3.11. — *Now, if x, b, Z_2 are fixed as previously, the set $\pi_1(\pi_2^{-1}(\Phi_2)) = \pi_1(x,b,Z_2)$ of functions Φ_1 such that $(\Phi_1,\Phi_2) \in \mathcal{C}^{\otimes}$ does not depend on the considered function $\Phi_2 \in \pi_2(x,b,Z_2)$. The class $\pi_1(x,b,Z_2)$ is then the set of convex functions Φ on $[0,1]$ such that*

(i) $\Phi(0) = 0$ *and* $\Phi(1) = x$,

(ii) $\Phi'(1) \leqslant b$ *and* $\Phi'(0) \geqslant -\frac{1}{Z_1(\Phi)+Z_2}$.

These two fundamental remarks make sure that a minimizing function of the functional K on $\mathcal{C}$ introduced in proposition 2.3.4 is necessarily linear on an interval $[1,\beta]$ and infinite after β. The minimizing problem is then reduced to the following

Problem 1. — *There exists a positive ϵ such that, for any real numbers x and b such that $b \geqslant x > -\log(2)$, any mass $0 < Z_2 \leqslant \mathcal{Z}_2(x,b)$ and any function $\Phi_1 \in \pi_1(x,b,Z_2)$,*

$$\epsilon\, Z_2 \leqslant \frac{1}{E_1(\Phi_1)}\left(\frac{M_1(\Phi_1)+M_2(x,b,Z_2)}{Z_1(\Phi_1)+Z_2}\right)^2.$$

2.3.2.2. *First projection.* — The next step will consist in fixing the mass of the function $\Phi_1 \in \pi_1(x,b,Z_2)$. As for Z_2, we may determine its possible values. We get

$$Z_1(\pi_1(x,b,Z_2)) = [F(x), \mathcal{Z}_1(x,b,Z_2)].$$

The exact value of $\mathcal{Z}_1(x,b,Z_2)$ will be given below.

Let Z_2 be a non negative real number. Let us consider the following sets of convex functions on the interval $[0,1]$:

- $\pi_1(x) = \left\{\Phi \,:\, \Phi(0)=0, \Phi(1)=x, \Phi'(0) > -\frac{1}{Z_1(\Phi)}\right\}$,
- $\tilde{\pi}_1(x,Z_1) = \pi_1(x) \cap \{Z_1(\cdot) = Z_1\}$,
- $\pi_1(x,Z_2) = \left\{\Phi \,:\, \Phi(0)=0, \Phi(1)=x, \Phi'(0) \geqslant -\frac{1}{Z_1(\Phi)+Z_2}\right\}$,
- $\pi_1(x,Z_1,Z_2) = \pi_1(x,Z_2) \cap \{Z_1(\cdot) = Z_1\}$,
- $\pi_1(x,b,Z_1,Z_2) = \pi_1(x,Z_1,Z_2) \cap \{\Phi \,:\, \Phi'(1) \leqslant b\}$.

Proposition 2.3.12. — *Let Z_1 and Z_2 be positive real numbers. Then we have the following existence conditions*

(i) $\pi_1(x) \neq \emptyset \Longleftrightarrow x > -\log(2)$

(ii) $\tilde{\pi}_1(x,Z_1) \neq \emptyset \Longleftrightarrow x > -\log(2)$ *and* $F(x) \leqslant Z_1 < F(-\log(2)) = 1/\log(2)$

(iii) $\pi_1(x,Z_2) \neq \emptyset \Longleftrightarrow x \geqslant \eta^{-1}(Z_2)$

(iv) $\pi_1(x,Z_1,Z_2) \neq \emptyset \Longleftrightarrow x \geqslant \eta^{-1}(Z_2)$ *and* $F(x) \leqslant Z_1 \leqslant F(\eta^{-1}(Z_2))$

(v) $\pi_1(x,b,Z_1,Z_2) \neq \emptyset \Longleftrightarrow b \geqslant x$ *and* $F(x) \leqslant Z_1 \leqslant \mathcal{Z}_1(x,b,Z_2)$

Moreover, the exact value of $\mathcal{Z}_1(x,b,Z_2)$ is given by

$$\mathcal{Z}_1(x,b,Z_2) = -\frac{1}{b} - Z_2 - \frac{1-\frac{x}{b}}{\eta^{-1}\left(\frac{Z_2-\frac{x}{b}\eta(x)}{1-\frac{x}{b}}\right)}.$$

Here, η is the homeomorphism introduced in (57).

Remark 2.3.13. — *If $b \leqslant b'$, $\pi_1(x,b,Z_1,Z_2) \subset \pi_1(x,b',Z_1,Z_2)$. The existence condition* (v) *then implies that the function $\mathcal{Z}_1(x,b,Z_2)$ is increasing on the variable b. As b goes to $+\infty$,* (v) *becomes*

$$F(x) \leqslant -Z_2 - \frac{1}{\eta^{-1}(Z_2)} = F\big(\eta^{-1}(Z_2)\big)$$

and we find condition (iii) *again i.e. $x \geqslant \eta^{-1}(Z_2)$.*

Proof of proposition 2.3.12. — We concentrate on the proof of (v). We get the other cases by taking limit values of the parameters.

Let Φ be in $\pi_1(x, b, Z_1, Z_2)$. Put

$$a \stackrel{\text{def.}}{=} -\frac{1}{Z_1 + Z_2}.$$

The fixed conditions on the values and derivatives at points 0 and 1 imply that $\Psi_{a,x,b} \leqslant \Phi \leqslant \overline{x}$, where

$$\Psi_{a,x,b}(t) = a\, t\, \mathbb{1}_{[0,\alpha]} + (x + b\,(t-1))\, \mathbb{1}_{]\alpha,1]}$$

is the 2-piecewise linear function Ψ such that $\Psi(0) = 0$, $\Psi(1) = x$, $\Psi'(0) = a$ and $\Psi'(1) = b$. Remark that the convexity of Φ implies that

$$b \geqslant x = \Phi(1) \geqslant \Phi'(0) \geqslant a$$

so that the previous function exists. The knot α of Ψ is then $\alpha = \frac{b-x}{b-a}$. Consequently,

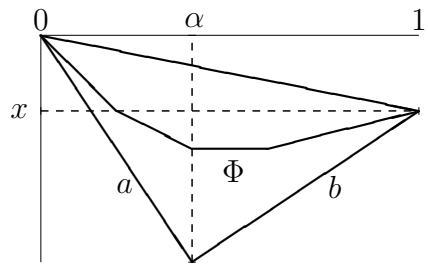


FIGURE 1.

$$F(x) \leqslant Z_1 \leqslant Z_1(\Psi_{a,x,b}). \tag{58}$$

Well, an easy computation shows that, if Ψ is the 2-piecewise linear function on $[0,1]$ with $\Psi(0) = 0$, $\Psi(1) = x$ and with knot α such that $\Psi(\alpha) = u$, the associated mass is

$$Z_1(\Psi) = \alpha\, F(u) + (1-\alpha)\, F_x(u)$$

where the function $F_x(u)$ has been introduced in definition 2.3.7. Hence,

$$Z_1(\Psi_{a,x,b}) = \frac{1 - e^{-a\,\alpha}}{a} + \frac{e^{-a\,\alpha} - e^{-x}}{b} \tag{59}$$

$$= \left(Z_1 + Z_2 + \frac{1}{b}\right)\left(e^{\left(1-\frac{x}{b}\right)\frac{1}{Z_1+Z_2+\frac{1}{b}}} - 1\right) + \frac{1 - e^{-x}}{b}. \tag{60}$$

Let's rewrite the condition

$$Z_1 \leqslant Z_1(\Psi_{a,x,b})$$

thanks to this previous expression. Put $X = \frac{1}{Z_1+Z_2+\frac{1}{b}}$. We get the following equivalent condition

$$\frac{1}{X} - Z_2 - \frac{1}{b} \leqslant \frac{e^{\left(1-\frac{x}{b}\right)X} - 1}{X} + \frac{1-e^{-x}}{b}$$

or

$$\eta\left(\tilde{X}\right) = -\frac{1}{\tilde{X}} + \frac{e^{-\tilde{X}} - 1}{\tilde{X}} \leqslant \tilde{Z}_2$$

if we put

$$\tilde{X} = -\left(1 - \frac{x}{b}\right) X$$

and

$$\tilde{Z}_2 = \left(1 - \frac{x}{b}\right)^{-1} \left(\frac{2 - e^{-x}}{b} + Z_2\right).$$

Finally, if we write properly the inequality

$$\tilde{X} \leqslant \eta^{-1}\left(\tilde{Z}_2\right),$$

we see that conditions

$$Z_1 \leqslant Z_1(\Psi_{a,x,b})$$

and

$$Z_1 \leqslant \mathcal{Z}_1(x, b, Z_2)$$

are equivalent. So, we proved the necessity.

Conversely, under the condition given in (v), we have, as it was seen before,

$$F(x) \leqslant Z_1 \leqslant Z_1(\Psi_{a,x,b})$$

where $a = -\frac{1}{Z_1+Z_2}$. Well, as β goes from α to 1, the mass $Z_1(\Phi_{x,\beta,b})$ of the 2-piecewise linear function $\Phi_{x,\beta,b}$, such that $\Phi'_{x,\beta,b}(1) = b$ and with knot β, achieves all values between $F(x)$ and $Z_1(\Psi_{a,x,b})$, and in particular Z_1. Consequently, for a well chosen

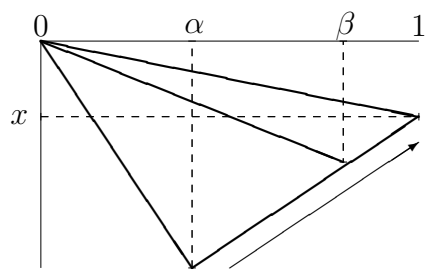


FIGURE 2.

β, $\psi_{x,\beta,b} \in \pi_1(x, b, Z_1, Z_2)$. And the proof is complete. □

By fixing Z_1, our functional is highly simplified and the minimizing problem is reduced to

Problem 2. — *There exists an $\epsilon > 0$ such that, for any $b \geqslant x > -\log(2)$, any masses $0 < Z_2 \leqslant \mathcal{Z}_2(x,b)$ and $F(x) \leqslant Z_1 \leqslant \mathcal{Z}_1(x,b,Z_2)$, and any function $\Phi \in \pi_1(x,b,Z_1,Z_2)$,*

$$\epsilon\, Z_2\,(Z_1+Z_2)^2 \leqslant \frac{\left(M_1(\Phi)+M_2(x,b,Z_2)\right)^2}{E_1(\Phi)}. \tag{61}$$

2.3.3. Mean and inverse mass functionals on 2-piecewise linear functions. — As we have already remarked, a minimizing function of the initial problem has to be linear on a compact interval $[1,\beta]$ and infinite on $]\beta,\infty]$. We now turn our attention to the extremal functions on the interval $[0,1]$, that is the minimizing functions of problem 2.

As we shall see, we are not able to determine the exact minimizing functions of the functional

$$K_{M,E}(\Phi) \stackrel{\text{def.}}{=} \frac{\left(M_1(\Phi)+M\right)^2}{E_1(\Phi)+E}, \tag{62}$$

with M, $E \geqslant 0$, in general. Nevertheless, up to the constant $1/2$ of proposition 2.3.4, these are linear with two pieces. To prove this, we will make use of an induction argument. The main step is to prove that, on the set of piecewise linear convex functions with less than three pieces, the extremal values of the functionals M_1 and E_1 are achieved by some 2-piecewise linear functions.

With this aim in view, we start with the characterization of the set $\pi_1^2(x,b,Z_1,Z_2)$ of the functions in $\pi_1(x,b,Z_1,Z_2)$ which are 2-piecewise linear and we compute precisely the functionals involved in our problem on this set.

2.3.3.1. *Precise values of the functionals for a linear function on an interval.* — Let Φ be the linear function on the interval $[\alpha,\beta]$ such that $\Phi(\alpha)=u$ and $\Phi(\beta)=v$. An easy computation of the corresponding functionals shows that

$$Z(\Phi) = (\beta-\alpha)\,S(-u,-v), \tag{63}$$

$$E(\Phi) = (\beta-\alpha)\,S(u,v), \tag{64}$$

$$M(\Phi) = \frac{(\beta-\alpha)^2}{v-u}\left(S(-u,-v)-e^{-v}\right) + \alpha\,(\beta-\alpha)\,S(-u,-v). \tag{65}$$

The function $S(u,v)$ is the slope of the secant line of the exponential function between the two points u and v :

$$S(u,v) \stackrel{\text{def.}}{=} \frac{e^u-e^v}{u-v}.$$

Remark that $S(-u,-v) = F_u(v) = F_v(u)$. An important property of this function for what follows is the differential equation

$$\partial_1 S(u,v) = \frac{1}{u-v}(e^u - S(u,v)) = \frac{1}{v-u}(S(u,v) - e^v) + S(u,v) \tag{66}$$

or, in other words,

$$(u-v)F_v'(u) = e^{-u} - F_v(u). \tag{67}$$

Hence, we can rewrite (65) to get

$$M(\Phi) = -(\beta-\alpha)^2\,\partial_1 S(-u,-v) + \beta\,(\beta-\alpha)\,S(-u,-v) \tag{68}$$

$$= (\beta-\alpha)^2\,\partial_1 S(-v,-u) + \alpha\,(\beta-\alpha)\,S(-u,-v). \tag{69}$$

This expression will be useful in section 2.3.5 to determine the order of the functional M_1.

2.3.3.2. *2-piecewise linear functions with fixed mass.* — We are interested here with the set of convex functions on $[0,1]$ with fixed values at 0 and 1 and fixed mass functional

$$\overline{\pi}_1(x, Z_1) = \{\Phi \,:\, \Phi(0) = 0,\, \Phi(1) = x,\, Z_1(\Phi) = Z_1\}, \quad Z_1 \geqslant F(x),$$

and with its subset of 2-piecewise linear functions $\overline{\pi}_1^2(x, Z_1)$. Recall that a 2-piecewise linear function $\Phi \in \pi_1(x)$ may be parametrized by its knot α and the value u at this point α. Then, the constraint

$$Z_1(\Phi) = \alpha\,F(u) + (1-\alpha)\,F_x(u) = Z_1$$

determines a curve. This curve itself may be globally parametrized by

$$\alpha = \alpha(u) = \frac{Z_1 - F_x(u)}{F(u) - F_x(u)}, \tag{70}$$

for u in an interval U. Precisely, $U = [u_0, u_1]$ if $x > 0$, $[u_1, u_0]$ if $x < 0$, where $u_0 = F_x^{-1}(Z_1)$ and $u_1 = F^{-1}(Z_1)$.

Remark that these 2-piecewise linear functions are either convex or concave. The convexity is automatically given by the necessary condition $Z_1 \geqslant F(x)$, which then make sure that

$$\alpha\,x \geqslant u. \tag{71}$$

The derivative

$$\alpha'(u) = -\frac{\frac{\partial Z_1}{\partial u}}{\frac{\partial Z_1}{\partial \alpha}} = -\frac{\alpha\,F'(u) + (1-\alpha)\,F_x'(u)}{F(u) - F_x(u)}$$

is strictly of the same sign as x, as F and F_x are strictly decreasing.

2.3.3.3. *Mean functional.* — A geometric argument points out the extremals of the mean functional M_1 on $\overline{\pi}_1(x, Z_1)$. The identical function $t \mapsto t$ being increasing, if the mass of Φ is fixed to $Z_1(\Phi) = Z_1$,

$$M_1(\Phi) = \int_0^1 t\, e^{-\Phi(t)}\, dt$$

is minimal as this mass is the most located near 0 as possible i.e. if Φ is the smaller as possible near the point 0. Hence, the minimizing function Φ is $\Phi_0 = \Phi_{0,u_0} \in \overline{\pi}_1^2(x, Z_1)$ with infinite slope at 0 i.e. the linear function joining $(0, u_0)$ to $(1, x)$. We may prove this result in a more rigorous way : if $\Phi \in \overline{\pi}_1^2(x, Z_1)$, there exists a β such that

$$\text{(i) } \Phi \geqslant \Phi_0 \text{ sur } [0, \beta]$$
$$\text{(ii) } \Phi \leqslant \Phi_0 \text{ sur } [\beta, 1].$$

Hence,

$$\begin{aligned}\int_0^1 t\left(e^{-\Phi} - e^{-\Phi_0}\right)(t)\, dt &\geqslant \int_0^\beta t\left(e^{-\Phi} - e^{-\Phi_0}\right)(t)\, dt + \beta \int_\beta^1 \left(e^{-\Phi} - e^{-\Phi_0}\right)(t)\, dt \\ &= \int_0^\beta (t-\beta)\left(e^{-\Phi} - e^{-\Phi_0}\right)(t)\, dt \geqslant 0.\end{aligned}$$

2.3.3.4. *Inverse mass functional.* — Thanks to the parametrization (70) of the set $\overline{\pi}_1^2(x, Z_1)$, we only have to study the variations of the function

$$E(u) = \alpha(u)\, F(-u) + (1 - \alpha(u))\, F_{-x}(-u). \tag{72}$$

Proposition 2.3.14. — *The function E is decreasing.*

Proof

◀ The presentation of the proof could certainly be improved and made clearer. The use of symmetries makes the computation tractable though at the expense of having to introduce quite complex notation. With further work, the above presentation might be reducible to a more simple and elegant form, which is not immediately apparent. This difficulty may be due to the notation used in the proof.

Recall that

$$\alpha(u) = \frac{Z_1 - F_x(u)}{F(u) - F_x(u)}$$

and that

$$\alpha'(u) = -\frac{\alpha\, F'(u) + (1-\alpha)\, F_x'(u)}{F(u) - F_x(u)}.$$

Consequently,

$$-E'(u) = \alpha\, A(u, x) + (1-\alpha)\, \tilde{A}(u, x) \tag{73}$$

where we put

$$A(u,x) = \frac{F(-u) - F_{-x}(-u)}{F(u) - F_x(u)}\, F'(u) + F'(-u)$$

and

$$\tilde{A}(u,x) = \frac{F(-u) - F_{-x}(-u)}{F(u) - F_x(u)}\, F'_x(u) + F'_{-x}(-u).$$

An interesting *symmetry* property is satisfied by the previous functions. That is

$$\tilde{A}(u,x) = e^x\, A(u-x,-x). \tag{74}$$

Indeed, coming back to the definition of F and F_x and making use of the differential equation (67), we get

$$A(u,x) = \frac{(u-x)\,(e^u-1) - u\,(e^u - e^x)}{(u-x)\,(1-e^{-u}) - u\,(e^{-x}-e^{-u})}\,\frac{1}{u}\left(e^{-u} - \frac{1-e^{-u}}{u}\right) - \frac{1}{u}\left(e^u - \frac{e^u-1}{u}\right)$$

and

$$\begin{aligned}\tilde{A}(u,x) = \frac{(u-x)\,(e^u-1) - u\,(e^u - e^x)}{(u-x)\,(1-e^{-u}) - u\,(e^{-x}-e^{-u})}\,\frac{1}{u-x}\left(e^{-u} - \frac{e^{-x}-e^{-u}}{u-x}\right)\\ - \frac{1}{u-x}\left(e^u - \frac{e^u-e^x}{u-x}\right).\end{aligned}$$

Which gives (74). Now, remark that $A(u,x)$ can be rewritten after a quite tedious development as

$$\begin{aligned}A(u,x) &= \frac{2}{u\,[(u-x)\,(1-e^{-u}) - u\,(e^{-x}-e^{-u})]}\,\psi_x(u)\\ &= \frac{2}{u^2\,(u-x)\,(S(-u,0) - S(-u,-x))}\,\psi_x(u),\end{aligned}$$

where the function $\psi_x(u)$ defined by

$$\psi_x(u) = x + (u-x)\,\cosh(u) - \sinh(u) + \sinh(x) - u\,\cosh(u-x) + \sinh(u-x)$$

has the same sign as x. And, thanks to (74), we get

$$\tilde{A}(u,x) = \frac{2\,e^x}{-(u-x)^2\,u\,(S(x-u,0) - S(x-u,x))}\,\psi_{-x}(u-x).$$

Well,

$$\psi_{-x}(u-x) = -\psi_x(u)$$

and

$$e^{-x}(S(x-u,0) - S(x-u,x)) = S(-u,-x) - S(-u,0).$$

So we have

$$\begin{aligned}-E'(u) &= \frac{2\,\psi_x(u)}{(S(-u,0)-S(-u,-x))}\left(\frac{\alpha}{u^2\,(x-u)}-\frac{1-\alpha}{u\,(x-u)^2}\right)\\ &= \frac{2\,\psi_x(u)}{u^2\,(x-u)^2\,(S(-u,0)-S(-u,-x))}\,(\alpha\,x-u).\end{aligned}$$

But $u \leqslant \alpha\,x$ is exactly the convexity constraint (71) of the function. Finally, as ψ_x has the same sign as x, the ratio

$$\frac{\psi_x(u)}{S(-u,0)-S(-u,-x)}$$

is always non negative. ▶

As a consequence, the maximun value of the function E is achieved at point

$$\min(U) = \begin{cases} u_0 & \text{if } x>0 \\ u_1 & \text{if } x<0. \end{cases}.$$

The corresponding extremal function is then

(i) Φ_0 if $x>0$,

(ii) $\Phi_1 = \overline{u_1}$ if $x<0$.

Here, $\overline{u_1}$ is, as in definition 2.3.7, the linear function on $[0,1]$ such that $\overline{u_1}(0)=0$ and $\overline{u_1}(1)=u_1$.

2.3.4. Minimizing functions of the variational problem. — We are going to prove that, in problem 2, it suffices to consider 2-piecewise linear functions on $[0,1]$. Let Φ be in $\pi_1(x,b,Z_1,Z_2)$.

2.3.4.1. *Approximation by piecewise linear functions.* — The first step consists in an approximation of Φ by a sequence $(\Phi_n)_n \subset \pi_1(x,b,Z_1,Z_2)$ of piecewise linear functions which converges uniformly to Φ. We get such an approximation by the following argument : given a partition σ of $[0,1]$, let the functions

$$\overline{\Phi}(\sigma) \geqslant \Phi \geqslant \underline{\Phi}(\sigma),$$

which are 1 or 2-piecewise linear between two points of the partition, be defined as follows. $\overline{\Phi}(\sigma)$ is the secant line of Φ between these two points, $\underline{\Phi}(\sigma)$ is the supremum of the two tangent lines of Φ in these points. Easily, these two functions converge uniformly to Φ as the partition size goes to 0. Now, on each interval between two consecutive points of the partition, the situation is similar to figure 2. We are hence able to exhibit a 2-piecewise linear function on the considered interval with the same mass as Φ. Globally, we get a piecewise linear $\tilde{\Phi}$ in $\pi_1(x,b,Z_1,Z_2)$ such

$$\left\|\tilde{\Phi}-\Phi\right\|_\infty \leqslant \max\left(\left\|\overline{\Phi}-\Phi\right\|_\infty, \|\underline{\Phi}-\Phi\|_\infty\right).$$

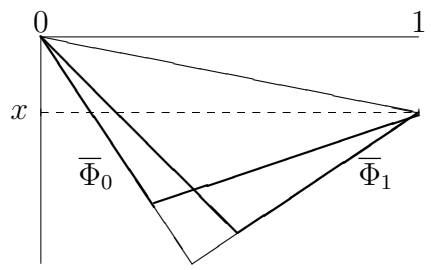


FIGURE 3.

As the partition size goes to 0, we get the expected sequence $(\Phi_n)_n$. As $K_M(\Phi_n)$ converges to $K_M(\Phi)$ as $n \to +\infty$, we may assume in what follows that the considered function $\Phi \in \pi_1(x, b, Z_1, Z_2)$ is piecewise linear.

2.3.4.2. *Extremal functions of the mean and inverse mass functionals.* — On the set $\pi_1(x, b, Z_1, Z_2)$, we have to replace the functions Φ_0 and Φ_1 introduced before by functions $\overline{\Phi}_0$ and $\overline{\Phi}_1$ satisfying the corresponding constraints. The latter functions are whose with mass Z_1 the most located as possible near 0 and 1 respectively (see figure 3).

Thanks to the preceding paragraph and the study of the mean and inverse mass functionals on 2-piecewise linear functions in the subsection 2.3.3, we will be able to find the extremal functions of these functionals on the whole set $\pi_1(x, b, Z_1, Z_2)$.

Lemma 2.3.15. — *For any function $\Phi \in \pi_1(x, b, Z_1, Z_2)$,*

$$M_1\big(\overline{\Phi}_0\big) \leqslant M_1(\Phi) \leqslant M_1\big(\overline{\Phi}_1\big).$$

Moreover,

$$E_1\big(\overline{\Phi}_1\big) \leqslant E_1(\Phi) \leqslant E_1\big(\overline{\Phi}_0\big), \quad \textit{if} \;\; x \geqslant 0,$$
$$E_1\big(\overline{\Phi}_0\big) \leqslant E_1(\Phi) \leqslant E_1\big(\overline{\Phi}_1\big), \quad \textit{if} \;\; -\log(2) < x \leqslant 0.$$

Remark 2.3.16. — *For $x > 0$, M_1 achieves its minimum value as E_1 achieves its maximum, so $\overline{\Phi}_0$ is the minimizing function of $K_M(\Phi)$ on $\pi_1(x, b, Z_1, Z_2)$. For $x < 0$, a closer look is necessary.*

Proof of lemma 2.3.15. — As seen before, we may assume Φ to be n-piecewise linear for some n. Let $I = [0, \gamma]$ be an interval on which Φ is 3-piecewise linear. We now describe some $(n-1)$-piecewise linear functions $\overleftarrow{\Phi}_{n-1}$ and $\overrightarrow{\Phi}_{n-1}$ in $\pi_1(x, b, Z_1, Z_2)$ which coincide with Φ outside I. On I, $\overleftarrow{\Phi}_{n-1}$ (resp. $\overrightarrow{\Phi}_{n-1}$) is the 2-piecewise linear function with the same mass $Z_1(\Phi)$ as Φ construct by extension of the slope of Φ at its first (resp. second) knot α (resp. β). We concentrate on the comparison of Φ and $\overleftarrow{\Phi}_{n-1}$. They coincide on $[\alpha, \gamma]$ and their restrictions to this interval are 2-piecewise linear functions. We are then allowed to make use of the results seen in subsection

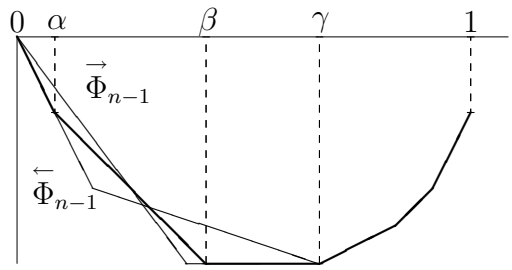


FIGURE 4.

2.3.3.3 and 2.3.3.4 to compare the values of the mean and inverse functionals on these two functions. The same things occurs for Φ and $\overrightarrow{\Phi}_{n-1}$ on $[0,\beta]$. On one side,

$$M_1\Big(\overleftarrow{\Phi}_{n-1}\Big) \leqslant M_1(\Phi) \leqslant M_1\Big(\overrightarrow{\Phi}_{n-1}\Big). \tag{75}$$

On the other side, according to the respective values of α, β and γ,

$$E_1\Big(\overleftarrow{\Phi}_{n-1}\Big) \leqslant E_1(\Phi) \leqslant E_1\Big(\overrightarrow{\Phi}_{n-1}\Big), \tag{76}$$

or

$$E_1\Big(\overrightarrow{\Phi}_{n-1}\Big) \leqslant E_1(\Phi) \leqslant E_1\Big(\overleftarrow{\Phi}_{n-1}\Big). \tag{77}$$

By induction, we end up with 2-piecewise linear functions $\overleftarrow{\Phi}_2$ and $\overrightarrow{\Phi}_2$ on $[0,1]$ which satisfy inequalities (75) and (76) or (77). The proof is complete if you remember that, for any 2-piecewise linear $\tilde{\Phi} \in \pi_1(x,b,Z_1,Z_2)$ on $[0,1]$, the expected comparisons of lemma 2.3.15 derive easily from the results proved in subsection 2.3.3. □

2.3.4.3. *Restriction to 2-piecewise linear functions.* — To prove proposition 2.3.4, it remains to show that we may only consider 2-piecewise linear functions on $[0,1]$ to minimize the functional K of problem 2. According to remark 2.3.16, we may assume that $-\log(2) < x < 0$. The argument is based on lemma 2.3.15. Let $\Phi \in \pi_1(x,b,Z_1,Z_2)$. Note $M = M_2(x,b,Z_2)$. Then

$$K_M(\Phi) = \frac{(M_1(\Phi)+M)^2}{E_1(\Phi)} \geqslant \frac{\Big(M_1(\tilde{\Phi}_0)+M\Big)^2}{E_1(\tilde{\Phi}_1)} \tag{78}$$

$$= \frac{E_1(\tilde{\Phi}_0)}{E_1(\tilde{\Phi}_1)}\, K_M\Big(\tilde{\Phi}_0\Big) \tag{79}$$

$$\geqslant \frac{E_1(\Phi_0)}{E_1(\Phi_1)}\, K_M\Big(\tilde{\Phi}_0\Big). \tag{80}$$

We now have to bound the last ratio from below, uniformly in the parameters x and Z_1.

Lemma 2.3.17. — *For any* $-\log(2) < x < 0$ *and* $Z_1 \geqslant 0$,

$$R(x, Z_1) \stackrel{def.}{=} \frac{E_1(\Phi_0)}{E_1(\Phi_1)} \geqslant 1/2.$$

Proof
◀ According to (72),

$$R(x, Z_1) = \frac{E_1(u_0)}{E_1(u_1)} = \frac{F_{-x}(-F_x^{-1}(Z_1))}{F(-F^{-1}(Z_1))}.$$

Now, $F_x(u) = e^{-x}\, F(u - x)$ and $F_x^{-1}(z) = x + F^{-1}(e^x\, z)$. Hence,

$$R(x, Z_1) = e^x\, \frac{F(-F^{-1}(e^x\, Z_1))}{F(-F^{-1}(Z_1))}.$$

But, the function $G(z) = F(-F^{-1}(z))$ is decreasing as

$$G'(z) = -\frac{F'(-F^{-1}(z))}{F'(F^{-1}(z))} \leqslant 0,$$

and then $R(x, Z_1) \geqslant e^x \geqslant 1/2$. ▶

Finally, for any $\Phi \in \pi_1(x, b, Z_1, Z_2)$,

$$K_M(\Phi) \geqslant 1/2\, K_M\Big(\tilde{\Phi}_0\Big).$$

Remark 2.3.18. — *It would be more interesting to minimize the ratio*

$$\frac{M_1(\Phi_0)}{M_1(\Phi_1)} \tag{81}$$

if it were possible. We would get

$$\begin{aligned} K_M(\Phi) &\geqslant \frac{\Big(M_1(\tilde{\Phi}_0) + M\Big)^2}{E_1(\tilde{\Phi}_1)} \\ &= \left(\frac{M_1(\tilde{\Phi}_0) + M}{M_1(\tilde{\Phi}_1) + M}\right)^2 K_M\Big(\tilde{\Phi}_1\Big) \\ &\geqslant \left(\frac{M_1(\Phi_0)}{M_1(\Phi_1)}\right)^2 K_M\Big(\tilde{\Phi}_1\Big). \end{aligned}$$

and the slope of function $\tilde{\Phi}_1$ *at point* 1 *is* b *as for the extremal function outside* $[0, 1]$. *The extremal functions would be* 2*-piecewise linear on a compact interval. Unfortunately, the ratio* (81) *goes to* 0 *as the fixed mass goes to infinity. Namely,*

the mean functional is related to the mass near 0. *So, the mean of* Φ_0 *is greater as* Z_1 *increases w.r.t.* Φ_1*'s. Precisely, if* $x = 0$ *(and then* $u_0 = u_1$*), we have*

$$\frac{M_1(\Phi_0)}{M_1(\Phi_1)} = \frac{1 - F(u_1)}{F(u_1) - e^{-u_1}} = \frac{u - 1 + e^{-u}}{1 - e^{-u} - u\, e^{-u}}$$

which is equivalent to $-1/u$ *as* u *goes to* $-\infty$.

2.3.5. Return to the initial problem for 3-piecewise linear functions. — In proposition 2.3.4, we introduced the set of 3-piecewise linear extremal functions $\mathcal{T}$. A function in $\mathcal{T}$ is given by the five following parameters : the knot α inside $]0, 1[$, the value u of the function at this point, the value x and the slope r at point 1_+ and the threshold β after which the function is infinite (see figure 5). We will note p and q the slopes at 0 and 1_-. The convexity constraints are

$$p = \frac{u}{\alpha} \leqslant x \quad \text{and} \quad q = \frac{x - u}{1 - \alpha} \leqslant r. \tag{82}$$

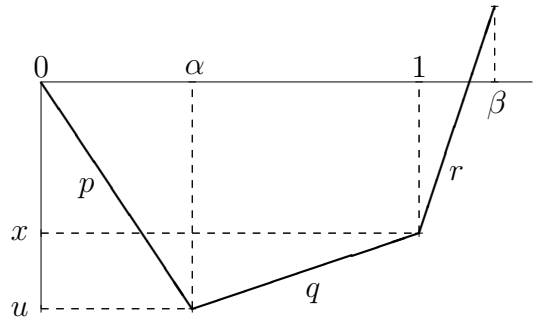


FIGURE 5.

Remark the value at β is $w = x + r\,(\beta - 1)$. So, we derive from paragraph 2.3.3.1 the explicit values of our functionals.

$$\begin{aligned}
M_1 &= \alpha^2\, \frac{1 - (u+1)\, e^{-u}}{u^2} + (1-\alpha)^2\, \frac{F_x(u) - e^{-x}}{x - u} + \alpha\,(1-\alpha)\, F_x(u),\\
M_2 &= \frac{1}{r^2}\, e^{-x}\, \big((r+1) - (\beta\, r + 1)\, e^{-r(\beta-1)}\big),\\
Z_1 &= \alpha\, F(u) + (1-\alpha)\, F_x(u),\\
Z_2 &= \frac{1}{r}\, e^{-x}\, \big(1 - e^{-r(\beta-1)}\big),\\
E_1 &= \alpha\, F(-u) + (1-\alpha)\, F_{-x}(-u),\\
K &= \frac{1}{E_1\, Z_2} \left(\frac{M_1 + M_2}{Z_1 + Z_2}\right)^2.
\end{aligned}$$

The condition of lemma 2.3.2, which makes sure that Φ has a convex extension with median 0, is

$$p\,Z = -u\left[F_x(u) - F(u) - \frac{F_x(u) + Z_2}{\alpha}\right] \geqslant -1. \tag{83}$$

We have to prove that K is bounded from below by a positive constant on $\mathcal{T}$ i.e. on the domain in $\mathbb{R}^5$ determined by the constraints (82) and (83).

2.3.5.1. *Proof of lemma 2.1.4.* — Let Φ be a convex function on $\mathbb{R}$ going to infinity at infinity. Recall we want to establish the estimate

$$\Phi \geqslant \Phi(m) - \log(2),$$

where m is the median of Φ. We may assume that $m = 0$ and that $\Phi(0) = 0$. Then the restriction of Φ on $[0, \infty[$ satisfies the extension condition of lemma 2.3.2 and, for any $t > 0$, the convex function

$$l(s) = \frac{\Phi(t)}{t}\, s\, \mathbb{1}_{[0,t]}(s) + (+\infty)\, \mathbb{1}_{]t,+\infty[}(s)$$

is still in class $\mathcal{C}$ as $l'(0) \geqslant \Phi'(0)$ and $Z(l) \leqslant Z(\Phi)$. Consequently,

$$l'(0)\, Z(l) = 1 - e^{-\Phi(t)} \geqslant -1$$

i.e.

$$\Phi(t) \geqslant -\log(2).$$

We get the lower bound for the restriction of Φ on $]-\infty, 0]$ by symmetry.

2.3.5.2. *Existence conditions.* — By analogy with the previous notations, we note

$$\pi_1(x, r, \beta) = \{\Phi \in \mathcal{T} \text{ defined on } [0, \beta] : \Phi(1) = x,\ \Phi'(1_+) = r\}.$$

The following proposition describes under which conditions there exists a function in $\mathcal{T}$ with fixed parameters x, r, β.

Proposition 2.3.19. — *Let $\beta \geqslant 1$ and $x \leqslant r$. If $x \geqslant 0$, $\pi_1(x, r, \beta) \neq \emptyset$. Assume now that $x < 0$ and note*

$$\epsilon(x) \overset{def.}{=} \frac{1 - 2\,e^x}{x} = e^x\, \eta(x).$$

Then $\pi_1(x, r, \beta)$ is not empty, if and only if, $-\log(2) < x < 0$ and one of the two following conditions is true

(i) $r \geqslant 1/\epsilon(x)$,

(ii) $\beta \leqslant 1 - \frac{1}{r}\log\left(1 - r\,\epsilon(x)\right)$.

Proof

◀ It is similar to that of lemma 2.3.8 and corollary 2.3.9 we omitted previously. If $\Phi \in \pi_1(x, r, \beta)$, then

$$\overline{x}\, \mathbb{1}_{[0,1]} + \Phi\, \mathbb{1}_{]1,\infty[} \in \pi_1(x, r, \beta),$$

and

$$x \geqslant \frac{-1}{F(x) + Z_2(x, r, \beta)},$$

for $Z_2(x, r, \beta) = e^{-x}\left(1 - e^{-r\,(\beta-1)}\right)/r$. It is exactly

$$\left(1 - e^{-r\,(\beta-1)}\right)/r \leqslant \epsilon(x).$$

And we get the expected conditions. ►

2.3.5.3. *The case of bounded x.* — Let $A \gg 1$ be a fixed constant and let $\Phi \in \mathcal{T}$ be such that the value x at 1 is less than A. Following lemma 2.1.4, the functions Z_1, M_1, E_1 et e^{-x} are of the same order as the constant function 1 and we only have to study the simpler function

$$H(r, \beta) = \frac{1}{\sqrt{Z_2}}\,\frac{1 + M_2}{1 + Z_2}, \tag{84}$$

where x, r, β satisfy the constraints of proposition 2.3.19. In (84), we wrote Z_2 and M_2 for $e^x\,Z_2$ and $e^x\,M_2$ which do not depend on x any more.

Well, the function $\epsilon(x)$ being increasing, if $-\log(2) < x \leqslant r < 0$,

$$f_r(x) = 1 - \frac{1}{r}\log\left(1 - r\,\epsilon(x)\right)$$

is increasing as well. Hence, the weakest constraint $\beta \leqslant f_r(x)$ is obtained as $x = r$. And we get the following domain

$$-\log(2) \leqslant r \text{ et } 1 \leqslant \beta \leqslant \beta_{\max}(r) \tag{85}$$

where

$$\beta_{\max}(r) = \begin{cases} -\log(2)/r & \text{si } -\log(2) \leqslant r < 0 \\ +\infty & \text{si } r \geqslant 0, \end{cases} \tag{86}$$

on which we have to minimize $H(r, \beta)$.

A variation computation. —

Lemma 2.3.20. — *The function $H(r, \beta)$ has no extremum value inside the domain* (85).

Proof

◄ Write M_2 and Z_2 in terms of parameters r and $w = r\,(\beta - 1)$. We get

$$Z_2 = \frac{1}{r}\left(1 - e^{-w}\right) \tag{87}$$

and

$$M_2 = \left(1 + \frac{1}{r}\right) Z_2 - \frac{w\,e^{-w}}{r^2}. \tag{88}$$

Assume on the contrary that there is an extremum inside the domain. The EULER-LAGRANGE equations for the extremum of H, as Z_2 is fixed, show that

$$(w+2)\,e^{-w}+w-2=0$$

i.e. $w=0$ or rather $\beta=1$ which is not inside the domain. ►

What about the boundary of the domain ?— We are going to prove in the following that function H is positively bounded from below on the boundary of domain (85).

First, $H(r,1)\equiv+\infty$ and so does $H(+\infty,\beta)$. The case $r=-\log(2)$ is a particular case of $\beta=1$.

It remains to have a look at the case $\beta=\beta_{\max}(r)$.

If $-\log(2)\leqslant r<0$,

$$Z_2=\frac{1}{r}\,(1-2\,e^r)=\epsilon(r),$$
$$M_2=\frac{1}{r^2}((r+1)-2\,(1-\log(2))\,e^r).$$

Put $v=\log(2)+r$. We get

$$Z_2\approx\frac{v}{\log(2)-v}$$

and

$$M_2\approx\frac{v}{(\log(2)-v)^2}.$$

So,

$$\frac{1}{\sqrt{Z_2}}\,\frac{1+M_2}{1+Z_2}$$

goes to $+\infty$ as v goes to 0 or $\log(2)$.

Now, if $r\geqslant 0$, we have to study the function $H(r,+\infty)$. Then we have

$$Z_2(r,+\infty)=\frac{1}{r}\text{ et }M_2(r,\infty)=\frac{r+1}{r^2}$$

Consequently,

$$H(r,\infty)\approx\frac{1+r}{\sqrt{r}}$$

is positively bounded from below as well.

2.3.5.4. *The case of x going to infinity.* —

The case $p<0$. — We concentrate here on the case when the derivative p at 0 is negative. Note $v=-u>0$. Fix Z_2 which is less than e^{-x}/x by the convexity of the function. We divide the domain D introduced in (83) into curves D_γ, $0<\gamma\leqslant 1$, defined by the following equation

$$v\left[F_x(-v)-F(-v)-\frac{F_x(-v)+Z_2}{\alpha}\right]=-\gamma. \tag{89}$$

As the left hand side of the preceding equation is maximal for $\alpha = 1$, there exists an α satisfying (89) if and only if

$$v\,(-Z_2 - F(-v)) \geqslant -\gamma$$

or equivalently if

$$0 < v \leqslant \psi_\gamma(Z_2) \leqslant \log(2). \tag{90}$$

Here, we put $\psi_\gamma = -\eta_\gamma^{-1}$ where the strictly increasing function η_γ is defined as follows

$$\eta_\gamma(u) \stackrel{\text{def.}}{=} \frac{e^{-u} - (\gamma + 1)}{u}.$$

Now, we get the parametrization of the curve D_γ

$$\forall 0 < v \leqslant \psi_\gamma(Z_2), \qquad \alpha = \alpha(v, Z_2) = \frac{F_x(-v) + Z_2}{F_x(-v) + z}$$

with $z = \eta_\gamma(-v) \geqslant Z_2$. We have

$$F_x(-v) = S(v, -x) = \frac{e^v - e^{-x}}{x + v}$$

with $x + v \approx x$ and $e^v - e^{-x} \approx 1 - e^{-x} \approx 1$, hence

$$F_x(-v) \approx 1/x.$$

It follows that

$$\alpha \approx \frac{1 + x\,Z_2}{1 + x\,z}.$$

Now, $Z_2 \leqslant e^{-x}/x$, and then $1 + x\,Z_2 \approx 1$. Finally,

$$\alpha \approx \frac{1}{1 + z\,x} \approx \frac{1}{1 + \lambda\,x} \tag{91}$$

if we introduce the new parameter

$$\lambda = z - Z_2 \geqslant 0.$$

Remark that $F(-v) \approx 1 \approx F(v)$ as these functions admit positive limits at points 0 and $\log(2)$. So,

$$Z_1 \approx \alpha + \frac{1 - \alpha}{x} = \alpha(1 - \frac{1}{x}) + \frac{1}{x} \approx \alpha + \frac{1}{x} \approx \frac{1}{1 + \lambda\,x} + \frac{1}{x}.$$

But

$$\frac{1}{1 + \lambda\,x} + \frac{1}{x} = \frac{1 + (\lambda + 1)\,x}{x\,(1 + \lambda\,x)} \approx \frac{\lambda + 1}{1 + \lambda\,x},$$

as $(\lambda + 1)\,x \geqslant x$ and then $1 + (\lambda + 1)\,x \approx (\lambda + 1)\,x$.

Moreover, $Z_2 = \circ(1/x) = \circ(Z_1)$, and finally,

$$Z \approx Z_1 \approx \frac{1 + \lambda}{1 + \lambda\,x}. \tag{92}$$

Now,

$$E_1 = \alpha F(v) + (1-\alpha) F_{-x}(v),$$

with

$$F_{-x}(v) = \frac{e^x - e^v}{x+v} \approx \frac{e^x}{x}.$$

As

$$1-\alpha = \frac{z - Z_2}{z + F_x(-v)} \approx \frac{x\,(z - Z_2)}{1 + x\,z} \approx \frac{\lambda\,x}{1+\lambda\,x}, \tag{93}$$

and we get

$$E_1 \approx \frac{1+\lambda\,e^x}{1+\lambda x}. \tag{94}$$

Let's now concentrate on M_1. Thanks to equation (65), we get

$$M_1 = \alpha^2\,\frac{1+(v-1)\,e^v}{v^2} + \frac{(1-\alpha)^2}{x+v}\left(S(v,-x) - e^{-x}\right) + \alpha\,(1-\alpha)\,S(v,-x). \tag{95}$$

As x and v are positive, the slope of the tangent of the exponential function at point $-x$ is less than the slope of the secant between $-x$ and v, i.e.

$$S(v,-x) \geqslant e^{-x}.$$

Hence, each term of the sum (95) is positive, and it suffices to find the order of each of them separately. First,

$$\frac{1+(v-1)\,e^v}{v^2} \approx 1.$$

Then,

$$S(v,-x) \approx \frac{1-e^{-x}}{x} \approx \frac{1}{x}$$

and $e^{-x} = \circ(1/x)$, so that

$$S(v,-x) - e^{-x} \approx \frac{1}{x}.$$

We get

$$\begin{aligned} M_1 &\approx \alpha^2 + \frac{(1-\alpha)^2}{x^2} + \frac{\alpha\,(1-\alpha)}{x} \\ &\approx \frac{1}{(1+\lambda\,x)^2} + \frac{1}{x^2}\left(\frac{\lambda\,x}{1+\lambda\,x}\right)^2 + \frac{1}{x}\,\frac{\lambda\,x}{(1+\lambda\,x)^2} \\ &\approx \frac{\lambda^2+\lambda+1}{1+\lambda^2\,x^2} \end{aligned}$$

or equivalently

$$M_1 \approx \frac{1+\lambda^2}{1+\lambda^2\,x^2}. \tag{96}$$

We are now able to bound the functional K on D_γ from below by a positive constant in the case $x \gg 1$.

Recall that $\succcurlyeq$ stands for $\geqslant$ up to a universal constant. From (92) and (96), it follows that $(M_1/Z)^2 \approx M_1$. Moreover, as $Z_2 \leqslant e^{-x}/x$, we have

$$K = \frac{1}{E_1\, Z_2}\left(\frac{M_1+M_2}{Z}\right)^2 \geqslant \frac{x\, e^x}{E_1}\left(\frac{M_1}{Z}\right)^2 \succcurlyeq x\, e^x\, \frac{1+\lambda x}{1+\lambda\, e^x}\, \frac{1+\lambda^2}{1+\lambda^2\, x^2}.$$

It remains to minimize the function $H(\lambda, x)$ uniformly on the parameters $\lambda \geqslant 0$ and $x \geqslant A$ for some value of the threshold A. But,

$$\begin{aligned} x\, e^x\, \frac{1+\lambda x}{1+\lambda\, e^x}\, \frac{1+\lambda^2}{1+\lambda^2\, x^2} &\approx x\, \frac{1}{e^{-x}+\lambda}\, \frac{1+\lambda^2}{1+\lambda\, x} \succcurlyeq x\, \frac{1}{1+\lambda}\, \frac{1+\lambda^2}{1+\lambda\, x} \\ &\succcurlyeq x\, \frac{1+\lambda}{1+\lambda\, x} = \frac{x+\lambda\, x}{1+\lambda\, x} \geqslant 1. \end{aligned}$$

The parameter γ was hidden up to now behind $z = \eta_\gamma(-v) \geqslant Z_2$ or $\lambda = z - Z_2$. Hence, the positive bound we get previously does not depend on γ and is valid on the whole domain D.

The case $p \geqslant 0$. — We concentrate in this paragraph on functions with non negative derivative at point 0. In that case,

$$0 \leqslant u \leqslant \alpha\, x$$

and

$$F(x) = Z_1(\overline{x}) \leqslant Z_1 \leqslant 1 = Z_1(0). \tag{97}$$

Hence $Z_1 \succcurlyeq \frac{1}{x}$ so that $Z_2 = \circ(Z_1)$. It follows

$$K \succcurlyeq \frac{x\, e^x}{E_1}\left(\frac{M_1}{Z_1}\right)^2. \tag{98}$$

Now, remember the results we got in paragraph 2.3.3. Recall that a 2-piecewise linear function on $[0,1]$ which takes value x at 1 is parametrized by the knot α and the value u at this point. When the mass Z_1 is fixed, we get a curve

$$\alpha = \alpha(u) = \frac{Z_1 - F_x(u)}{F(u) - F_x(u)},$$

where u lies in $U = [u_0, u_1]$ (as $x \geqslant 0$) with $u_0 = F_x^{-1}(Z_1)$ and $u_1 = F^{-1}(Z_1)$. According to (97), $u_0 \leqslant 0 \leqslant u_1$, so that $0 \in U$. We had proven (still because $x \geqslant 0$) that $E_1 = E_1(u)$ is decreasing and $M_1 = M_1(u)$ increasing. Consequently, $\frac{(M_1)^2}{E_1}$ is an increasing function of parameter u which achieves its minimum (when the slope p is non negative at 0) at $u = 0$. And we may restrict ourselves to the case $u = 0$ in (98). Let's have a look at the remaining function

$$K(\alpha, x) = \frac{x\, e^x}{E_1}\left(\frac{M_1}{Z_1}\right)^2$$

where

$$\begin{aligned}
E_1 &= \alpha + (1-\alpha)\,\frac{e^x - 1}{x} \approx \alpha + (1-\alpha)\,\frac{e^x}{x},\\
Z_1 &= \alpha + (1-\alpha)\,\frac{1-e^{-x}}{x} \approx \alpha + \frac{1-\alpha}{x},\\
M_1 &= \frac{\alpha^2}{2} + \frac{(1-\alpha)^2}{x}\left(\frac{1-e^{-x}}{x} - e^{-x}\right) + \alpha\,(1-\alpha)\,\frac{1-e^{-x}}{x}\\
&\approx \alpha^2 + \frac{(1-\alpha)^2}{x^2} + \frac{\alpha\,(1-\alpha)}{x}.
\end{aligned}$$

Finally,

$$\begin{aligned}
K(\alpha, x) &\approx \frac{1}{\alpha\, x\, e^{-x} + (1-\alpha)}\left(\frac{\alpha^2\, x^2 + (1-\alpha)^2 + \alpha\,(1-\alpha)\, x}{\alpha\, x + 1}\right)^2\\
&\succcurlyeq \frac{1}{\alpha + (1-\alpha)}\left(\frac{(1+\alpha\, x)^2}{1+\alpha\, x}\right)^2 \succcurlyeq (1+\alpha\, x)^2 \succcurlyeq 1.
\end{aligned}$$

The proof of proposition 2.3.5 is complete.

2.4. Cheeger inequality and Bobkov's argument

In this section, we present a new proof of theorem 2.1.1 thanks to a closer analysis of Bobkov's argument (see [Bob99]). This approach gives rise to an exact estimate of the constant D, which may be chosen equal to 16. We start with some generalities on Cheeger inequality and then come to the proof of the result.

2.4.1. Preliminaries on Cheeger inequality. — Let μ be a probability measure on Euclidean space $\mathbb{R}^n$. For any measurable set A, the surface measure of A is

$$\mu_s(\partial A) = \liminf_{\varepsilon \to 0} \frac{\mu(A^\varepsilon) - \mu(A)}{\varepsilon},$$

where $A^\varepsilon = \{x \in \mathbb{R}^n \ : \ |x - a| < \varepsilon, \text{ for some } a \in A\}$ is called the ε-neighborhood of A. The measure μ satisfies a Cheeger inequality if there exists a constant $C_1 > 0$ such that, for any measurable set,

$$\min\left(\mu(A), \mu(A^c)\right) \leqslant C_1\, \mu_s(\partial A). \tag{99}$$

This inequality is nothing but a uniform control of the measure of any Borel set such that $\mu(A) \leqslant 1/2$ by its surface measure. It was introduced by Cheeger to get an estimate of the spectral gap of the Laplace-Beltrami operator on a Riemannian manifold (see [Che70]). As the Gaussian isoperimetric inequality, Cheeger

inequality has a functional equivalent version. This functional CHEEGER inequality says that, for any compactly supported smooth function f on $\mathbb{R}^n$ with median m,

$$\int |f-m|\,d\mu \leqslant C_1 \int |\nabla f|\,d\mu. \tag{100}$$

The median m is precisely the median of the law of f under measure μ. This formulation shows that CHEEGER inequality is a $\mathbf{L}^1$ or isoperimetric version of POINCARÉ inequality. This is the reason of the constants' indices 1 and 2.

Formally, we get the equivalence between (99) and (100) as follows. First, if $\mu(A) \leqslant 1/2$ then 0 is a median of $\mathbb{1}_A$, so that (100) applied to such indicator functions is exactly (99). Conversely, if g is smooth enough with median 0, using co-area formula on the manifold $\{g>0\}$, we get

$$\int_{\{g>0\}} g\,d\mu = \int_0^{+\infty} \mu(\{g>t\})\,dt \leqslant C_1 \int_0^{+\infty} \mu_s(\{g=t\})\,dt = C_1 \int_{\{g>0\}} |\nabla g|\,d\mu,$$

as $\mu(g>t) \leqslant 1/2$. Then, (100) follows by applying the previous inequality to both $f-m$ and $m-f$.

2.4.1.1. *Isoperimetric function for measures on the real line.* — Assume

$$\mu(dx) = Z^{-1} \exp(-\Phi(x))\,dx$$

is a BOLTZMANN probability measure on $\mathbb{R}$ with phase function Φ continuous. Following S.BOBKOV, note $F(t) = \mu(]-\infty,t])$ the distribution function of μ (which is continuously differentiable and strictly increasing) and define the *isoperimetric function* of μ by

$$\iota_\mu(p) = F' \circ F^{-1}(p), \quad p \in [0,1].$$

For the exponential measure $\epsilon(dx) = \exp(-|x|)\,dx/2$, the function $\iota(p)$ is the function $\min(p, 1-p)$ appearing in CHEEGER inequality, whereas for the Gaussian distribution it is the function $\mathcal{U}(p)$ of the Gaussian isoperimetric inequality (see chapter 1).

Why this terminology *isoperimetric function*? Trivially, if H is a half-line, we have $\mu_s(\partial H) = \iota_\mu(\mu(H))$. Rigorously, the isoperimetric function of the measure μ is

$$I_\mu(p) = \inf_{\mu(A)=p} \mu_s(\partial A).$$

By definition, I_μ is the greatest function I such that, for any BOREL set A,

$$I(\mu(A)) \leqslant \mu_s(\partial A). \tag{101}$$

So, requiring ι_μ equals to I_μ is exactly saying that inequality (101) for $I = I_\mu$ has extremal sets with any fixed measure p and that these sets are half-lines. According to [Bob96b] (see also chapter 1 and [Bob96a]), it is the case when μ is the Gaussian measure. It is known that the same thing occurs for the exponential measure. Hence, the isoperimetric inequality of $\epsilon(dx)$ is exactly

$$\min(\epsilon(A), 1-\epsilon(A)) \leqslant \epsilon_s(\partial A)$$

with half-lines as extremal sets (see [Tal91]). So the best constant in CHEEGER inequality for $\epsilon(dx)$ is 1.

Another consequence is that CHEEGER inequality appears like a comparison with the typical model of the exponential measure. This situation is very similar to the one of Gaussian measure with respect to BOBKOV's isoperimetric inequality. In [BL96a], BAKRY and LEDOUX deduce from this Gaussian isoperimetric inequality of constant C_1 (of $\mathbf{L}^1$ type) the corresponding $\mathbf{L}^2$ type inequality, i.e. logarithmic SOBOLEV inequality of constant $C_2 = 2\,C_1^2$. The constant 2 comes from the logarithmic SOBOLEV constant of the Gaussian model. The same argument may be applied to CHEEGER and POINCARÉ inequalities. Whereas this result is well known, we now sketch the copied proof in this new context in order to emphasize the unity of methods to compare inequalities of $\mathbf{L}^1$ and $\mathbf{L}^2$ types (in the case of diffusion operators).

2.4.1.2. CHEEGER *and* POINCARÉ *inequalities.* — It is easy to be convinced the POINCARÉ's constant of the exponential measure $\epsilon(dx)$ is $C_2(\epsilon) = 4$. Trivially, the MUCKENHOUPT functional takes value 1 in this example. Hence, by (50), $C_2(\epsilon) \leqslant 4$. Applying POINCARÉ inequality to the monomials x^k for odd k, we get $C_2(\epsilon) \geqslant 4$ as k goes to infinity. Remark that it shows the constant 4 in MUCKENHOUPT estimate is optimal.

The following argument was already detailed in the proof of lemma 1.2.7 and comes from [BL96a]. In this paragraph, μ is a probability measure on $\mathbb{R}^n$. Let f be a fixed smooth function. We may assume that the law of f has a positive density G'. Remark that, in the case $r \leqslant m$, 0 is a median of $g \overset{\text{def.}}{=} \mathbb{1}_{]-\infty,r]} \circ f$. In the other case $r \geqslant m$, 1 is a median. Applying the functional CHEEGER inequality to g, we get the differential inequality

$$\min\left(G(r), 1 - G(r)\right) \leqslant C_1 \theta(r) G'(r)$$

where G is the distribution function of the law of f under μ and

$$\theta(r) = \mathbb{E}_\mu(\sqrt{\mathbf{\Gamma}(f)}/f = r)$$

is a version of the conditional expectation of $\sqrt{\mathbf{\Gamma}(f)}$ given f. Note F_ϵ the distribution function of $\epsilon(dx)$. In terms of the function $k = G^{-1} \circ F_\epsilon$ which has the same law under ϵ as f under μ, the previous inequality says exactly that

$$k' \leqslant C_1\, \theta \circ k. \tag{102}$$

The POINCARÉ inequality for ϵ applied to k, conjugate with (102), gives rise to the POINCARÉ inequality

$$\mathbf{Var}_\mu(f) \leqslant 4\, C_1^2 \int |\nabla f|^2\, d\mu.$$

2.4.2. One-dimensional CHEEGER inequality. — We have seen before that CHEEGER inequality may be considered as a comparison with the typical model of the exponential measure. As this model is also the typical model of log-concave

measures, it is not surprising that these measures satisfy a CHEEGER inequality. As for POINCARÉ inequality, to determine the exact value of the best constant $C_1(\mu)$ for a log-concave measure is of great interest. This highly non trivial question is still open in dimension more than 2. In dimension 1, the answer was given by BOBKOV and HOUDRÉ (see [BH97, Bob99]). We briefly present their results in the following.

2.4.2.1. CHEEGER *constant for* BOLTZMANN *measures on* $\mathbb{R}$. — Let

$$\mu(dx) = Z^{-1} \exp(-\Phi(x))\, dx$$

be a BOLTZMANN probability measure on $\mathbb{R}$ with Φ continuous and F be its distribution function. Then

***Theorem 2.4.1* (Bobkov-Houdré,** [BH97]). — *The* CHEEGER*'s constant is given by*

$$C_1 \overset{def.}{=} \sup_A \frac{\min(\mu(A), 1-\mu(A))}{\mu_s(\partial A)} = \sup_{x\in\mathbb{R}} \frac{\min(F(x), 1-F(x))}{F'(x)}.$$

In other terms, we may consider only half-lines in the supremum.

Sketch of proof. — For $A =]-\infty, x]$, we get

$$C_1 \geqslant K_\mu \overset{\text{def.}}{=} \sup_{x\in\mathbb{R}} \frac{\min(F(x), 1-F(x))}{F'(x)}.$$

Hence, we may assume that K_μ is finite. The function $k = F^{-1} \circ F_\epsilon$ has the same law under ϵ as the identity function $x \mapsto x$ under μ. Moreover, one easily check the LIPSCHITZ bound

$$k' \leqslant K_\mu.$$

Now, if f is of median m under μ, $g = f \circ k$ is of median m under ϵ and $g' \leqslant K_\mu\, f' \circ k$. Applying CHEEGER inequality with constant 1 for ϵ to the function k, we get

$$\int |f-m|\, d\mu = \int |g-m|\, d\epsilon \leqslant \int |g'|\, d\epsilon \leqslant K_\mu \int |f' \circ k|\, d\epsilon = K_\mu \int |f'|\, d\mu.$$

Finally, $C_1 \leqslant K_\mu$. □

2.4.2.2. CHEEGER *constant for log-concave measure on* $\mathbb{R}$. — In the case Φ convex, theorem 2.4.1 may be stated with much more precision. This is due to BOBKOV's characterization of BOLTZMANN and log-concave measures in terms of their isoperimetric functions ι. In few words, there is a one-to-one correspondence between the family of BOLTZMANN probability measures μ (up to translations) and the family of positive continuous functions ι on $]0,1[$ vanishing on the boundary. For a measure μ with median 0, it is given by

$$\iota = F' \circ F^{-1} \tag{103}$$

and

$$F^{-1}(u) = \int_{\frac{1}{2}}^{u} \frac{dp}{\iota(p)}, \quad u \in [0,1]. \tag{104}$$

If we assume as previously that the function Φ we consider is finite (i.e. the domain of μ is the whole line) then the function $1/\iota$ has to be non integrable at 0 and 1. The fundamental trick is that this correspondence maps the family of log-concave measures into the family of concave isoperimetric functions ι.

Theorem 2.4.2 **(Bobkov,** [Bob99]**).** — *Let $\mu(dx) = Z^{-1} \exp(-\Phi(x))\, dx$ be a log-concave probability measure of median m on the real line. Then, the* CHEEGER*'s constant is*

$$C_1 = \frac{1}{2} \exp(\Phi(m)).$$

In other words, the CHEEGER *inequality admits the two half-lines determined by m as extremal sets of measure $1/2$.*

Sketch of proof. — Once we understand the characterization in terms of isoperimetric functions, everything is done. Namely,

$$C_1 = \sup_{x\in\mathbb{R}} \frac{\min(F(x), 1-F(x))}{F'(x)} = \sup_{p\in]0,1[} \frac{\min(p, 1-p)}{\iota(p)}.$$

As ι is concave and coincides with the secant 2-piecewise linear function

$$2\,\iota(1/2) \min(p, 1-p)$$

at points $0, 1/2$ and 1, the latter lies below ι everywhere. Consequently,

$$C_1 \leqslant (2\,\iota(1/2))^{-1} = 1/2 \exp(\Phi(m)).$$

□

A new look at lemma 2.1.4 shows that if μ is log-concave on the real line, then its density is bounded from above by the inverse of its CHEEGER constant.

2.4.3. Proof of our result thanks to this tools. — BOBKOV proved in [Bob99] that

$$C_1^2 \leqslant 3\,\mathbf{Var}(\mu)$$

and then end up to the estimate

$$C_2 \leqslant 12\,\mathbf{Var}(\mu).$$

But we may use his method as well to get a similar bound in terms of the functional we studied. The proof is even simpler.

Theorem 2.4.3. — *Let $\mu(dx) = Z^{-1} \exp(-\Phi(x))\, dx$ be a log-concave probability measure of median m on the real line. Then*

$$C_1 \leqslant 2 \int |x-m| \mu(dx)$$

and consequently

$$C_2 \leqslant 16 \left(\int |x-m| \mu(dx) \right)^2.$$

Proof

◀ We may assume $m = 0$. According to (104), we have

$$\int_{-\infty}^{+\infty} |x|\, F'(x)\, dx = \int_0^1 \left|F^{-1}(u)\right| du = \int_0^1 \left| \int_{\frac{1}{2}}^{u} \frac{dp}{\iota(p)} \right| du.$$

Define $\iota_\alpha(p) = \iota(1/2) + \alpha\,(p - 1/2)$, the linear line with slope α which coincides with ι at point $1/2$. As ι is concave (and $\mathcal{C}^1$), it lies below its tangent line at point $1/2$. Let $\overline{\alpha} = \iota'(1/2)$. We get $\iota \leqslant \iota_{\overline{\alpha}}$ and $\int |x|\, \mu(dx) \geqslant H(\overline{\alpha})$ where we put

$$H(\alpha) \stackrel{\text{def.}}{=} \int_0^1 \left| \int_{\frac{1}{2}}^{u} \frac{dp}{\iota_\alpha(p)} \right| du, \quad \alpha \in]-2\,\iota(1/2), 2\,\iota(1/2)[.$$

Remark that this interval contains $\overline{\alpha}$ by the concavity of ι and that $H(\alpha)$ is well defined on it.

As a function of α, $1/\iota_\alpha(p)$ is convex, and consequently $H(\alpha)$ is convex as well. Moreover, thanks to a change of variable, we notice that $H(\alpha) = H(-\alpha)$, so that $H(\alpha) \geqslant H(0)$. Remark that $\iota_0 \equiv \iota(1/2)$ and we get

$$\int |x|\, \mu(dx) \geqslant \frac{1}{\iota(1/2)} \int_0^1 |u - 1/2|\, du \geqslant \exp\left(\Phi(m)\right)/4 = C_1/2.$$

▶

CHAPITRE 3

LIFTINGS OF A DIRICHLET FORM, CURVATURE CREATION AND SOBOLEV INEQUALITY

3.1. Introduction

The natural structure in which a SOBOLEV inequality may be stated is the general context of DIRICHLET forms. But to make things simpler, and since the methods developed below take place in a more restrictive setting, we shall restrict ourselves here to the case of elliptic operators on a smooth manifold.

Let M be a connected d-dimensional smooth manifold, and let (g) be a (co)metric tensor on it, that is, in a local system of coordinates, for any point $x \in M$, $(g^{ij})(x)$ is a symmetric definite positive matrix, and the coefficients of the matrix are smooth with respect to the x-variable. Then, the square field operator is defined on smooth functions f, g by

$$\mathbf{\Gamma}(f,g)(x) = \sum_{ij} g^{ij}(x) \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j}(x).$$

Moreover, we suppose we are given a measure $\mu(dx)$ on the manifold M, which is absolutely continuous with respect to the LEBESGUE measure in any local coordinates, with smooth density u.

The DIRICHLET form associated with this system is then

$$\mathcal{E}(f,f) = \int_M \mathbf{\Gamma}(f,f)\mu(dx).$$

A SOBOLEV inequality with dimension n is then valid if there exists two constants A and B such that, for any smooth compactly supported function $f : M \mapsto \mathbb{R}$, one has

$$(\int f^{2n/(n-2)}\mu(dx))^{(n-2)/n} \leqslant A \int f^2 \mu(dx) + B \int \mathbf{\Gamma}(f,f)\mu(dx). \tag{105}$$

If the measure is finite, we may always choose the measure μ to be a probability measure, up to a trivial change in the constants A and B. Moreover, it is then obvious that $A \geqslant 1$, at least when the constant functions are limits in the DIRICHLET norm $\left(\|\cdot\|_{2,\mu}^2 + \mathcal{E}(\cdot,\cdot)\right)^{\frac{1}{2}}$ of smooth functions with compact support. This is true

in particular in the compact case, but still holds true if the manifold is complete for the metric induced by the RIEMANN metric. We shall always assume that this normalization of the measure is made if it is finite in what follows, in order to be able to compare the different values of A and B for different models. Moreover, it is true in this case that if a SOBOLEV inequality holds with some constants $A \geqslant 1$ and B, then it holds with $A = 1$ if and only if there exists a spectral gap (see [ABC$^+$00] or [Bak94], for example).

If the manifold is compact, then such an inequality always holds true with any $n \geqslant d$. From the previous remark, it also holds with $A = 1$. Then, the best value we may expect for B depends then on global properties of the manifold, and more precisely of the pair (g, μ) which characterizes our model.

For example, let M be the standard unit sphere in the Euclidean space $\mathbb{R}^{d+1}$. The natural metric is inherited from the Euclidean metric : for a function f defined on M, we extend it in a neighbourhood of the sphere such that it is independent of the radius to a function F, and we take the restriction of $|\nabla F|^2$ to M to be $\mathbf{\Gamma}(f, f)$. This metric gives the Euclidean length to the velocity vector of any path on the sphere. Then, we may take the measure μ to be the uniform measure of the sphere, that is the unique probability measure on M which is invariant by rotations. Then, a SOBOLEV inequality holds for $d = n$ with $A = 1$ and $B^{-1} = d(d-2)/4$. This optimal result is due to AUBIN (see [Aub82]).

It must be noted that a SOBOLEV inequality may or may not hold on a non compact manifold. For example, on the Euclidean space $\mathbb{R}^d$, with $\mathbf{\Gamma}(f, f) = |\nabla f|^2$, and $\mu(dx) = dx/a_d$, where a_d is the surface of the unit ball in $\mathbb{R}^{d+1}$, then the SOBOLEV inequality holds with $n = d$, $A = 0$ and $B^{-1} = d(d-2)/4$, while on the d-dimensional hyperbolic space H^d with $\mathbf{\Gamma}(f, f) = |\nabla f|^2$, (computed in the Riemannian metric), and $\mu(dx) = m(dx)/a_d$, $m(dx)$ being the RIEMANN measure and a_d as before, the SOBOLEV inequality holds with $A = -1$ and $B^{-1} = d(d-2)/4$.

The choice of the normalizing constant in the measures μ in the above examples is made to get simple expressions for A and B. The remarkable symmetry between the case of spheres, Euclidean and hyperbolic spaces is then due to a invariance property of SOBOLEV inequalities under conformal transformations, and any of these three models is a conformal image of each other.

On the other hand, if we come back to the Euclidean case, but if we change the measure μ to the standard Gaussian measure in $\mathbb{R}^d$

$$\mu(dx) = \exp(-|x|^2/2)dx/(2\pi)^{d/2}$$

then no SOBOLEV inequality holds, and these inequalities are then replaced by logarithmic SOBOLEV inequalities.

There are few methods to control SOBOLEV inequalities. Using BAKRY-EMERY curvature-dimension criterion is one of them, which gives for example the optimal constant $d(d-2)/4$ for spheres (see [Aub82, BÉ85b, Bak94, Led00]).

In order to describe this curvature-dimension criterion, we have to introduce a few notions. First, let us introduce the diffusion generator $\mathbf{L}$ associated to our pair

(g, μ). It may be defined on smooth functions f by the following integration by parts formula :

$$\forall g, \quad \int \mathbf{L}(f) g \, d\mu = - \int \mathbf{\Gamma}(f, g) \, d\mu.$$

The operator $\mathbf{L}$ is then a second order differential operator with no 0-order terms. It's expression in a local system of coordinates is not very simple in terms of (g, μ), but is quite simple if we introduce the LAPLACE-BELTRAMI operator Δ associated to the co-metric g. Then it is

$$\mathbf{L}(f) = \Delta(f) + \mathbf{\Gamma}(\log u, f),$$

where u denotes the density of μ with respect to the RIEMANN measure. In this case, we may recover $\mathbf{\Gamma}(f, g)$ and μ from $\mathbf{L}$ by the fact that

$$\mathbf{\Gamma}(f, g) = \frac{1}{2}(\mathbf{L}(fg) - f\mathbf{L}(g) - g\mathbf{L}(f)),$$

and, for any smooth compactly supported f, $\int \mathbf{L}(f)\mu(dx) = 0$. Therefore, the operator $\mathbf{L}$ entirely characterizes the pair (g, μ), and we shall talk of a SOBOLEV inequality related to $\mathbf{L}$ as well as related to the pair (g, μ).

With the help of $\mathbf{L}$, we may introduce the new bilinear operator $\mathbf{\Gamma_2}$ which may be defined as

$$\mathbf{\Gamma_2}(f, g) = \frac{1}{2}(\mathbf{\Gamma}(f, g) - \mathbf{\Gamma}(f, \mathbf{L}g) - \mathbf{\Gamma}(g, \mathbf{L}f)).$$

It's expression in a local system of coordinates is even worse than the expression of $\mathbf{L}$, but may be written in a simpler way with the help of a bit of Riemannian geometry. If we introduce the LEVI-CIVITA connection ∇ associated with the metric (g), and the RICCI tensor Ric of this Riemannian structure, then we have

$$\mathbf{\Gamma_2}(f, f) = |\nabla\nabla f|^2 + (\mathrm{Ric} - \nabla\nabla \log u)(df, df).$$

(See [Bak94] or [ABC+00] for more details, and below for explicit computations on the sphere.)

Then, for any real numbers ρ and $n \geqslant 1$, the curvature-dimension hypothesis $\mathrm{CD}(\rho, n)$ is the following :

$$\forall f, \quad \mathbf{\Gamma_2}(f, f) \geqslant \rho \, \mathbf{\Gamma}(f, f) + \frac{1}{n} \, (\mathbf{L}f)^2. \tag{106}$$

Necessarily, we have $n \geqslant d$. In the case $n > d$, this amounts to say that, as symmetric tensors,

$$\mathrm{Ric} - \nabla\nabla \log u \geqslant \rho g + \frac{1}{n-d} \nabla \log u \otimes \nabla \log u.$$

The case $n = d$ implies that μ is the RIEMANN measure or equivalently $\mathbf{L} = \Delta$, and this is simpler

$$\mathrm{Ric} \geqslant \rho g,$$

which just means that the RICCI tensor is bounded below by the constant ρ.

Then, the curvature-dimension criterion for SOBOLEV inequalities says that, if $\mathrm{CD}(\rho, n)$ holds for some $\rho > 0$ and some $n \geqslant 2$, then first the measure μ has to be finite, and, when normalized to be a probability measure, the SOBOLEV inequality (105) holds with $A = 1$ and $B^{-1} = \frac{\rho}{n-1}\frac{n(n-2)}{4}$. If we notice that, on the d-dimensional sphere, $CD(d-1, d)$ holds, then one recover the sharp constant of [Aub82].

Let us point out one of the main defect of this criterion : basically, it compares the model we are working with with a sphere, through the use of a RICCI curvature tensor associated to our pair (g, μ). If this RICCI tensor is not bounded from below by a positive constant, then, it gives nothing.

Here we propose a new approach based on a simple idea : if the pair (g, μ) does not verify the BAKRY-EMERY criterion, we get around this difficulty by constructing a pair $(\overline{g}, \overline{\mu})$ in a higher dimension with good curvature properties. To be useful, this lifting $(\overline{g}, \overline{\mu})$ has to project *essentially* on the operator. We give some criterion on the pair (g, μ) to construct such a lifted pair $(\overline{g}, \overline{\mu})$. The basic property of this lifting is that the operator $\mathbf{L}$ itself is not the projection of the lifted operator, and it is this feature which enables us to relax the curvature condition.

The fundamental trick here is that we perform this lifting through a conformal change of the basic metric g. This makes the computations a bit tedious, but in order to understand where this comes from, we start with the presentation of a simple example which comes from the analysis on the sphere and classical geometry.

We devote section 3.3.1 to the abstract construction of the lifted operator and give some criterion to obtain a SOBOLEV inequality for $\mathbf{L}$ by this means. The following sections consist in a gradual particularization in order to take advantage of the general result. The proof of the main result is given in the last section.

3.2. A simple example in spherical geometry

In this section, we present a fundamental example (in spherical geometry) of a diffusion generator without curvature-dimension but which still satisfies a SOBOLEV inequality. The latter may be obtained from the curvature-dimension criterion in higher dimension.

3.2.1. Spherical Laplacian. — Let $S^m \subset \mathbb{R}^{m+1}$ be the unit sphere of m dimensions and consider Δ_{S^m} the LAPLACE-BELTRAMI operator for the induced Riemannian metric inherited from the Euclidean metric of $\mathbb{R}^{m+1}$. To make explicit computations, we shall use the stereographical projection, which has the advantage of representing the spherical metric as a conformal change of the Euclidean one.

Let us recall that the projection $x = (x^1, \ldots, x^m) \in \mathbb{R}^m$ of $\overline{x} = (\overline{x}^1, \ldots, \overline{x}^{m+1}) \in S^m$ is given by

$$x^i = \frac{\overline{x}^i}{1 - \overline{x}^{m+1}},$$

for $i = 1, \ldots, m$. Conversely, we have

$$\left| \begin{array}{ll} \overline{x}^j & = \frac{2x^j}{\overline{r}^2+1}, \text{ for } j = 1, \ldots m, \\ \overline{x}^{m+1} & = 1 - \frac{2}{\overline{r}^2+1}, \end{array} \right.$$

where $\overline{r}^2 \stackrel{\text{def.}}{=} (x^1)^2 + \ldots + (x^m)^2$. The spherical metric (g_{ij}) on $\mathbb{R}^m$ is then easy to obtain by computing the induced basis $\left(\frac{\partial}{\partial x^i}\right)$ and the associated GRAMM matrix. So

$$g_{ij} = \left(\frac{2}{\overline{r}^2 + 1}\right)^2 \delta_{ij}.$$

Since we are interested in co-metrics, our square field operator here is

$$\mathbf{\Gamma}(f, f) = \left(\frac{\overline{r}^2 + 1}{2}\right)^2 |\nabla f|^2,$$

where $|\nabla f|^2$ is the Euclidean length of the vector field $\nabla f = (\frac{\partial f}{\partial x^i})$.

Also, the RIEMANN measure is

$$\sigma(dx) = \left(\frac{2}{\overline{r}^2 + 1}\right)^m dx,$$

where dx denotes the usual LEBESGUE measure in $\mathbb{R}^m$.

We often will use the notation $\overline{v} = \frac{1-\overline{r}^2}{1+\overline{r}^2} = -\overline{x}^{m+1} \in]-1, 1]$, so that

$$g_{ij} = (1 + \overline{v})^2 \delta_{ij}.$$

We hence see that Euclidean and spherical metrics are conformal. This computation simply hides the fact that the stereographical projection is the restriction to the unit sphere of the inversion in $\mathbb{R}^{m+1}$ which maps the unit sphere onto the hyperplane $\{\overline{x}^{m+1} = 0\}$, and that inversions are conformal transformations in $\mathbb{R}^{m+1}$.

Let us recall that the LAPLACE-BELTRAMI operator is the trace of the second covariant derivative. A simple way to get an explicit expression is to use the conformal representation.

The general computation of LAPLACE-BELTRAMI operator by conformal transformations is easy, and comes from the definition of Δ_{S^m} by use of the pair(g, μ). In fact, one may easily verify that, if $\mathbf{L}$ is the operator associated to the pair (g, μ), the operator associated to the pair $(h^{-2}g, h^m\mu)$ is

$$\Delta_{h,m}(f) = h^{-2}(\mathbf{L}(f) + (m-2)\mathbf{\Gamma}(\log h, f)).$$

Here, with $\mathbf{L}$ being the Euclidean operator Δ and $h = (1 + \overline{v})$, we get, for $f \in \mathcal{C}_c^\infty(\mathbb{R}^m)$,

$$\Delta_{S^m} f = (1 + \overline{v})^{-2} \left(\sum_{i=1}^m \partial_{ii}^2 f - (m-2)\,(1 + \overline{v})^{-1} \sum_{i=1}^m x^i\, \partial_i f \right). \tag{107}$$

The normalizing constant ω_m which changes the RIEMANN measure into a probability measure is

$$\omega_m = \frac{2\,\pi^{\frac{m+1}{2}}}{\Gamma\big(\frac{m+1}{2}\big)},$$

which is the surface of the unit sphere in $\mathbb{R}^{m+1}$, and therefore our reference measure shall be

$$\sigma_m(dx) = (1+\overline{v})^m \frac{dx}{\omega_m}.$$

The spherical LAPLACE-BELTRAMI operator has a RICCI curvature which is very simple, since

$$\mathrm{Ric} = (m-1)g$$

and in particular the inequality $\mathrm{CD}(m-1,m)$ holds. Consequently, we get the following SOBOLEV inequality : for all function $f \in \mathcal{C}_c^\infty(\mathbb{R}^m)$,

$$\|f\|^2_{\frac{2m}{m-2},\sigma_m} \leqslant \|f\|^2_{2,\sigma_m} + \frac{4}{m\,(m-2)}\|(1+\overline{v})^{-2}\,|\nabla f|^2\|_{1,\sigma_m}, \tag{108}$$

where $\|\cdot\|_{p,\mu}$ stands for the norm in $\mathbf{L}^p(d\mu)$.

3.2.2. Projection of the DIRICHLET form. — Write $\mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^p$ for n,p positive integers and note, for $x \in \mathbb{R}^m$, $x = (y,z)$ with $y \in \mathbb{R}^n$ and $z \in \mathbb{R}^p$. In order to project the system (g,σ_m) from $\mathbb{R}^m$ to $\mathbb{R}^n$, we just consider functions f on $\mathbb{R}^m$ which depends only on the first coordinate y. The new measure $\sigma_{m,n}$ is just the usual projection of the measure σ_m. We then get a measure and a DIRICHLET form on $\mathbb{R}^n$ which corresponds to a system (g_1,μ_1) in $\mathbb{R}^n$. Precisely, we get the measure

$$\sigma_{m,n}(dy) = (1+v)^q \frac{dy}{\omega_{n,q}},$$

where $q \stackrel{\text{def.}}{=} \frac{m+n}{2}$. Here, we posed

$$v = \frac{1-r^2}{1+r^2}$$

with $r^2 = (y^1)^2 + \ldots + (y^n)^2$ and

$$\omega_{n,q} \stackrel{\text{def.}}{=} \int_{\mathbb{R}^n} \left(\frac{2}{1+|y|^2}\right)^q dy = \pi^{\frac{n}{2}}\,2^q\,\frac{\Gamma\big(q-\frac{n}{2}\big)}{\Gamma(q)}.$$

For the projection of the DIRICHLET form, we get

$$\int_{\mathbb{R}^m} (1+\overline{v})^{-2}\,|\nabla f(y)|^2\,\sigma_m(dx) = \frac{\omega_{p,m-2}\,\omega_{n,q}}{2^{q-2}\,\omega_m}\int_{\mathbb{R}^n} (1+v)^{-2}\,|\nabla f(y)|^2\,\sigma_{m,n}(dy).$$

Therefore, we see that the projected square field operator is proportional to $(1+v)^{-2}|\nabla f|^2$, and is again the square field operator of an n-dimensional sphere, and the metric is again the spherical metric g_{S^n}. The important remark here is that the

diffusion operator $\mathbf{L}_{m,n}$ associated to the pair $(g_{S^n}, \sigma_{m,n})$ is not the projection of the spherical LAPLACE-BELTRAMI operator Δ_{S^m}. This operator does not in fact projects on $\mathbb{R}^n$, since if f is a function which depends only on y, $\Delta_{S^m}(f)$ may in general depend on z as well.

From the SOBOLEV inequality on the sphere S^m, we directly get by projection the following. For all $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$,

$$\|f\|^2_{\frac{2m}{m-2},\sigma_{m,n}} \leqslant \|f\|^2_{2,\sigma_{m,n}} + B_{m,n} \left\|(1+v)^{-2} |\nabla f|^2\right\|_{1,\sigma_{m,n}}, \tag{109}$$

where

$$B_{m,n} = \frac{4\,\omega_{p,m-2}\,\omega_{n,q}}{m\,(m-2)\,2^{q-2}\,\omega_m} = \frac{16\,(m-1)}{m\,(m+n-2)\,(m+n-4)}. \tag{110}$$

Now, the measure $\sigma_{m,n}$ and the DIRICHLET form $\left\|(1+v)^{-2} |\nabla f|^2\right\|_{1,\sigma_{m,n}}$ define the operator

$$\mathbf{L}_{m,n} = \Delta_{S^n} + \nabla\Phi \tag{111}$$

with

$$\Phi = \frac{p}{2} \log(1+v)$$

(remember that $p = m-n$). Hence, the operator $\mathbf{L}_{m,n}$ satisfies a SOBOLEV inequality.

But the associated pair $(g_{S^n}, \sigma_{m,n})$ does not satisfy any $\mathrm{CD}(\rho, k)$ criterion for any $k \in [1, \infty]$.

Recall that an operator $\mathbf{L} = \Delta + \nabla\Phi$ on a Riemannian manifold of dimension n satisfies a $\mathrm{CD}(\rho, k)$ inequality for $k > n$ if

$$A \stackrel{\text{def.}}{=} (k-n)\,(\mathrm{Ric} - \nabla\nabla\Phi - \rho\, g) - \nabla\Phi \otimes \nabla\Phi \geqslant 0. \tag{112}$$

Here Ric and ∇ are related to the Riemannian structure defining Δ. We hence have to estimate the eigenvalues of the bilinear form A. This will be really easier thanks to some striking properties of the function v we have to emphasize. That is

$$\mathbf{\Gamma}(v) = \mathbf{\Gamma}_{S^n}(v, v) = 1 - v^2 \tag{113}$$

and

$$\nabla\nabla v = -v g_{S^n}, \tag{114}$$

in the spherical metric. The first identity is straightforward and the second may be done either from a direct computation in the Riemannian structure, either from general formula for the Hessian through conformal transformations, as described below in formula 132. Notice that if we take the trace in the second formula, we obtain $\Delta_{S_n}(v) = -nv$, which shows that v is an eigenvector of Δ_{S_n} associated to the first non zero eigenvalue.

Moreover,

$$\nabla\Phi = \frac{p}{2}\,(1+v)^{-1} \nabla v$$

and

$$\nabla\nabla\Phi = \frac{p}{2}\left((1+v)^{-1}\nabla\nabla v - (1+v)^{-2}\nabla v\otimes\nabla v\right).$$

Hence,

$$A = a(v)g + b(v)\,\nabla v\otimes\nabla v \tag{115}$$

where

$$a(v) = (k-n)\left((n-1-\rho) + \frac{p}{2}\frac{v}{1+v}\right)$$

and

$$b(v) = \frac{p}{2}\left(k-n-\frac{p}{2}\right)(1+v)^{-2}.$$

So, the eigenvalues of A are nothing but $a(v)$ on $(\nabla v)^{\perp}$ and $a(v)+b(v)\mathbf{\Gamma}(v)$ in the direction ∇v. Since $a(v)$ goes to $-\infty$ when v goes to -1, (112) can never occur for $\mathbf{L}_{m,n}$.

Let's now look into the eigenvalue $c(v) = a(v) + b(v)|\nabla v|^2$ in the direction ∇v. Thanks to (113), we can write

$$c = \alpha + \frac{\beta+\gamma v}{1+v}$$

with $\alpha = (k-n)(n-1-\rho)$, $\beta = \frac{p}{2}\left(k-n-\frac{p}{2}\right)$ and $\gamma = \left(\frac{p}{2}\right)^2$. Requiring c to be non negative thus leads to $\beta \geqslant \gamma$, that is $k \geqslant n+p$. Now, for $k = n+p$, $c = \alpha+\gamma$ is non negative provided that

$$\rho \leqslant (n-1) + \frac{p}{4} = \frac{3\,(n-1)+(m-1)}{4}.$$

Therefore, the inequality $\mathrm{CD}(\rho, n+p)$ is valid with $\rho = \frac{3\,(n-1)+(m-1)}{4}$ provided that we restrict our attention on functions of v (i.e. radial functions). We shall see in the next paragraph that this corresponds to curvature-dimension inequalities for non-symmetric JACOBI operators.

3.2.3. JACOBI operators. — Let us recall that the JACOBI operator with parameters m and n ($m \geqslant n$) acts on a $\mathcal{C}^1$ function f on $]-1,1[$ by

$$\tilde{\mathbf{L}}_{m,n}\,f(v) = \left(1-v^2\right)f''(v) - \frac{1}{2}((m+n)\,v + (n-m))\,f'(v).$$

It is a diffusion operator on $]-1,1[$ associated with the pair

$$\tilde{\mathbf{\Gamma}}_{m,n}(f,f)(v) = \left(1-v^2\right)f'^2(v)$$

and

$$\tilde{\sigma}_{m,n}(dv) = (1-v)^{\frac{n}{2}-1}\,(1+v)^{\frac{m}{2}-1}\,\frac{dv}{Z_{m,n}}.$$

The normalization constant $Z_{m,n}$ can be computed as

$$Z_{m,n} = 2^{\frac{m+n}{2}-1} \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m+n}{2}\right)}, \tag{116}$$

where Γ stands here for the classical Γ function.

It was already shown in [Bak96] that such an operator satisfies the previous

$$\mathrm{CD}\left(\frac{3\,(n-1)+(m-1)}{4}, m\right)$$

inequality, and deduced from this a SOBOLEV inequality for $\tilde{\mathbf{L}}_{m,n}$ with constant

$$B'_{m,n} = \frac{16\,(m-1)}{m\,(m-2)\,(m+3\,n-4)}. \tag{117}$$

Now, if the consider the previous operator $\mathbf{L}_{m,n}$ and let it act on functions of v only (that is on functions which depend only on $|y|^2$), we get

$$\mathbf{L}_{m,n} f(v) = f''(v)\mathbf{\Gamma}_{S^n}(v) + f'(v)\mathbf{L}_{m,n}(v).$$

But

$$\mathbf{L}_{m,n}(v) = \Delta_{S_n}(v) + \frac{p}{2(1+v)}\mathbf{\Gamma}_{S^n}(v) = -\frac{1}{2}((m+n)v + (n-m)).$$

Therefore,

$$\mathbf{L}_{m,n}(f(v)) = \tilde{\mathbf{L}}_{m,n}(f)(v).$$

Therefore, the action of $\mathbf{L}_{m,n}$ on functions of v coincides with the JACOBI operator. But the SOBOLEV inequality for $\tilde{\mathbf{L}}_{m,n}$ obtained by projection gives a constant

$$B_{m,n} = \frac{16(m-1)}{m(m+n-2)(m+n-4)},$$

while the constant obtained directly from the $\mathrm{CD}(\rho, m)$ criterion gives

$$B'_{m,n} = \frac{16(m-1)}{m(m-2)(m+3n-4)},$$

and one may directly check that $B'_{m,n} < B_{m,n}$.

Let us notice that this construction of JACOBI operators $\tilde{\mathbf{L}}_{m,n}$ from an m-dimensional sphere by projection of the DIRICHLET structure is different from the one obtained in [Bak96]. The latter was based on a projection of the LAPLACE-BELTRAMI operator itself from a sphere with dimension $m+n-1$.

The structure $(g_{S^n}, \sigma_{m,n})$ may also be projected on the sphere S^{n-1}, using functions which do not depend on the radial part. Then, we obtain on the sphere S^{n-1} the structure $(c_{n,m} g_{S^{n-1}}, \sigma_{S^{n-1}})$, where

$$c_{n,m} = Z_{m-2,n-2}/Z_{m,n} = \frac{(n+m-2)(m+n-4)}{4(m-2)(n-2)}.$$

We then obtain a SOBOLEV inequality for the sphere S^{n-1}, but with an exponent m instead of $n-1$, and a constant which may be explicitely computed. But the optimal constant of the SOBOLEV inequality on the $(n-1)$-dimensional sphere with exponent m is $\frac{4}{(m-2)(n-1)}$ (see [Bak94] or [Led00]), and one may check that the constant such obtained is weaker. The reason why this is true is that the optimal functions for the SOBOLEV inequality on S^{n-1} are far from being optimal when lifted on the sphere S^m through this projection.

In fact, we could imagine that projecting a pair (g, m) on one of it's factor in a product amount to average the RICCI tensor on the other variable. This is wrong, since the above projection of the sphere S^m gives a pair with a RICCI tensor which is not bounded below, and on the other hand, the two projection of this projected structure, on the radius and on the angular part, give in turn rise to strictly positive RICCI tensors. Therefore, this strange behaviour of RICCI tensors under projections of DIRICHLET forms allows us to establish SOBOLEV inequalities in situations where no $\mathrm{CD}(\rho, n)$ inequality occurs.

Remark also that the previous SOBOLEV inequality we obtained for the operator $\mathbf{L}_{m,n}$ is a priori only valid on $\mathbb{R}^n$ for an integer $m \geqslant n+1$. One of the features of the tools that we develop below is that we shall be able to extend it to any $m \geqslant n+1$, even when it is not an integer.

3.3. Weighted SOBOLEV inequality and lifting of a DIRICHLET form

As we saw in the previous example, it is possible for a pair (g, μ) with finite measure μ which has a RICCI curvature which is not bounded from below to be the projection of another pair $(\overline{g}, \overline{\mu})$ on a higher dimensional space which satisfies a $\mathrm{CD}(\overline{\rho}, m)$ inequality with a strictly positive $\overline{\rho}$ and finite m. In such a situation, then the pair satisfies a SOBOLEV inequality, and therefore the manifold has to be compact (or at least of finite diameter for the natural distance induced by g, see [BL96b]). The question we address in this section is to determine simple conditions for this to occur. As we shall see, this question is too complicated to answer in the general setting, and we shall restrict our attention to a more restrictive setting. We shall propose a generic construction, which leads to a set of equations which are quite difficult to handle. In the end, we shall give a complete answer only in a construction which mimics the case studied in the previous example issued from spheres, but we propose the complete set of equation, since even some small modifications of the basic example could lead to completely different conditions.

The precise general computations are quite tedious and are done in the section 3.4 below. The basic remark is that, in the previous example, it was enough to consider functions which are radial in the extra dimensions, and therefore it would have been enough to consider 1-dimensional extensions of our basic manifold.

3.3.1. General construction of a lifted DIRICHLET form. — Suppose we are given a pair (g, μ) on a given n-dimensional manifold. We shall assume that the

measure μ has a density $\exp(\Phi)$ with respect to the Riemannian measure σ related to g. Then, we consider the space $\overline{M} \stackrel{\text{def.}}{=} M \times \mathbb{R}_+^*$ on which we put the metric with square field operator

$$\overline{\mathbf{\Gamma}}(f,f)(x,r) = e^{-2\overline{\tau}} \left(e^{2\tau}\, \mathbf{\Gamma}_x(f,f) + (\frac{\partial f}{\partial r})^2 \right),$$

associated to the measure with density $\exp \overline{\Phi}$ with respect to the Riemannian measure associated to this metric on $\overline{M}$.

The problem is now to devise conditions on (g,σ) which insure the existence of good functions $\overline{\tau}(x,r)$, $\tau(x)$, $\overline{\Phi}(x,r)$ on the product structure. In such a generality, the problem is out of reach, but we shall restrict our choice to functions $\overline{\tau}(x,r)$, $\tau(x)$, $\overline{\Phi}(x,r)$ which are of the form $\overline{\tau}(\Phi(x),r)$, $\tau(\Phi(x))$ and $\overline{\Phi}(\Phi(x),r)$. Our aim is to describe conditions on those three functions, of two and one variables respectively, which insures a $\mathrm{CD}(\overline{\rho},m)$ inequality.

The lifted Riemannian metric $\overline{g}$ is actually obtained after three steps :

– a first conformal transformation from $g_{..}$ to $\check{g}_{ij} \stackrel{\text{def.}}{=} e^{-2\,\tau} g_{ij}$.

– a simple tensorization : $\hat{g} \stackrel{\text{def.}}{=} \check{g} \oplus 1$.

– another conformal step : $\overline{g}_{kl} \stackrel{\text{def.}}{=} e^{2\,\overline{\tau}} \hat{g}_{kl}$.

Considering all the objects w.r.t. $\overline{g}$, the operator $\overline{\mathbf{L}}$ can be simply written

$$\overline{\mathbf{L}} = \overline{\Delta} + \overline{\nabla\Phi}.$$

We have also an explicit expression

$$\overline{\sigma}(dy,dr) = e^{(n+1)\overline{\tau} - n\,\tau} \sigma(dy) \otimes dr$$

of the lifted Riemannian measure.

Let's now summarize under which conditions on the function Φ the generator $\overline{\mathbf{L}}$ satisfies a $\mathrm{CD}(\overline{\rho}, n+p)$ inequality for some real numbers $p > 1$ and $\overline{\rho} > 0$. With this aim in view, we need to introduce some notations. The functions $\overline{\tau}(\varphi,r)$, $\overline{\Phi}(\varphi,r)$ on $\mathbb{R} \times \mathbb{R}_+^*$ and $\tau(\varphi)$ on $\mathbb{R}$ being fixed, define the following related functions.

On one side,

$$
\begin{aligned}
\nu_1(\varphi,r) &\stackrel{\text{def.}}{=} -\partial_\varphi \overline{\tau},\\
\nu_2(\varphi,r) &= \partial_\varphi(\tau-\overline{\tau}),\\
\pi_1(\varphi,r) &= -\partial^2_{\varphi\varphi}\overline{\tau} + (n-2)\,\partial_\varphi\tau\,\partial_\varphi\overline{\tau} - (n-1)\left(\partial_\varphi\overline{\tau}\right)^2 - \partial_\varphi\overline{\Phi}\,\partial_\varphi\overline{\tau},\\
\pi_2(\varphi,r) &= \partial^2_{\varphi\varphi}(\tau-\overline{\tau}) - (n-2)\left(\partial_\varphi\tau\right)^2\\
&\qquad + (2\,n-3)\,\partial_\varphi\tau\,\partial_\varphi\overline{\tau} - (n-1)\left(\partial_\varphi\overline{\tau}\right)^2 - \partial_\varphi\overline{\Phi}\left(\partial_\varphi\overline{\tau}-\partial_\varphi\tau\right),\\
\chi_1(\varphi,r) &= e^{-2\,\tau}\Big\{-n\,\partial^2_{rr}\overline{\tau} - \partial^2_{rr}\overline{\Phi} + \partial_r\overline{\Phi}\,\partial_r\overline{\tau} - (p-1)^{-1}\left(\partial_r\overline{\Phi}\right)^2 - \overline{\rho}\,e^{2\,\overline{\tau}}\Big\},\\
\chi_2(\varphi,r) &= -e^{-2\,\tau}\left(\partial^2_{rr}\overline{\tau} + (n-1)\left(\partial_r\overline{\tau}\right)^2 + \partial_r\overline{\Phi}\,\partial_r\overline{\tau} + \overline{\rho}\,e^{2\,\overline{\tau}}\right),\\
\kappa(\varphi,r) &= (n-2)\,\partial_\varphi\tau - (n-1)\partial_\varphi\overline{\tau} - \partial_\varphi\overline{\Phi},\\
\eta(\varphi,r) &= (n-2)\left\{\partial^2_{\varphi\varphi}\tau + \left(\partial_\varphi\tau\right)^2\right\} - (n-1)\left\{\partial^2_{\varphi\varphi}\overline{\tau} - \left(\partial_\varphi\overline{\tau}\right)^2\right\}\\
&\qquad - 2\,(n-1)\,\partial_\varphi\tau\,\partial_\varphi\overline{\tau} - \partial^2_{\varphi\varphi}\overline{\Phi} + 2\,\partial_\varphi\overline{\Phi}\left(\partial_\varphi\overline{\tau}-\partial_\varphi\tau\right) - (p-1)^{-1}\left(\partial_\varphi\overline{\Phi}\right)^2,\\
\zeta(\varphi,r) &= (n-1)\left\{\partial_r\overline{\tau}\,\partial_\varphi\overline{\tau} - \partial^2_{\varphi r}\overline{\tau}\right\} - \partial^2_{\varphi r}\overline{\Phi}\\
&\qquad + \partial_\varphi\overline{\Phi}\,\partial_r\overline{\tau} + \partial_r\overline{\Phi}\,\partial_\varphi\overline{\tau} - (p-1)^{-1}\,\partial_\varphi\overline{\Phi}\,\partial_r\overline{\Phi}
\end{aligned}
$$

are functions on $\mathbb{R}\times\mathbb{R}^*_+$. On the other side, we will consider some functions on $M\times\mathbb{R}^*_+$

$$
\begin{aligned}
a(\cdot,r) &\stackrel{\text{def.}}{=} \nu_1(\Phi,r)\,\Delta\Phi + \pi_1(\Phi,r)\,\mathbf{\Gamma}(\Phi) + \chi_1(\Phi,r),\\
\xi(\cdot,r) &= \nu_2(\Phi,r)\,\Delta\Phi + \pi_2(\Phi,r)\,\mathbf{\Gamma}(\Phi) + \chi_2(\Phi,r)
\end{aligned}
$$

and finally

$$
\theta(\cdot,r) = \left|\begin{array}{ll} e^{-2\,\tau}\,a^{-1}(\cdot,r)\,\zeta^2(\Phi,r) & \text{if } a\neq 0,\\ 0 & \text{otherwise.}\end{array}\right.
$$

With these notations,

Theorem 3.3.1 **(Curvature-dimension for the lifted operator)**
The lifted operator $\overline{\mathbf{L}} = \overline{\Delta}+\overline{\nabla\Phi}$ *defined as before satisfies a* $\mathrm{CD}(\overline{\rho},n+p)$ *inequality for some real numbers* $p>1$ *and* $\overline{\rho}>0$ *if and only if the three following conditions are verified*

i. $a(y,r)\geqslant 0$ *on* $M\times\mathbb{R}^*_+$,

ii. *for any* $(y,r)\in M\times\mathbb{R}^*_+$, *we have*

$$
a(y,r)=0 \Rightarrow \zeta(\Phi(y),r)\,\nabla\Phi(y)=0,
$$

iii. *for any* $r\in\mathbb{R}^*_+$, *the quadratic form*

$$
h \stackrel{\text{déf.}}{=} \mathrm{Ric} + \xi(\cdot,r)\,g + \kappa(\Phi,r)\,\nabla\nabla\Phi + \left\{\eta(\Phi,r)-\theta(\cdot,r)\right\}\nabla\Phi\otimes\nabla\Phi \tag{118}
$$

is non negative.

Here, Ric stands for the RICCI tensor associated to the metric g. The proof of this theorem will be given in section 3.4.

Remark 3.3.2. — *When*

$$\zeta \equiv 0, \tag{119}$$

the conditions are reduced to $a \geqslant 0$ *and*

$$\mathrm{Ric} + \xi(\cdot, r)\, g + \kappa(\Phi, r)\, \nabla\nabla\Phi + \eta(\Phi, r)\, \nabla\Phi \otimes \nabla\Phi \geqslant 0 \tag{120}$$

and the problem is relatively easier to handle.

Notice that if we stick to this case, then the equation $\zeta = 0$ is a (partial) differential equation which relates $\overline{\tau}$ and $\overline{\Phi}$. We shall see below that the fact that the DIRICHLET form on $\overline{M}$ projects on the DIRICHLET form on M shows that τ is determined from $\overline{\tau}$ and $\overline{\Phi}$, and therefore the only real choice in this case is the choice of the function $\overline{\tau}$. But even that amounts to a really complicated system, and we shall just explore some examples.

3.3.2. Projection conditions. — To get a SOBOLEV inequality on M, we shall apply a SOBOLEV inequality on $\overline{M}$ on functions which do not depend on r. One could expect to derive the latter inequality by means of the $\mathrm{CD}(\overline{\rho}, m)$ criterion for the pair $(\overline{g}, \overline{\mu})$ given by conditions (i), (ii) and (iii) of theorem 3.3.1.

Nevertheless, the manifold $\overline{M}$ is not complete in general for the new metric as in the example described before and we have to apply the curvature-dimension criterion with a little care. It appears that we may apply the $\mathrm{CD}(\overline{\rho}, m)$ criterion on the manifold with boundary $M \times [\varepsilon, 1/\varepsilon]$ provided we use NEUMANN boundary conditions. A closer look shows that there is no problem with the extra boundary terms.

The limit inequality for a function f only depending on $y \in M$ gives rise to a SOBOLEV inequality for $\mathbf{L}$ provided that

– $\overline{\mu}$ projects on μ,

– the DIRICHLET form of $\overline{\mathbf{L}}$ also projects on $\mathbf{L}$'s one, i.e.

$$\|\overline{\mathbf{\Gamma}}(f)\|_{1,\overline{\mu}} = \|\mathbf{\Gamma}(f)\|_{1,\mu}.$$

Hence, we get the following

Corollary 3.3.3 **(Projected SOBOLEV inequality)**. — *Suppose the two functions* $\overline{\tau}(\varphi, r)$ *and* $\overline{\Phi}(\varphi, r)$ *satisfy the following condition : for all fixed* $\varphi \in \mathbb{R}$,

$$\left(1 - \frac{n}{2}\right) \log\left(\int_0^\infty e^{\overline{\Phi} + (n+1)\,\overline{\tau}}\, dr\right) + \frac{n}{2} \log\left(\int_0^\infty e^{\overline{\Phi} + (n-1)\,\overline{\tau}}\, dr\right) = \varphi, \tag{121}$$

and define

$$\tau(\varphi) \stackrel{def.}{=} \frac{1}{2}\left(\log\int_0^\infty e^{\overline{\Phi}+(n+1)\,\overline{\tau}}\,dr - \log\int_0^\infty e^{\overline{\Phi}+(n-1)\,\overline{\tau}}\,dr\right). \tag{122}$$

Let now Φ be a function on M such that, for this choice of $\overline{\Phi}$, $\overline{\tau}$ and τ, conditions (i), (ii) *and* (iii) *of theorem 3.3.1 are verified for some $\overline{\rho} > 0$ and $m \geqslant n+1$. Then $\mathbf{L} = \Delta + \nabla\Phi$ satisfies the following* SOBOLEV *inequality of dimension m. For any function $f \in \mathcal{C}_c^\infty(M)$,*

$$\|f\|^2_{\frac{2m}{m-2},\mu} \leqslant \|f\|^2_{2,\mu} + \frac{4}{m\,(m-2)}\,\frac{m-1}{\overline{\rho}}\,\|\mathbf{\Gamma}(f)\|_{1,\mu}. \tag{123}$$

Remark 3.3.4. — *Under conditions* (121) *and* (122),

$$\int e^{\overline{\Phi}}\,d\overline{\sigma} = 1$$

provided that $\int e^{\Phi}\,d\sigma = 1$.

It seems too difficult to take advantage of the conditions of theorem 3.3.1 in general. So we have to particularize the choice of the shape functions $\overline{\tau}$, $\overline{\Phi}$ and τ.

3.3.3. A family of structure functions. — For fixed $\alpha, \beta \in \mathbb{R}$, consider the vector space of functions f on $\mathbb{R}\times\mathbb{R}_+^*$ which are homogeneous in the sense that, for all $(x,r) \in \mathbb{R}\times\mathbb{R}_+^*$,

$$f(x,r) = x^\alpha\, f(1, r/x^\beta).$$

Such a function is characterized by the one variable function $F(r) = f(1,r)$, and we have $f(x,r) = x^\alpha\, F(r/x^\beta)$.

As a consequence, fix a positive function F on $\mathbb{R}$. Define

$$\omega_{\lambda,\delta} \stackrel{\text{def.}}{=} \int_0^\infty F(r)^\lambda\, r^\delta\, dr.$$

Choose $\overline{\tau}$ and $\overline{\Phi}$ such that

$$\overline{\tau} = \overline{\tau}_0 + \alpha\,\varphi + G\big(r\,e^{-\beta\,\varphi}\big). \tag{124}$$

and

$$\overline{\Phi} = \overline{\Phi}_0 + \gamma\,\varphi + G_1\big(r\,e^{-\beta\,\varphi}\big). \tag{125}$$

Now, it is quite clear that a change of r into r^ν changes β into β/ν, such that the choice of the parameter β is not really important.

This form allows us to compute explicitely the projection conditions of corollary 3.3.3.

Unfortunately, the equation $\zeta = 0$ leads to a RICCATI equation in the derivative of the function $G_1(\log(r))$ with coefficients depending on the function G, which we have been unable to solve in general.

In what follows, we shall restrict our attention to the case where

$$\overline{\tau}(\varphi, r) = \overline{\tau}_0 - \log(\exp(-2\varphi/p) + r^2/2),$$

and

$$\overline{\Phi} = (p-1)(\overline{\tau} + \log(r)).$$

Then, we have $\tau = \tau_0 + \frac{2}{p}\varphi$, and the projection conditions are satisfied provided that

$$\exp(p\,\overline{\tau}_0) = \left(\frac{\omega}{\omega'}\right)^{\frac{n}{2}} \frac{1}{\omega}, \quad \tau_0 = \overline{\tau}_0 - \frac{1}{2}\log\left(\frac{\omega}{\omega'}\right),$$

where

$$\omega = \int_0^\infty \frac{r^{p-1}}{(1+r^2/2)^{n+p}}\,dr, \ \omega' = \int_0^\infty \frac{r^{p-1}}{(1+r^2/2)^{n+p-2}}\,dr,$$

and we have

$$\omega = 2^{-n-\frac{p}{2}} Z_{2n+p,p}, \ \omega' = 2^{2-n-\frac{p}{2}} Z_{2n+p-4,p},$$

where the constants $Z_{m,n}$ had been defined in (116).

Now, for these functions, we may explicitely compute the quantities $a(y,r)$ and $h(y,r)$ which characterize the fact that the structure $(\overline{g}, \overline{\mu})$ satisfy a $\mathrm{CD}(\overline{\rho}, n+p)$ inequality, following theorem 3.3.1.

It turns out that, in this case and up to positive multiplicative functions, the expression a has the form $A_1 r^2 + B_1$, where A_1 and B_1 are functions on M, and h has the form $C_1 + r^2 D_1 + E_1 r^4$, where C_1, D_1 and E_1 are symmetric tensors. If we write down the conditions that these two functions and these three tensors are positive, which is enough to ensure the positivity of the quantities a and h on the product space, we end up with the following five conditions, written in terms of the function $\Psi = \frac{2}{p}\Phi$. We then obtain the following

Theorem 3.3.5. — *Assume that the following conditions holds for the function $\Psi = \frac{2}{p}\Phi$ on M :*

$$\begin{aligned}
&-\Delta\Psi + (n-1)\,\mathbf{\Gamma}(\Psi,\Psi) - n\,e^{-2\tau_0-\Psi} \geqslant 0,\\
&-\Delta\Psi - p\mathbf{\Gamma}(\Psi,\Psi) + (n+2p-2)e^{-2\tau_0-\Psi} - e^{2(\overline{\tau}_0-\tau_0)}\,\overline{\rho} \geqslant 0,\\
&\mathrm{Ric} + \Delta(\Psi)\,g + (n-2)\left\{\nabla\nabla\Psi - \mathbf{\Gamma}(\Psi,\Psi)\,g + \nabla\Psi\otimes\nabla\Psi\right\} \geqslant 0\\
&2\,\mathrm{Ric} + (n-p-2)\{\nabla\nabla\Psi + \nabla\Psi\otimes\nabla\Psi\}\\
&\qquad\qquad + \left\{\Delta(\Psi) + (p+1)\,\mathbf{\Gamma}(\Psi) - (2n+p-2)\,e^{-2\tau_0-\Psi}\right\} g \geqslant 0\\
&\mathrm{Ric} - p\left\{\nabla\nabla\Psi + \nabla\Psi\otimes\nabla\Psi\right\} + \left\{p\,e^{-2\tau_0-\Psi} - e^{2(\overline{\tau}_0-\tau_0)}\,\overline{\rho}\right\} g \geqslant 0.
\end{aligned}$$

Then, with the functions $\overline{\tau}$, τ and $\overline{\Phi}$ defined as above, the pair

$$\left(\overline{g} = \exp(2\overline{\tau})\,(\exp(2\tau)g \oplus 1), \exp(\overline{\Phi})d\overline{\sigma}\right)$$

satisfies a $\mathrm{CD}(\overline{\rho}, m = n+p)$ inequality and projects onto $(g, \exp(\Phi)d\sigma)$, where $d\sigma$ and $d\overline{\sigma}$ are the Riemannian measures associated with g and $\overline{g}$ respectively.

Therefore, if $\overline{\rho} > 0$, then the pair $(g, \exp(\Phi)d\sigma)$ satisfies a SOBOLEV *inequality with coefficients $A = 1$ and $B = \frac{4}{m(m-2)}\frac{m-1}{\overline{\rho}}$.*

Let us notice that in this example $\tau = \tau_0 + \Psi$, and that the third inequality is just the fact that the metric $\exp(2\tau)g$ has a positive RICCI curvature.

In the case of the previous example, if the starting metric is the metric of the sphere, the metric obtained by this method is just the Euclidean metric.

3.3.4. Weighted SOBOLEV inequality in Euclidean spaces. — In the previous computations, we saw that the (co)metric $\check{g}$ transformed from the starting (co)metric by the conformal rate $\exp(2\tau)$ had positive RICCI curvature. Therefore, it would be interesting to start from a flat metric (or at least a metric with 0 RICCI tensor), and to see what kind of weighted SOBOLEV inequalities we could obtain in this case. To make things simpler, we just address the case where the starting space is the usual Euclidean space $\mathbb{R}^n$. In this case, $\check{g}$ is the Euclidean metric and, up to a constant, the measure μ is $\exp((n + p/2)\tau)dx$. We are going here to explore for which functions τ we may expect to obtain by this method a SOBOLEV inequality for the pair $(\exp(-2\tau)\,\check{g}, \exp((n + p/2)\tau)dx/Z)$.

In this particular case, the five conditions are reduced to three : in fact, the third equation is satisfied by definition, and if we take the trace in the fourth inequality, and combine with the result obtained by taking the trace of the third one, then we obtain the first inequality. This simplification would have been the same under the condition that the scalar curvature of the (co)metric $\exp(2\tau)g$ is 0.

It appears then that the computations are easier from the point of view of the flat metric. This explains why the following statements are made in the Euclidean case.

Proposition 3.3.6. — *Suppose $(M, \check{g})$ is a riemannian manifold of dimension n with flat* RICCI *curvature and let τ be a function on M such that*

$$\left\{-\check{\Delta}\tau + \check{\mathbf{\Gamma}}(\tau) - (2\,n + p - 2)\,e^{\tau}\right\}\check{g} + (n + p - 2)\left\{\check{\nabla}\tau \otimes \check{\nabla}\tau - \check{\nabla}\check{\nabla}\tau\right\} \geqslant 0 \tag{126}$$

$$\begin{aligned}\left\{-\check{\Delta}\tau - (n + p - 2)\check{\mathbf{\Gamma}}(\tau) + p\,e^{\tau} - \rho'\,e^{2\,\tau}\right\}\check{g} \\ + (n + p - 2)\left\{\check{\nabla}\tau \otimes \check{\nabla}\tau - \check{\nabla}\check{\nabla}\tau\right\} \geqslant 0\end{aligned} \tag{127}$$

$$-\check{\Delta}\tau - (n + p - 2)\check{\mathbf{\Gamma}}(\tau) + (n + 2\,p - 2)\,e^{\tau} - \rho'\,e^{2\,\tau} \geqslant 0, \tag{128}$$

for some $p > 1$ and $\rho' > 0$. Then the probability measure $d\mu = e^{\left(n+\frac{p}{2}\right)\tau}\,d\check{\sigma}/Z$ satisfies the following weighted SOBOLEV *inequality of dimension $m = n + p$. For any $f \in \mathcal{C}_c^{\infty}(M)$,*

$$\|f\|^2_{\frac{2m}{m-2},\mu} \leqslant \|f\|^2_{2,\mu} + \frac{16\,(m-1)}{m\,(m+n-2)\,(m+n-4)}\,\frac{m-1}{\rho'}\,\|e^{-2\,\tau}\,\check{\mathbf{\Gamma}}(f)\|_{1,\mu}.$$

Proof

◀ This is exactly writing $D_1, C_1, B_1 \geqslant 0$ in terms of $\check{g}$ and remarking that $A_1 \geqslant 0$

is thus given by $A_1 = (2\,n + p - 2)\,\mathrm{Tr}(E_1)$. ►

We are going now to search which radial functions in the Euclidean space $\mathbb{R}^n$ satisfy the conditions of proposition 3.3.6. Such a function can be written $\tau = \tau(v)$ where the function

$$v = \frac{1 - |y|^2}{1 + |y|^2}$$

satisfies

$$\begin{aligned}
&\check{\nabla}\check{\nabla} v = -(1+v)^2\,\check{g} + 2\,(1+v)^{-1}\check{\nabla} v \otimes \check{\nabla} v \\
&\check{\mathbf{\Gamma}}(v) = (1+v)^3(1-v) \\
&\check{\Delta} v = -(1+v)^2\,(n - 2\,(1-v)),
\end{aligned}$$

in the Euclidean metric. Hence, the quadratic forms (126) and (127) turn out to be linear combinations of $\check{g}$ and $\check{\nabla} v \otimes \check{\nabla} v$. Now, the eigenvalues of $\alpha\,\check{g} + \beta\,\check{\nabla} v \otimes \check{\nabla} v$ are easy to determine : these are α and $\alpha + \beta\,\check{\mathbf{\Gamma}}(v)$.

We finally obtain the five necessary and sufficient conditions (to be verified simultaneously) for the function τ of the variable $v \in]-1, 1[$ to satisfy the conditions of proposition 3.3.6 :

$$\begin{aligned}
&-(1-v^2)\big\{\tau'' - \tau'^2\big\} - 2\,(1-v)\,\tau' + (2\,n + p - 2)\,\big\{\tau' - (1+v)^{-2}e^{\tau}\big\} \geqslant 0, \\
&(n+p-1)\,\big[-(1-v^2)\big\{\tau'' - \tau'^2\big\} - 2\,(1-v)\,\tau'\big] \\
&\qquad\qquad + (2\,n + p - 2)\,\big\{\tau' - (1+v)^{-2}e^{\tau}\big\} \geqslant 0, \\
&-(1-v^2)\,\big\{\tau'' + (n+p-2)\tau'^2\big\} + \{(2\,n + p - 2) - 2\,(1-v)\}\,\tau' \\
&\qquad\qquad + (p - \rho'\,e^{\tau})\,(1+v)^{-2}\,e^{\tau} \geqslant 0, \\
&-(n+p-1)\,(1-v^2)\,\tau'' + \{(2\,n + p - 2) - 2\,(n+p-1)(1-v)\}\,\tau' \\
&\qquad\qquad + (p - \rho'\,e^{\tau})\,(1+v)^{-2}\,e^{\tau} \geqslant 0, \\
&-(1-v^2)\,\big\{\tau'' + (n+p-2)\tau'^2\big\} + \{n - 2\,(1-v)\}\,\tau' \\
&\qquad\qquad + \{(n + 2\,p - 2) - \rho'\,e^{\tau}\}\,(1+v)^{-2}\,e^{\tau} \geqslant 0.
\end{aligned}$$

Eventually, the latter are given by the three simpler ones

$$\tau' - (1+v)^{-2}\,e^{\tau} \geqslant 0 \tag{129}$$

$$-\tau'' + \tau'^2 - 2\,(1+v)^{-1}\tau' \geqslant 0 \tag{130}$$

$$-(1-v^2)\,\tau'^2 + (2 - \lambda\,e^{\tau})(1+v)^{-2}\,e^{\tau} \geqslant 0 \tag{131}$$

where we posed $0 < \lambda = \rho'/(n+p-1)$.

Remark 3.3.7. — *The equality case in* (129), (130) *and* (131) *is given by the function*

$$\tau(v) = \log\left(\frac{2}{\lambda + r^2}\right)$$

for $r^2 \stackrel{def.}{=} \frac{1-v}{1+v}$. *Moreover, by coupling* (129) *and* (131), *one easily obtains that*

$$e^{\tau} \leqslant \frac{2}{\lambda + r^2}.$$

Let's consider again the example we were studying in section 3.2 on $\mathbb{R}^n$. Recall that

$$\mathbf{L}_{m,n} = \Delta_{S^n} + \nabla\Phi$$

with

$$\Phi = \frac{p}{2}\log(1+v).$$

The number $p = m - n$ was a positive integer. We can define the operator $\mathbf{L}_{m,n}$ as well if $p \geqslant 1$ is a real number instead of an integer. The preceding approach allows to extend the SOBOLEV inequality obtained for $p \in \mathbb{N}^*$ to all values of p. If we pose $\tau \stackrel{\text{def.}}{=} \log(1+v)$, then the measure μ introduced in proposition (3.3.6) is the reversible measure $\sigma_{m,n}$ associated to $\mathbf{L}_{m,n}$. Moreover, the metric g associated to $\mathbf{L}_{m,n}$ is the spherical metric and $\exp(2\tau)g$ is the Euclidean one. Finally, by definition, the function

$$\tau = \log\left(\frac{2}{1+r^2}\right)$$

is the equality case in (129), (130) and (131) if we pose $\rho' = n + p - 1 = m - 1$. Hence, proposition (3.3.6) applies and we get the expected SOBOLEV inequality for $\sigma_{m,n}$

$$\|f\|^2_{\frac{2m}{m-2},\sigma_{m,n}} \leqslant \|f\|^2_{2,\sigma_{m,n}} + B_{m,n}\|e^{-2\tau}|\nabla f|^2\|_{1,\sigma_{m,n}}$$

where, as before,

$$B_{m,n} = \frac{16(m-1)}{m(m+n-2)(m+n-4)}.$$

This generalizes the previous result for integer values of $m \geqslant n+1$ to all real values greater than $n+1$.

3.4. Proof of the main result

We are going to devote this section to the proof of the fundamental theorem 3.3.1.

3.4.1. Some preliminaries on conformal transformations. — Let (M, g) be a Riemannian manifold. We now briefly expose what kind of changes the geometrical objects undergo when the underlying metrics is replaced by a new one, say $\check{g}$, supposed to be conformal to g. These well known results can be found in a large literature. We refer for instance to [Heb97].

The geometrical objects the expression of which we will need subsequently are Hessian tensor, Laplacian and RICCI curvature. The formulas given above make use of the *exponential conformal rate* $\check{\tau}$ defined by $\check{g}_{..} = e^{2\check{\tau}}\, g_{..}$, where $\check{\tau}$ is a given smooth function on M. If f is a $\mathcal{C}^\infty(M)$ function, we then have

Hessian tensor

$$\check{\nabla}^{\cdot}\check{\nabla}^{\cdot} f = e^{-4\check{\tau}}[\nabla^{\cdot}\nabla^{\cdot} f - 2\,\nabla^{\cdot} f \odot \nabla^{\cdot}\check{\tau} + \mathbf{\Gamma}(f,\check{\tau})\, g^{\cdot\cdot}] \tag{132}$$

In other terms, the action on the 1-form dh gives

$$\check{\nabla}^{\cdot}\check{\nabla}^{\cdot} f(dh,dh) = e^{-4\check{\tau}}[\nabla^{\cdot}\nabla^{\cdot} f(dh,dh) - 2\,\mathbf{\Gamma}(f,h)\mathbf{\Gamma}(\check{\tau},h) + \mathbf{\Gamma}(f,\check{\tau})\,\mathbf{\Gamma}(h,h)]. \tag{133}$$

Laplace-Beltrami operator

$$\check{\Delta} f = e^{-2\check{\tau}}\,[\Delta f + (n-2)\,\mathbf{\Gamma}(f,\check{\tau})] \tag{134}$$

Ricci tensor

$$\check{\mathrm{Ric}}^{\cdot\cdot} = e^{-4\check{\tau}}[\mathrm{Ric}^{\cdot\cdot} - \Delta(\check{\tau})\, g^{\cdot\cdot} + (n-2)\{\nabla^{\cdot}\check{\tau}\otimes\nabla^{\cdot}\check{\tau} - \mathbf{\Gamma}(\check{\tau})\, g^{\cdot\cdot} - \nabla^{\cdot}\nabla^{\cdot}\check{\tau}\}]. \tag{135}$$

In equation (132), the notation $\nabla^{\cdot} f \odot \nabla^{\cdot}\check{\tau}$ stands for the symmetrized tensor product, i.e.

$$\nabla^{\cdot} f \odot \nabla^{\cdot}\check{\tau} \overset{\text{def.}}{=} \frac{1}{2}(\nabla^{\cdot} f \otimes \nabla^{\cdot}\check{\tau} + \nabla^{\cdot}\check{\tau}\otimes\nabla^{\cdot} f).$$

Remark 3.4.1. — *From now on, we use the notation $g_{\cdot\cdot}$ for the metric tensor acting on vector fields, i.e. in the sections of the fiber bundle $\otimes_2^0(M)$, and $g^{\cdot\cdot}$ for the dual one, in the sections of $\otimes_0^2(M)$, acting on 1-forms, i.e. the ones we used up to now. Recall that, in coordinates, the matrix expression of $g^{\cdot\cdot}$ is nothing but the inverse of $g_{\cdot\cdot}$. This notation will be extended to all tensors, the upper points representing an action on 1-forms, the under ones an action on vector fields. The metric isomorphism between tangent $\mathbf{T}M$ and cotangent $\mathbf{T}^*M$ gives rise to the symmetric operations of taking the index up and down.*

3.4.2. Ricci curvature of the lifted operator. — Recall that the curvature-dimension inequality $\mathrm{CD}(\overline{\rho}, n+p)$ for $\overline{\mathbf{L}}$ can be rewritten as

$$(p-1)\left(\overline{\mathrm{Ric}} - \overline{\nabla\nabla\Phi} - \overline{\rho}\,\overline{g}\right) - \overline{\nabla\Phi}\otimes\overline{\nabla\Phi} \geqslant 0.$$

Hence, we will first of all compute the Ricci curvature of the lifted operator. This computation will follow the three steps of the construction of $\overline{g}$.

3.4.2.1. *From g to $\check{g}$.* — Recall that $\check{g}_{\cdot\cdot} = e^{2\check{\tau}} g_{\cdot\cdot}$ with $\check{\tau} = -\tau$. Hence, by equation (135) conjugate to the elementary following remarks

$$\Delta\tau = \partial_{\varphi}\tau\,\Delta\Phi + \partial^2_{\varphi\varphi}\tau\,\mathbf{\Gamma}(\Phi), \tag{136}$$

$$\mathbf{\Gamma}(\tau) = \left(\partial_{\varphi}\tau\right)^2\mathbf{\Gamma}(\Phi), \tag{137}$$

$$\nabla\nabla\tau = \partial_{\varphi}\tau\,\nabla\nabla\Phi + \partial^2_{\varphi\varphi}\tau\,\nabla\Phi\otimes\nabla\Phi, \tag{138}$$

$$\nabla\tau\otimes\nabla\tau = \left(\partial_{\varphi}\tau\right)^2\nabla\Phi\otimes\nabla\Phi, \tag{139}$$

we obtain

$$\check{\mathrm{Ric}} = e^{4\tau}\Big[\mathrm{Ric} + \Big\{\partial_\varphi\tau\,\Delta\Phi + \Big(\partial^2_{\varphi\varphi}\tau - (n-2)\,\big(\partial_\varphi\tau\big)^2\Big)\,\mathbf{\Gamma}(\Phi)\Big\}\,g\Big] + (n-2)\,e^{4\tau}\Big[\partial_\varphi\tau\,\nabla\nabla\Phi + \Big(\partial^2_{\varphi\varphi}\tau + \big(\partial_\varphi\tau\big)^2\Big)\,\nabla\Phi\otimes\nabla\Phi\Big] \tag{140}$$

From now on, we will omit the two upper dots.

3.4.2.2. *The tensorization step.* — This is easy to see that

$$\hat{\mathrm{Ric}} = \left(\begin{array}{c|c} \check{\mathrm{Ric}} & 0 \\ \hline 0 & 0 \end{array}\right).$$

3.4.2.3. *Second conformal transformation.* — This step requires a little more attention. By (135), we still have

$$\overline{\mathrm{Ric}} = e^{-4\overline{\tau}}\Big[\hat{\mathrm{Ric}} - \hat{\Delta}(\overline{\tau})\,\hat{g} + (n-2)\Big\{\hat{\nabla}\overline{\tau}\otimes\hat{\nabla}\overline{\tau} - \hat{\mathbf{\Gamma}}(\overline{\tau})\,\hat{g} - \hat{\nabla}\hat{\nabla}\overline{\tau}\Big\}\Big]. \tag{141}$$

We then have to compute the different terms of this sum. Regarding first order terms, we can use

$$\hat{\nabla}_y\overline{\tau} = e^{2\tau}\,\nabla\overline{\tau} = e^{2\tau}\,\partial_\varphi\overline{\tau}\,\nabla\Phi$$
$$\text{and}\qquad \hat{\nabla}_r\overline{\tau} = \partial_r\overline{\tau},$$

to conclude that

$$\hat{\nabla}\overline{\tau}\otimes\hat{\nabla}\overline{\tau} = \left(\begin{array}{c|c} e^{4\tau}\,\big(\partial_\varphi\overline{\tau}\big)^2\,\nabla\Phi\otimes\nabla\Phi & e^{2\tau}\,\partial_\varphi\overline{\tau}\,\partial_r\overline{\tau}\,\nabla\Phi \\ \hline e^{2\tau}\,\partial_\varphi\overline{\tau}\,\partial_r\overline{\tau}\;{}^{\mathrm{t}}\nabla\Phi & \big(\partial_r\overline{\tau}\big)^2 \end{array}\right) \tag{142}$$

and

$$\hat{\mathbf{\Gamma}}(\overline{\tau}) = e^{2\tau}\,\big(\partial_\varphi\overline{\tau}\big)^2\,\mathbf{\Gamma}(\Phi) + (\partial_r\overline{\tau})^2. \tag{143}$$

Let's now look into the second order terms. We obtain $\hat{\nabla}_y\hat{\nabla}_y\overline{\tau}$ by conformal change from g to $\check{g}$ and using (132). Hence, after easy computation like (136) to (139), we can write

$$\hat{\nabla}_y\hat{\nabla}_y\overline{\tau} = e^{4\tau}\big[\partial_\varphi\overline{\tau}\,\nabla\nabla\Phi - \partial_\varphi\tau\,\partial_\varphi\overline{\tau}\,\mathbf{\Gamma}(\Phi)\,g + \big(\partial^2_{\varphi\varphi}\overline{\tau} + 2\,\partial_\varphi\tau\,\partial_\varphi\overline{\tau}\big)\,\nabla\Phi\otimes\nabla\Phi\big]. \tag{144}$$

Moreover,

$$\hat{\nabla}_y\hat{\nabla}_r\overline{\tau} = e^{2\tau}\partial^2_{\varphi r}\overline{\tau}\nabla\Phi \tag{145}$$

and

$$\hat{\nabla}_r\hat{\nabla}_r\overline{\tau} = \partial^2_{rr}\overline{\tau}. \tag{146}$$

Taking the trace of this Hessian w.r.t. the metrics

$$\hat{g}_{..} = \left(\begin{array}{c|c} e^{-2\tau}\, g_{..} & 0 \\ \hline 0 & 1 \end{array}\right),$$

one gets the hat Laplacian

$$\hat{\Delta}\overline{\tau} = e^{2\tau}\left[\partial_\varphi \overline{\tau}\,\Delta\Phi + \left(\partial^2_{\varphi\varphi}\overline{\tau} - (n-2)\,\partial_\varphi\tau\,\partial_\varphi\overline{\tau}\right)\mathbf{\Gamma}(\Phi) + e^{-2\tau}\,\partial^2_{rr}\overline{\tau}\right]. \tag{147}$$

Finally, using (141) and the previous computations, we find the expected (awful) expression of the $\overline{\mathrm{Ric}}$ tensor. On the one hand,

$$e^{4(\overline{\tau}-\tau)}\overline{\mathrm{Ric}}_{yy} = \mathrm{Ric} + \overline{\nu}(y,r)\,g + \overline{\pi}(y,r)\,\nabla\nabla\Phi + \overline{\chi}(y,r)\,\nabla\Phi\otimes\nabla\Phi \tag{148}$$

with first

$$\begin{aligned} \overline{\nu} &= \partial_\varphi(\tau-\overline{\tau})\,\Delta\Phi \\ &+ \left\{\partial^2_{\varphi\varphi}(\tau-\overline{\tau}) - (n-2)\left(\partial_\varphi\tau\right)^2 + (2\,n-3)\,\partial_\varphi\tau\,\partial_\varphi\overline{\tau} - (n-1)\left(\partial_\varphi\overline{\tau}\right)^2\right\}\mathbf{\Gamma}(\Phi) \\ &\qquad - e^{-2\tau}\left(\partial^2_{rr}\overline{\tau} + (n-1)\left(\partial_r\overline{\tau}\right)^2\right), \end{aligned} \tag{149}$$

secondly

$$\overline{\pi} = (n-2)\,\partial_\varphi\tau - (n-1)\partial_\varphi\overline{\tau} \tag{150}$$

and thirdly

$$\overline{\chi} = (n-2)\left\{\partial^2_{\varphi\varphi}\tau + \left(\partial_\varphi\tau\right)^2\right\} - (n-1)\left\{\partial^2_{\varphi\varphi}\overline{\tau} - \left(\partial_\varphi\overline{\tau}\right)^2\right\} - 2\,(n-1)\,\partial_\varphi\tau\,\partial_\varphi\overline{\tau}. \tag{151}$$

On the other hand,

$$e^{4\overline{\tau}-2\tau}\;\overline{\mathrm{Ric}}_{yr} = (n-1)\left\{\partial_r\overline{\tau}\,\partial_\varphi\overline{\tau} - \partial^2_{\varphi r}\overline{\tau}\right\}\nabla\Phi \tag{152}$$

and

$$\begin{aligned} e^{4\overline{\tau}-2\tau}\;\overline{\mathrm{Ric}}_{rr} &= -\partial_\varphi\overline{\tau}\,\Delta\Phi \\ &+ \left\{-\partial^2_{\varphi\varphi}\overline{\tau} + (n-2)\,\partial_\varphi\tau\,\partial_\varphi\overline{\tau} - (n-1)\left(\partial_\varphi\overline{\tau}\right)^2\right\}\mathbf{\Gamma}(\Phi) - n\,e^{-2\tau}\,\partial^2_{rr}\overline{\tau}. \end{aligned} \tag{153}$$

3.4.3. Curvature-dimension criterion for the lifted operator. — We now will finally determine the conditions to be verified in order that the operator $\overline{\mathbf{L}}$ satisfies a $\mathrm{CD}(\overline{\rho}, n+p)$ inequality. As we have seen before, this inequality requires the twice contravariant tensor

$$\overline{A} \stackrel{\text{def.}}{=} (p-1)\left(\overline{\mathrm{Ric}} - \overline{\nabla\nabla\Phi} - \overline{\rho}\,\overline{g}\right) - \overline{\nabla\Phi}\otimes\overline{\nabla\Phi} \tag{154}$$

to be positive. To determine $\overline{\mathrm{Ric}}$ was the aim of the previous section. We now have to compute the other terms of (154).

By (132), we can write

$$\overline{\nabla\nabla\Phi} = e^{-4\,\overline{\tau}}\Big[\hat{\nabla}\hat{\nabla}\overline{\Phi} - 2\,\hat{\nabla}\overline{\Phi}\odot\hat{\nabla}\overline{\tau} + \hat{\mathbf{\Gamma}}\big(\overline{\Phi},\overline{\tau}\big)\,\hat{g}\Big].$$

Thanks to the computations previously done for $\overline{\tau}$ the reader can easily adapt for $\overline{\Phi}$, we have

$$\begin{aligned}
&\hat{\nabla}_y\hat{\nabla}_y\overline{\Phi} = e^{4\,\tau}\big[\partial_\varphi\overline{\Phi}\,\nabla\nabla\Phi - \partial_\varphi\tau\,\partial_\varphi\overline{\Phi}\,\mathbf{\Gamma}(\Phi)\,g + \big(\partial^2_{\varphi\varphi}\overline{\Phi} + 2\,\partial_\varphi\tau\,\partial_\varphi\overline{\Phi}\big)\,\nabla\Phi\otimes\nabla\Phi\big],\\
&\hat{\nabla}_y\overline{\Phi}\odot\hat{\nabla}_y\overline{\tau} = e^{4\,\tau}\,\partial_\varphi\overline{\Phi}\,\partial_\varphi\overline{\tau}\,\nabla\Phi\otimes\nabla\Phi,\\
&\hat{\mathbf{\Gamma}}\big(\overline{\Phi},\overline{\tau}\big)\,\hat{g}_{yy} = e^{2\,\tau}\,\big(e^{2\,\tau}\,\partial_\varphi\overline{\Phi}\,\partial_\varphi\overline{\tau}\,\mathbf{\Gamma}(\Phi) + \partial_r\overline{\Phi}\,\partial_r\overline{\tau}\big)\,g
\end{aligned}$$

and consequently

$$\begin{aligned}
e^{4\,(\overline{\tau}-\tau)}\,\overline{\nabla}_y\overline{\nabla}_y\overline{\Phi} = \partial_\varphi\overline{\Phi}\nabla\nabla\Phi + \big\{\partial_\varphi\overline{\Phi}\,\big(\partial_\varphi\overline{\tau} - \partial_\varphi\tau\big)\,\mathbf{\Gamma}(\Phi) + e^{-2\,\tau}\partial_r\overline{\Phi}\,\partial_r\overline{\tau}\big\}\,g\\
+ \big\{\partial^2_{\varphi\varphi}\overline{\Phi} - 2\,\partial_\varphi\overline{\Phi}\,\big(\partial_\varphi\overline{\tau} - \partial_\varphi\tau\big)\big\}\nabla\Phi\otimes\nabla\Phi.
\end{aligned}\tag{155}$$

Moreover, one can check that

$$\overline{\nabla}_r\overline{\nabla}_y\overline{\Phi} = e^{2\,\tau-4>\overline{\tau}}\,\big(\partial^2_{\varphi r}\overline{\Phi} - \partial_\varphi\overline{\Phi}\,\partial_r\overline{\tau} - \partial_r\overline{\Phi}\,\partial_\varphi\overline{\tau}\big)\,\nabla\Phi\tag{156}$$

and

$$\overline{\nabla}_r\overline{\nabla}_r\overline{\Phi} = e^{2\,\tau-4\,\overline{\tau}}\,\big\{\partial_\varphi\overline{\Phi}\,\partial_\varphi\overline{\tau}\,\mathbf{\Gamma}(\Phi) + e^{-2\,\tau}\,\big(\partial^2_{rr}\overline{\Phi} - \partial_r\overline{\Phi}\,\partial_r\overline{\tau}\big)\big\}.\tag{157}$$

The last term of (154) is, as like as (142),

$$\overline{\nabla\Phi}\otimes\overline{\nabla\Phi} = e^{4\,(\tau-\overline{\tau})}\left(\begin{array}{c|c}
\big(\partial_\varphi\overline{\Phi}\big)^2\,\nabla\Phi\otimes\nabla\Phi & e^{-2\,\tau}\,\partial_\varphi\overline{\Phi}\,\partial_r\overline{\Phi}\,\nabla\Phi\\
\hline
e^{-2\,\tau}\,\partial_\varphi\overline{\Phi}\,\partial_r\overline{\Phi}\;{}^{\mathrm{t}}\nabla\Phi & e^{-4\,\tau}\,\big(\partial_r\overline{\Phi}\big)^2
\end{array}\right).\tag{158}$$

Now, $\overline{A}\in\bigotimes^2_0\big(\overline{M}\big)$ is of the form

$$\overline{A} = \left(\begin{array}{c|c}
A & X\\
\hline
{}^{\mathrm{t}}\mathrm{X} & \tilde{a}
\end{array}\right)$$

where $A\in\bigotimes^2_0(M)$, $X\in\mathbf{T}M$ and $\tilde{a}\in\mathcal{C}^\infty(M)$. Hence, the condition $\overline{A}\geqslant 0$ is equivalent to the following conditions. First, $\tilde{a}$ and A must be non negative. And secondly,

i. every zero point y of $\tilde{a}$ (noted $y\in\mathcal{Z}(\tilde{a})$) is a singular point of X,

ii. If $y\notin\mathcal{Z}(\tilde{a})$, then

$$X\otimes X\leqslant\tilde{a}\,A,$$

as bilinear forms in $\bigotimes^2_0(M)$.

The motivated reader should verify the explicit values of A, X and $\tilde{a}$ are those given below. The values of the different used functions are those specified before theorem 3.3.1.

The matrix

The matrix A is given by

$$\frac{e^{4\,(\overline{\tau}-\tau)}}{p-1}\,A = \mathrm{Ric} + \xi(\cdot,r)\,g + \kappa(\Phi,r)\,\nabla\nabla\Phi + \eta(\Phi,r)\,\nabla\Phi\otimes\nabla\Phi$$

with

$$\xi(\cdot,r) = \nu_2(\Phi,r)\,\Delta\Phi + \pi_2(\Phi,r)\,\mathbf{\Gamma}(\Phi) + \chi_2(\Phi,r).$$

The vector field

$$\frac{e^{4\,\overline{\tau}-2\,\tau}}{p-1}\,X = \zeta(\Phi,r)\,\nabla\Phi.$$

The function

$$\frac{e^{4\,\overline{\tau}-2\,\tau}}{p-1}\,\tilde{a} = \nu_1(\Phi,r)\,\Delta\Phi + \pi_1(\Phi,r)\,\mathbf{\Gamma}(\Phi) + \chi_1(\Phi,r).$$

We hence find the conditions given in theorem 3.3.1.

ANNEXE A

CONCENTRATION DE LA MESURE ET INÉGALITÉ DE SOBOLEV LOGARITHMIQUE

Cet appendice est tiré du livre [ABC$^+$00] sur les inégalités de SOBOLEV logarithmique, écrit en collaboration avec sept autres doctorants du laboratoire de Statistique et Probabilités de l'Université Paul SABATIER. Le sortir de son contexte présente l'inconvénient de rendre la lecture parfois plus difficile. De nombreuses références sont en effet internes au livre. Lorsque nous renverrons au *livre*, il sera toujours question de [ABC$^+$00]. Nous espérons néanmoins que cet inconvénient n'est que mineur.

A.1. Introduction

Pour une mesure de probabilité sur la tribu borélienne d'un espace métrique, la masse du r-voisinage A_r de tout ensemble A converge vers 1 quand r tend vers $+\infty$. Pour certaines mesures, cette convergence a lieu à une vitesse très rapide. Plus précisément, partant d'une partie A quelconque de masse $1/2$, un faible grossissement de A (au sens de la métrique) emplit presque tout l'espace (au sens de la mesure), c'est-à-dire est de masse proche de 1. Ce phénomène est appelé concentration de la mesure. En particulier, la notion de concentration gaussienne concerne les mesures de probabilité pour lesquelles cette convergence est plus rapide que dans un modèle gaussien.

A l'origine, on établissait des propriétés de concentration comme une conséquence immédiate d'inégalités isopérimétriques (voir [GM83, LT91]). Nous présentons notamment ici la valeur exacte de la fonction de concentration gaussienne à l'aide de l'inégalité isopérimétrique correspondante. Ce type d'inégalités consiste, pour fixer les idées, en une majoration de la mesure d'un ensemble A par celle de son bord (ou de manière équivalente par celle de son r-voisinage A_r). Pour la concentration, ce contrôle n'est exigé que pour r grand. Cette propriété peut ainsi être vérifiée même en l'absence d'inégalité isopérimétrique.

TALAGRAND a récemment obtenu des inégalités de concentration dans des espaces produits, dont une application importante en statistique est l'obtention d'intervalles de confiance très précis pour le supremum (sur une classe de fonctions) de moyennes

empiriques (voir [Tal96a]). Une des difficultés techniques pour appréhender ces inégalités est qu'en général la propriété de concentration ne se tensorise pas. Un outil très utile pour pallier cet inconvénient est alors de faire appel à une inégalité de SOBOLEV logarithmique comme intermédiaire (voir [Led99]). Cette inégalité engendre en effet une concentration gaussienne (argument de HERBST) et se conserve par passage au produit (voir le chapitre 3 du livre [ABC+00]). Le lien entre la concentration gaussienne et l'existence d'une inégalité de SOBOLEV logarithmique est en fait beaucoup plus fort, et l'un des objectifs de ce chapitre est de décrire précisément les relations existant entre inégalités de SOBOLEV logarithmiques, concentration gaussienne et intégrabilité exponentielle du carré des fonctions lipschitziennes.

Nous introduisons dans un premier temps la notion de concentration gaussienne en insistant sur sa véritable signification : la décroissance au moins gaussienne de la masse des queues des lois des fonctions lipschitziennes (la queue d'une loi de probabilité ν sur $\mathbb{R}$ est la fonction $r \mapsto \nu(|\cdot| \geqslant r)$). Nous retrouvons la forme de la fonction de concentration de la loi gaussienne par une méthode de semi-groupe basée sur le critère de courbure qui ne fait pas appel à l'inégalité isopérimétrique. Puis, nous présentons en détail l'argument de HERBST dans les cadres continu et discret. Cette méthode constitue la clé de voûte de l'utilisation des inégalités de SOBOLEV logarithmiques dans l'étude de la concentration. Enfin, nous exposons le théorème de WANG qui peut être vu comme l'« inverse » de l'argument de HERBST en courbure minorée.

A.2. La concentration de la mesure

Nous exposons dans cette section les bases concernant le phénomène de concentration de la mesure. Dans un premier temps, nous justifions la définition abstraite de la concentration en nous appuyant sur l'exemple des mesures gaussiennes et nous en profitons pour introduire la propriété de concentration gaussienne d'une mesure. Nous donnons alors une formulation fonctionnelle de la concentration soulignant l'importance des fonctions lipschitziennes et de leurs médianes dans l'étude de cette propriété. Nous abordons ensuite une discussion permettant d'expliquer une propriété déroutante au premier abord : une inégalité de concentration gaussienne autour d'*un* point, pour la loi d'une fonction lipschitzienne, engendre une inégalité de même type autour de *tout autre* point. Enfin, nous déduisons d'une telle inégalité l'intégrabilité exponentielle du carré de la fonction considérée.

A.2.1. Exemple des mesures gaussiennes. — Nous introduirons dans la section A.2.2 la propriété de concentration pour une mesure de probabilité sur les boréliens d'un espace métrique. L'idée sous-jacente est simple bien que les définitions présentées dans un cadre abstrait demandent un certain formalisme. Aussi nous attachons-nous tout d'abord à l'exemple des mesures gaussiennes afin de bien faire comprendre le problème.

Sur l'espace euclidien $\mathbb{R}^2$ (par exemple), on considère la mesure gaussienne

$$\mu(dx, dy) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dxdy.$$

L'objectif que nous poursuivons est d'étudier la *décroissance de la masse à l'infini.* La fonction la plus simple exprimant cette décroissance est

$$\alpha(r) = \mu(B^c(0, r)) = e^{-r^2/2},$$

où, d'une part, $B(0, r)$ est la boule centrée à l'origine et de rayon r et, d'autre part, A^c désigne le complémentaire d'un ensemble A. Aussi la masse est-elle principalement *concentrée* autour de 0. Dans le cadre général, cette décroissance sera encore donnée par une fonction exprimant la masse du complémentaire du r-voisinage d'un ensemble fixé A.

On peut de plus imposer que la fonction $\alpha(r)$ donne la décroissance minimale sur toutes les parties A considérées. Un exemple simple montre qu'il faut alors imposer aux parties A d'avoir une masse suffisante. En effet, si γ désigne la mesure gaussienne standard sur $\mathbb{R}$, pour tout r positif,

$$\sup\{\gamma((A_r)^c); A \text{ borélien}\} = 1$$

(prendre $A_x =]-\infty, x]$, quand x tend vers $-\infty$). Nous imposerons donc aux parties considérées d'avoir une masse au moins $1/2$.

Remarque A.2.1. — *Considérer des parties de masse plus grande que ϵ au lieu de $1/2$, pour un $0 < \epsilon < 1$, donnerait lieu comme dans la suite à des inégalités de déviation mais conduirait à des détails techniques que nous nous abstiendrons de présenter ici.*

A.2.2. La fonction de concentration. — Soit (X, d) un espace métrique et μ une mesure de probabilité sur les boréliens de X. Pour une partie A de X, nous désignons par $A_r = \{d(\cdot, A) < r\}$ le r-voisinage de A. La *fonction de concentration* de μ est alors, pour $r \geqslant 0$,

$$\alpha_\mu(r) = \sup\{\mu((A_r)^c); \mu(A) \geqslant 1/2\} = 1 - \inf\{\mu(A_r); \mu(A) \geqslant 1/2\}.$$

Cette notion a été introduite par Milman et Schechtman dans [MS86]. Par définition même, la fonction $1 - \alpha_\mu$ donne une minoration uniforme de la masse des r-voisinages de toutes les parties de mesure significative. Plus précisément, pour tout borélien A de masse au moins $1/2$ et pour tout $r \geqslant 0$,

$$\mu(A_r) \geqslant 1 - \alpha_\mu(r). \tag{159}$$

Lemme A.2.2. — *$\alpha_\mu(r)$ décroît vers 0 quand r tend vers $+\infty$.*

Preuve du lemme A.2.2. — Comme α_μ est décroissante, il suffit, pour ε donné, de trouver r tel que $\mu(A_r) \geqslant 1 - \varepsilon$ uniformément sur les parties de masse $1/2$ au moins. Soit donc $\varepsilon < 1/2$. Comme, pour un point $x \in X$, $\cup_{r>0} B(x, r) = X$, il existe une

boule B telle que $\mu(B) \geqslant 1-\varepsilon$. Toute partie A telle que $\mu(A) \geqslant 1/2$ rencontre alors B et donc, si $r \geqslant \operatorname{diam}(B)$, A_r recouvre B et $\mu(A_r) \geqslant 1-\varepsilon$. □

A.2.3. La fonction de concentration d'une variable gaussienne. — Notons γ_{m,σ^2} la loi gaussienne sur $\mathbb{R}$ de moyenne m et de variance σ^2. Nous allons déterminer sa fonction de concentration. Commençons par la mesure gaussienne standard γ. La *forme intégrée de l'inégalité isopérimétrique* pour la mesure gaussienne (voir [Led96, Led99]) nous assure que, pour tout borélien A et tout $r \geqslant 0$,

$$\gamma(A_r) \geqslant \Phi\big(\Phi^{-1}(\gamma(A)) + r\big). \tag{160}$$

Ici, Φ désigne la fonction de répartition de γ, à savoir

$$\Phi(r) = \int_{-\infty}^{r} e^{-u^2/2} \frac{du}{\sqrt{2\pi}}.$$

En particulier, si $\gamma(A) \geqslant 1/2$, il vient $\gamma(A_r) \geqslant \Phi(r)$. D'autre part, l'égalité est réalisée dans (160) lorsque A est un demi-espace. En particulier, si $A =]-\infty, 0]$, $\gamma(A_r) = \Phi(r)$. On tire aisément de ces deux remarques la valeur de la fonction de concentration

$$\alpha_\gamma(r) = 1 - \Phi(r).$$

Il est alors facile d'en déduire que $\alpha_{\gamma_{m,\sigma^2}}(r) = 1 - \Phi(r/\sigma)$.

A.2.4. Concentration gaussienne d'une mesure. — Nous allons maintenant nous consacrer à l'étude des mesures μ qui se concentrent plus vite qu'une mesure gaussienne. Plus précisément, cela signifie qu'il existe une variance σ^2 telle que, pour r assez grand,

$$\alpha_\mu(r) \leqslant \alpha_{\gamma_{0,\sigma^2}}(r) = 1 - \Phi(r/\sigma). \tag{161}$$

Or la *queue* $1-\Phi(r) = \int_r^\infty e^{-u^2/2} du/\sqrt{2\pi}$ de la mesure gaussienne peut être encadrée par

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

Cela permet de reformuler (161) en la définition suivante.

Définition A.2.3. — *Nous dirons qu'une mesure de probabilité μ sur (X, d) vérifie* une propriété de concentration gaussienne *s'il existe deux constantes c et C strictement positives telles que, pour tout $r > 0$, $\alpha_\mu(r) \leqslant Ce^{-r^2/c}$. De manière équivalente, cela signifie que pour tout borélien A tel que $\mu(A) \geqslant 1/2$ et tout $r > 0$,*

$$\mu(A_r) \geqslant 1 - Ce^{-r^2/c}.$$

A.2.5. De l'importance des fonctions lipschitziennes. — Nous venons de donner la définition du phénomène de concentration de la mesure d'un point de vue ensembliste. Nous nous penchons ici sur une formulation fonctionnelle de cette propriété. La bonne classe de fonctions à considérer est la classe des fonctions lipschitziennes.

Définition A.2.4. — *Une fonction F sur un espace métrique (X, d) est dite* lipschitzienne *s'il existe une constante C telle que, pour tous points x et y de X,*

$$|F(x) - F(y)| \leqslant Cd(x, y). \tag{162}$$

Pour une telle fonction, la plus petite constante C vérifiant (162) *est notée*

$$\|F\|_{Lip} = \sup_{x \neq y} \frac{|F(x) - F(y)|}{d(x, y)}$$

et dite norme de Lipschitz *de F (bien que ce ne soit en fait qu'une semi-norme).*

Proposition A.2.5. — *Soient (X, d) un espace métrique et μ une mesure sur les boréliens de X. La minoration uniforme des mesures des r-voisinages (159) est équivalente à la déviation uniforme des fonctions lipschitziennes par rapport à leurs médianes. Cette dernière propriété signifie que, pour toute fonction F lipschitzienne, toute médiane m de F et tout $r \geqslant 0$,*

$$\mu(F \geqslant m + r) \leqslant \alpha_\mu \left(r / \|F\|_{Lip}\right). \tag{163}$$

Preuve

◀ Supposons (159) vérifiée. Remarquons que, par homogénéité, on peut supposer $\|F\|_{Lip} \leqslant 1$. On vérifie alors par l'inégalité triangulaire que, pour $r \geqslant 0$,

$$\{F \leqslant m\}_r \subset \{F < m + r\}.$$

Par conséquent,

$$\begin{aligned} \mu(F \geqslant m + r) &= 1 - \mu(F < m + r) \\ &\leqslant 1 - \mu\left(\{F \leqslant m\}_r\right) \leqslant \alpha_\mu(r), \end{aligned}$$

puisque $\mu\left(\{F \leqslant m\}\right) \geqslant 1/2$.

Inversement, supposons que (163) est vérifiée. On considère un ensemble A de mesure $\mu(A) \geqslant 1/2$ et la fonction $F = d(\cdot, A)$. Clairement, $\|F\|_{Lip} \leqslant 1$ et

$$\mu(F = 0) = \mu\left(\bar{A}\right) \geqslant \mu(A) \geqslant 1/2.$$

Donc 0 est une médiane de F et par suite, pour $r \geqslant 0$ fixé,

$$\mu(F \geqslant r) \leqslant \alpha_\mu(r).$$

Mais, si B^c désigne le complémentaire d'une partie B, $\{F \geqslant r\} = \{F < r\}^c = (A_r)^c$. Il en découle que $1 - \mu(A_r) \leqslant \alpha_\mu(r)$, ce qui est exactement l'inégalité (159) attendue. ▶

L'inégalité (163) est dite *inégalité de déviation de F par rapport à la médiane m.* Si on l'applique successivement à $-F$ et à F on obtient l'inégalité

$$\mu(|F-m| \geqslant r) \leqslant 2\alpha_\mu\left(r/\|F\|_{Lip}\right), \tag{164}$$

dite *inégalité de concentration* de la loi de F.

A.2.6. La concentration comme localisation de la mesure. — Ce qu'on appelle *phénomène de concentration gaussienne* pour la mesure μ est une localisation de la loi de toute contraction (ou fonction lipschitzienne) sur les compacts au sens de la décroissance au moins gaussienne des queues (rappelons que la queue d'une loi ν est la fonction $r \mapsto \nu(|\cdot| \geqslant r)$). Au premier abord, le terme « concentration » semble contenir une notion dynamique d'accumulation de la mesure autour d'*un unique* point, par exemple la médiane dans l'inégalité (164). En fait, il n'en est rien. Une inégalité de concentration gaussienne autour d'*un* point implique une inégalité de même type autour de *tous* les points. C'est ce que précise le lemme suivant :

Lemme A.2.6. — *Soit $(\mathbb{E}, \mathcal{F}, \mu)$ un espace de probabilité et soit F une fonction mesurable sur E. Supposons que F satisfait une inégalité de concentration gaussienne autour d'un réel a, c'est-à-dire que, pour deux constantes $c, C > 0$, F vérifie, pour tout $r \geqslant 0$,*

$$\mu(|F-a| \geqslant r) \leqslant Ce^{-r^2/c}. \tag{165}$$

Alors, pour tout autre réel b, il existe d'autres constantes $\tilde{c}, \tilde{C} > 0$ telles que, pour tout $r \geqslant 0$,

$$\mu(|F-b| \geqslant r) \leqslant \tilde{C}e^{-r^2/\tilde{c}}. \tag{166}$$

Preuve

◀ Cela consiste simplement à remarquer, d'une part, que si $r \geqslant 2|a-b|$, alors

$$\{|F-b| \geqslant r\} \subset \{|F-a| \geqslant r/2\}$$

et donc $\mu(|F-b| \geqslant r) \leqslant Ce^{-r^2/4c}$. D'autre part, pour $r \leqslant 2|a-b|$, la majoration est une trivialité :

$$\mu(|F-b| \geqslant r) \leqslant 1 \leqslant e^{|a-b|^2/c}e^{-r^2/4c}.$$

On obtient donc (166) avec les constantes $\tilde{c} = 4c$ et $\tilde{C} = \max\left(C, e^{|a-b|^2/c}\right)$. ▶

Remarque A.2.7. — *Notons que ce résultat ne fait intervenir aucune notion métrique sur l'espace $\mathbb{E}$.*

Dans la pratique, une inégalité de concentration (165) par rapport à un réel a permet, les constantes c et C étant connues, de trouver r pour lequel la fonction F prend ses valeurs dans l'intervalle $[a-r, a+r]$ mis à part sur un ensemble de mesure $1/100$ par exemple. Ce genre de contrôle est souvent utilisé lorsque a est la moyenne ou n'importe quelle médiane de F. En fait, si l'on dispose d'une inégalité (165) pour un a quelconque, cela impose à a de ne pas être « trop loin » de la

moyenne et des médianes de F. Le lemme précédent permet d'obtenir à partir de (165) une inégalité similaire de constantes $\tilde{c}$ et $\tilde{C}$ pour la moyenne de F. Ce qui est notable est que ces constantes $\tilde{c}$ et $\tilde{C}$ (obtenues pour $\mathbf{E}_\mu(F)$) ne dépendent plus de a mais seulement de c et C et s'expriment explicitement en fonction de ces dernières. En effet, retranscrivant la preuve précédente, on voit que, si $|a-b| \leqslant \lambda$, l'inégalité (166) est vérifiée pour

$$\tilde{C} = \max\left(C, e^{\lambda^2/c}\right) \text{ et } \tilde{c} = 4c. \tag{167}$$

Il ne reste plus qu'à remarquer que, si l'inégalité de concentration (165) est satisfaite pour a, alors $|\mathbf{E}_\mu(F) - a| \leqslant C\sqrt{\pi c}/2$ puisque

$$\mathbf{E}_\mu(|F-a|) = \int_0^\infty \mu(|F-a| \geqslant r)dr \leqslant C\int_0^\infty e^{-r^2/c}dr \leqslant C\frac{\sqrt{\pi c}}{2}.$$

Cela conduit aux constantes $\tilde{C} = \exp(\pi C^2/4)$ et $\tilde{c} = 4c$.

On obtient un résultat similaire en remplaçant la moyenne par n'importe quelle médiane de F. Il suffit pour cela de s'assurer que toutes les médianes m de F sont localisées autour de a, ou plus précisément qu'elles vérifient $|m-a| \leqslant \sqrt{c\log(4C)}$, et d'utiliser alors (167) pour $\lambda = \sqrt{c\log(4C)}$. Ce contrôle des médianes s'obtient de la manière suivante. Tout d'abord, par définition de λ, $Ce^{-\lambda^2/c} < 1/2$ et par conséquent $\mu(|F-a| \geqslant \lambda) < 1/2$. On en déduit en particulier que $\mu(F \geqslant a+\lambda) < 1/2$ et $\mu(F \leqslant a-\lambda) < 1/2$. Par suite, toute médiane est dans l'intervalle escompté.

A.2.7. Intégrabilité exponentielle. — Nous présentons maintenant une conséquence importante de la propriété de concentration gaussienne.

Proposition A.2.8. — *Soit F une application mesurable sur $(\mathbb{E}, \mathcal{F}, \mu)$ telle qu'il existe un réel a et des réels $c, C > 0$ pour lesquels, pour tout $r \geqslant 0$,*

$$\mu(|F-a| \geqslant r) \leqslant Ce^{-r^2/c},$$

pour tout $r > 0$. Alors, pour tout $\epsilon < 1/c$, l'espérance $\mathbf{E}_\mu\left(e^{\epsilon|F|^2}\right)$ est finie.

Preuve

◀ De façon générale, si Y est une variable positive sur $\mathbb{E}$ et f une fonction sur $\mathbb{R}^+$ à dérivée positive et localement intégrable, une utilisation du théorème de Fubini permet d'écrire :

$$\int_{\mathbb{E}} f(Y)d\mu = f(0) + \int_0^\infty f'(r)\mu(Y \geqslant r)dr.$$

Appliquée à $f(r) = e^{\varepsilon r^2}$ et $Y = |F|$, cette égalité entraîne que

$$\begin{aligned}\mathbf{E}_\mu\left(e^{\varepsilon|F|^2}\right) &= 1 + \int_0^\infty 2\varepsilon r e^{\varepsilon r^2}\mu(|F| \geqslant r)dr \\ &\leqslant e^{\varepsilon a^2} + \int_{|a|}^\infty 2\varepsilon r e^{\varepsilon r^2}\mu(|F| \geqslant r)dr.\end{aligned}$$

Il suffit alors pour conclure de remarquer que, si $r \geqslant |a|$, on a

$$\mu(|F| \geqslant r) \leqslant \mu(|F - a| \geqslant r - |a|) \leqslant Ce^{-(r-|a|)^2/c}.$$

►

Ainsi, la propriété de concentration gaussienne de μ entraîne l'*intégrabilité exponentielle* sous μ du carré de toute fonction lipschitzienne.

Remarque A.2.9. — *On pourrait obtenir des résultats similaires au lemme A.2.6 et à la proposition A.2.8 en remplaçant la fonction e^{-r^2} par une classe plus vaste de fonctions de concentration, en particulier celles du type e^{-r^p}, pour $p > 0$.*

A.3. Concentration via le critère de courbure

Nous donnons ici une preuve directe (c'est-à-dire n'ayant pas recours à l'inégalité isopérimétrique) de la forme de la fonction de concentration gaussienne. Rappelons que γ désigne la mesure gaussienne standard sur $\mathbb{R}$.

Théorème A.3.1. — *Soit F une fonction lipschitzienne sur $\mathbb{R}$ de norme de* Lipschitz $\|F\|_{Lip} \leqslant 1$. *Alors, pour tout $r \geqslant 0$,*

$$\gamma(F \geqslant \mathbf{E}_\gamma(F) + r) \leqslant e^{-r^2/2}. \tag{168}$$

Remarque A.3.2. — *Les arguments utilisés ici pour la mesure gaussienne sont transposables à toute mesure de probabilité μ pour laquelle on dispose d'un semi-groupe de diffusion (définition 2.6.1 du* livre*) symétrique dans $\mathbf{L}^2(\mu)$ satisfaisant au critère de courbure ((5.9) dans le* livre*) . En particulier, ils permettent de prouver la concentration gaussienne pour les mesures $\mu_\Phi(dx) = Z_\Phi^{-1}e^{-\Phi(x)}dx$ sur $\mathbb{R}^n$ décrites dans l'exemple fondamental 5.3.1 du* livre

telles que Hess $\Phi \geqslant \rho Id$ avec $\rho > 0$ (voir la remarque 5.3.6 du livre*).*

Preuve du théorème A.3.1. — Soit F une fonction de norme $\|F\|_{Lip} \leqslant 1$. La fonction F admet des moments exponentiels de tous ordres par rapport à γ car, pour tout λ positif, $\mathbf{E}_\gamma\left(e^{\lambda F}\right) \leqslant e^{\lambda F(0)}\mathbf{E}_\gamma\left(e^{\lambda|x|}\right) < \infty$. Afin de majorer $\gamma(F - \mathbf{E}_\gamma(F) \geqslant r)$, appliquons l'inégalité de Markov. On obtient ainsi, pour tout réel strictement positif λ,

$$\gamma(F - \mathbf{E}_\gamma(F) \geqslant r) = \gamma\left(e^{\lambda(F-\mathbf{E}_\gamma(F))} \geqslant e^{\lambda r}\right) \leqslant e^{-\lambda r}\mathbf{E}_\gamma(e^{\lambda(F-\mathbf{E}_\gamma(F))}).$$

On doit donc contrôler la transformée de LAPLACE de $F - \mathbf{E}_\gamma(F)$ sous γ. Pour ce faire, le point clé est de montrer l'estimation sous-gaussienne

$$\mathbf{E}_\gamma\big(e^{\lambda(F-\mathbf{E}_\gamma(F))}\big) \leqslant \exp\big(\lambda^2/2\big). \tag{169}$$

En effet, il en découle immédiatement que, pour tout $\lambda > 0$,

$$\gamma((F - \mathbf{E}_\gamma(F)) \geqslant r) \leqslant \exp\big(-\lambda r + \lambda^2/2\big). \tag{170}$$

Une optimisation en λ conduit alors à l'inégalité de déviation (168).

Prouvons maintenant l'estimation (169). Dans un premier temps, nous supposons que F est $\mathcal{C}^\infty$ bornée, ce qui rend licites les calculs qui suivent. On peut évidemment considérer ici que F est de moyenne nulle. Soit $(\mathbf{P}_t)_{t\geqslant 0}$ le semi-groupe d'ORNSTEIN-UHLENBECK qui, rappelons-le, est une diffusion réversible pour γ (voir paragraphe 2.3 du *livre*). Fixons un réel λ strictement positif et considérons la transformée de LAPLACE de $\mathbf{P}_t F$ en λ : $H(t) \overset{\text{déf.}}{=} \mathbf{E}_\gamma\big(e^{\lambda \mathbf{P}_t F}\big)$. La fonction H est dérivable et sa dérivée vaut $H'(t) = \lambda \mathbf{E}_\gamma\big(e^{\lambda \mathbf{P}_t F}\mathbf{L}\mathbf{P}_t F\big)$. Comme $(\mathbf{P}_t)_{t\geqslant 0}$ est ergodique (voir la définition 5.5.1 du *livre*) et $\mathbf{E}_\gamma(F)$ est nulle, par convergence dominée, $H(t)$ tend vers 1 quand t tend vers $+\infty$. Par suite,

$$H(t) = 1 - \int_t^\infty H'(s)ds.$$

Par la propriété de diffusion (voir (2.13) dans le *livre*),

$$\begin{aligned} H'(s) &= \lambda \int e^{\lambda \mathbf{P}_s F}\mathbf{L}\mathbf{P}_s F d\gamma \\ &= -\lambda \int \mathbf{\Gamma}\big(\mathbf{P}_s F, e^{\lambda \mathbf{P}_s F}\big) d\gamma \\ &= -\lambda^2 \int \mathbf{\Gamma}(\mathbf{P}_s F) e^{\lambda \mathbf{P}_s F} d\gamma. \end{aligned}$$

Il en résulte que

$$H(t) = 1 + \lambda^2 \int_t^\infty ds \int \mathbf{\Gamma}(\mathbf{P}_s F) e^{\lambda \mathbf{P}_s F} d\gamma.$$

Or ici, nous avons $\mathbf{\Gamma}(\mathbf{P}_s F) = |\nabla(\mathbf{P}_s F)|^2$ et, de plus,

$$\nabla(\mathbf{P}_s F) = \nabla \int_{\mathbb{R}} F\Big(e^{-s}x + \sqrt{1-e^{-2s}}y\Big)\gamma(dy) = e^{-s}\mathbf{P}_s(\nabla F),$$

où $\mathbf{P}_s(\nabla F)$ est à comprendre coordonnée par coordonnée. Et donc, par l'inégalité de JENSEN, on obtient[1]

$$|\nabla(\mathbf{P}_s F)|^2 = e^{-2s}|\mathbf{P}_s(\nabla F)|^2 \leqslant e^{-2s}\mathbf{P}_s\big(|\nabla F|^2\big).$$

[1] Ceci revient exactement à affirmer que le semi-groupe d'ORNSTEIN-UHLENBECK est de courbure 1 ; voir la proposition 5.4.1 du *livre*.

Remarquons enfin que l'hypothèse $\|F\|_{Lip} \leqslant 1$ équivaut à $|\nabla F| \leqslant 1$. On obtient donc

$$H(t) \leqslant 1 + \lambda^2 \int_t^\infty e^{-2s} H(s) ds,$$

et le lemme de GRONWALL entraîne la majoration

$$H(t) \leqslant \exp\left(\lambda^2 \int_t^\infty e^{-2s} ds\right) = \exp\left(\frac{\lambda^2}{2} e^{-2t}\right).$$

On aboutit finalement à l'inégalité souhaitée : $H(0) = \mathbf{E}_\gamma\big(e^{\lambda F}\big) \leqslant e^{\lambda^2/2}$.

L'extension de l'estimation (169) à toute fonction F telle que $\|F\|_{Lip} \leqslant 1$ découle du lemme d'approximation suivant.

Lemme A.3.3. — *Soit (X, d) un espace métrique et μ une mesure de probabilité sur les boréliens de X. Soit $(F_n)_{n\in\mathbb{N}}$ une suite de fonctions intégrables uniformément lipschitziennes (pour tout n, $\|F_n\|_{Lip} \leqslant 1$) tendant presque sûrement vers F, et telle que, pour tout entier n et tout réel positif λ,*

$$\mathbf{E}_\mu\big(e^{\lambda(F_n - \mathbf{E}_\mu(F_n))}\big) \leqslant e^{\lambda^2/2}. \tag{171}$$

Alors F est intégrable et vérifie l'inégalité (171).

Commençons par admettre ce lemme et par étendre l'estimation (169) à toutes les fonctions lipschitziennes. Tout d'abord, soient G une fonction lipschitzienne bornée telle que $\|G\|_{Lip} \leqslant 1$ et $(\rho_\varepsilon)_\varepsilon$ une famille régularisante (voir [Bré83]). Alors $G_\varepsilon \overset{\text{déf.}}{=} \rho_\varepsilon * G$ est une fonction $\mathcal{C}^\infty$ bornée et $\|G_\varepsilon\|_{Lip} \leqslant \|G\|_{Lip} \leqslant 1$. Ainsi, G_ε vérifie (169) et, d'après le lemme, il en est de même de G, puisque $(G_\varepsilon)_\varepsilon$ converge presque sûrement vers G. L'extension à toutes les fonctions lipschitziennes est obtenue de la même manière à l'aide de l'approximation suivante : si F est lipschitzienne de norme $\|F\|_{Lip} \leqslant 1$, $F_n \overset{\text{déf.}}{=} \max(-n, \min(F, n))$ est bornée, $\|F_n\|_{Lip} \leqslant \|F\|_{Lip} \leqslant 1$ et F_n tend vers F presque sûrement.

Preuve du lemme A.3.3. — Montrons, dans un premier temps, l'uniforme intégrabilité de la suite $(F_n)_{n\in\mathbb{N}}$. Pour ce faire, nous allons contrôler uniformément les espérances $\mathbf{E}_\mu(F_n^2)$. Or,

$$\mathbf{E}_\mu\big(F_n^2\big) = \mathbf{E}_\mu\big((F_n - \mathbf{E}_\mu(F_n))^2\big) + (\mathbf{E}_\mu(F_n))^2. \tag{172}$$

Pour contrôler le terme $\mathbf{E}_\mu\big((F_n - \mathbf{E}_\mu(F_n))^2\big)$, remarquons que l'inégalité (171) entraîne, comme cela a été fait en (170), l'inégalité de concentration

$$\mu(|F_n - \mathbf{E}_\mu(F_n)| \geqslant r) \leqslant 2e^{-r^2/2}. \tag{173}$$

En conséquence, on obtient la majoration suivante

$$\mathbf{E}_\mu\big((F_n - \mathbf{E}_\mu(F_n))^2\big) = 2\int_0^\infty r\mu(|F_n - \mathbf{E}_\mu(F_n)| \geqslant r) dr \leqslant 4\int_0^\infty r e^{-r^2/2} dr \leqslant 4.$$

D'autre part, on peut trouver une borne uniforme des espérances des F_n. En effet, soient m assez grand pour que $\mu(|F| \leqslant m) \geqslant 3/4$ et n_0 tel que, pour tout $n \geqslant n_0$, $\mu(|F_n - F| \geqslant 1) \leqslant 1/4$. Alors $\mu(|F_n| \leqslant m+1) \geqslant 1/2$, pour $n \geqslant n_0$. Soit maintenant r_1 tel que $2e^{-r_1^2/2} < 1/2$. L'ensemble $\{|F_n| \leqslant m+1\} \backslash \{|F_n - \mathbf{E}_\mu(F_n)| \geqslant r_1\}$ est non vide puisque de mesure strictement positive et donc

$$|\mathbf{E}_\mu(F_n)| \leqslant r_1 + m + 1,$$

pourvu que $n \geqslant n_0$.

On vient ainsi de montrer que

$$\sup_n \mathbf{E}_\mu\big(F_n^2\big) < \infty$$

et donc que la suite est uniformément intégrable. Il en découle alors que $\mathbf{E}_\mu(|F|) < \infty$ et que F_n converge vers F dans $\mathbf{L}^1(d\mu)$. Par conséquent, grâce au lemme de FATOU,

$$\mathbf{E}_\mu\big(e^{\lambda F}\big) \leqslant \liminf_{n\to\infty} \mathbf{E}_\mu\big(e^{\lambda F_n}\big) \leqslant \liminf_{n\to\infty} e^{\lambda \mathbf{E}_\mu(F_n) + \lambda^2/2} = e^{\lambda \mathbf{E}_\mu(F) + \lambda^2/2}.$$

Ceci termine la preuve du lemme A.3.3. □

La preuve du théorème A.3.1 est maintenant achevée. □

A.4. Concentration et inégalité de SOBOLEV logarithmique

A.4.1. L'argument de HERBST. — Nous venons de mettre en lumière le rôle que joue la transformée de LAPLACE dans l'étude de la concentration. Cette remarque permet d'étendre le résultat précédent aux mesures satisfaisant à une inégalité de SOBOLEV logarithmique, plutôt qu'au critère de courbure. La méthode utilisée est due à HERBST. La preuve présentée ici est valable pour toute mesure sur $\mathbb{R}^n$ à condition qu'elle satisfasse une inégalité de SOBOLEV logarithmique. Cette preuve pourrait aussi s'étendre à la mesure invariante d'une diffusion abstraite sur une algèbre standard définie en (2.11) dans le *livre*.

Théorème A.4.1. — *Soit μ une mesure de probabilité sur $\mathbb{R}^n$ vérifiant l'inégalité de* SOBOLEV *logarithmique*

$$\forall\, f \in \mathcal{C}_b^\infty(\mathbb{R}^n), \qquad \mathbf{Ent}_\mu(f^2) \leqslant c\mathbf{E}_\mu(|\nabla f|^2);$$

alors, pour toute fonction lipschitzienne F telle que $\|F\|_{Lip} \leqslant 1$,

$$\mu(F \geqslant \mathbf{E}_\mu(F) + r) \leqslant e^{-r^2/c} \tag{174}$$

et

$$\mu(|F - \mathbf{E}_\mu(F)| \geqslant r) \leqslant 2e^{-r^2/c}. \tag{175}$$

Preuve

◀ De la même manière que dans la preuve précédente, on suppose dans un premier temps que F est $\mathcal{C}^\infty$ bornée telle que $\|F\|_{Lip} \leqslant 1$.

Soit $H(\lambda) = \mathbf{E}_\mu\big(e^{\lambda F}\big)$ la transformée de LAPLACE de F. Appliquant l'inégalité de SOBOLEV logarithmique à $f^2 = e^{\lambda F}$ et ceci pour tout $\lambda > 0$, on obtient

$$\lambda H'(\lambda) - H(\lambda)\log H(\lambda) \leqslant c\frac{\lambda^2}{4}\int e^{\lambda F}|\nabla F|^2 d\mu \leqslant c\frac{\lambda^2}{4}H(\lambda) \tag{176}$$

puisque

$$\left|\nabla e^{\frac{\lambda}{2}F}\right|^2 = \frac{\lambda^2}{4}e^{\lambda F}|\nabla F|^2 \text{ et } |\nabla F| \leqslant 1.$$

De plus, la transformée de LAPLACE $H(\lambda)$ est strictement positive et l'inégalité (176) n'est autre que

$$\frac{H'(\lambda)}{\lambda H(\lambda)} - \frac{\log H(\lambda)}{\lambda^2} \leqslant \frac{c}{4}.$$

En posant $K(\lambda) = (1/\lambda)\log H(\lambda)$ pour $\lambda > 0$, cela revient à dire que

$$K'(\lambda) \leqslant c/4.$$

Un développement à l'ordre un nous assure de la convergence de $K(\lambda)$ vers $\mathbf{E}_\mu(F)$, quand λ tend vers 0. Finalement, comme K est continue sur $[0,\lambda]$, dérivable sur $]0,\lambda[$,

$$K(\lambda) - K(0) = \int_0^\lambda K'(s)ds \leqslant (c/4)\lambda,$$

ce qui correspond à l'estimation sous-gaussienne de la transformée de LAPLACE

$$H(\lambda) \leqslant e^{\lambda \mathbf{E}_\mu(F) + (c/4)\lambda^2}. \tag{177}$$

L'inégalité de MARKOV permet aisément d'en déduire l'inégalité de déviation (174) (voir pour cela le début de la preuve du théorème A.3.1). L'extension à toutes les fonctions lipschitziennes s'opère encore grâce au lemme A.3.3. ► Signalons que

par une méthode similaire (basée elle aussi sur un contrôle de la transformée de LAPLACE), on peut montrer (voir [AS94]) que l'inégalité de POINCARÉ induit une concentration en e^{-r} et l'intégrabilité exponentielle du module des fonctions lipschitziennes. LATALA et OLESZKIEWICZ ont récemment introduit des inégalités permettant d'interpoler entre inégalités de POINCARÉ et de SOBOLEV logarithmique (voir [LO00]). Ces nouvelles inégalités conduisent à une concentration en e^{-r^α} pour $1 \leqslant \alpha \leqslant 2$.

A.4.2. Concentration pour les moyennes empiriques. — Nous pouvons maintenant terminer la discussion entamée dans la section A.2.6 sur la signification du terme *concentration*.

Soient μ une loi de probabilité sur $\mathbb{R}^n$ vérifiant une inégalité de SOBOLEV logarithmique de constante c et $(X_i)_{i\in\mathbb{N}}$ une suite de vecteurs aléatoires indépendants identiquement distribués de loi μ. La propriété de concentration de μ implique une

concentration *dynamique* de la suite des mesures empiriques $1/N \sum_{i=1}^{N} \delta_{X_i}$ autour de μ, c'est-à-dire que, pour toute fonction f lipschitzienne sur $\mathbb{R}^n$ telle que $\|f\|_{Lip} \leqslant 1$,

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \mathbf{E}_\mu(f)\right| \geqslant r\right) \leqslant 2\exp\left(-\frac{Nr^2}{4c}\right). \tag{178}$$

En effet, d'après le théorème 3.2.2 du *livre*, la loi de $(X_1, \ldots, X_N)$ vérifie encore une inégalité de SOBOLEV logarithmique de constante c. De plus, la fonction $F_N \overset{\text{déf.}}{=} \frac{1}{N}\sum_{i\leqslant N} f(x_i)$ est lipschitzienne de rapport $1/\sqrt{N}$ sur $\mathbb{R}^{nN}$. L'inégalité (178) découle alors de l'inégalité de concentration (175) appliquée à $\sqrt{N}F_N$.

Remarquons que TALAGRAND obtient dans [Tal96a] des résultats similaires uniformes sur une classe de fonctions qui s'avèrent d'un grand intérêt en statistique.

A.4.3. Extension au cadre discret. — L'argument de HERBST introduit au paragraphe A.4.1 lorsque le gradient vérifie la propriété de dérivation

$$\nabla\Phi(f) = \Phi'(f)\,\nabla f \tag{179}$$

peut être légèrement adapté afin de couvrir le cadre des espaces discrets où (179) n'a plus lieu. Nous nous intéressons ici au cube $\mathbb{E} = \{-1, 1\}^n$ muni de la mesure produit $\mu = \boldsymbol{\beta}^{\otimes n}$, où $\boldsymbol{\beta}$ désigne comme précédemment la mesure de BERNOULLI uniforme sur l'espace à deux points $\{-1, 1\}$, $\boldsymbol{\beta} = 1/2(\delta_{-1} + \delta_1)$.

A.4.3.1. *Inégalité de* SOBOLEV *logarithmique.* — Sur $\{-1, 1\}$, l'opérateur $\mathbf{L}$ introduit au paragraphe 2.2 du *livre* engendre un semi-groupe de MARKOV de mesure réversible $\boldsymbol{\beta}$. Notant $\nabla f \overset{\text{déf.}}{=} f \circ \tau - f$ où $\tau(x) = -x$, il est facile de voir que $\mathbf{L}(f) = 1/2\,\nabla f$.

Sur $\mathbb{E}$ maintenant, la mesure μ est réversible pour le semi-groupe de générateur

$$\mathbf{L}(f) = 1/2 \sum_{i=1}^{n} \nabla_i f,$$

où $\nabla_i f \overset{\text{déf.}}{=} f \circ \tau_i - f$ avec $\tau_i(x_1, \ldots, x_n) = (x_1, \ldots, x_{i-1}, -x_i, x_{i+1}, \ldots, x_n)$. L'opérateur carré du champ associé à ce générateur est

$$\boldsymbol{\Gamma}(f) = 1/4 \sum_{i=1}^{n} (\nabla_i f)^2 \overset{\text{déf.}}{=} 1/4\,|\nabla f|^2.$$

D'après le théorème 1.3.2 du *livre*, qui établit l'inégalité de SOBOLEV logarithmique pour $\boldsymbol{\beta}$, et le corollaire 3.2.3 du *livre* qui permet sa tensorisation, la mesure μ vérifie l'inégalité de SOBOLEV logarithmique suivante : pour toute fonction f sur $\mathbb{E}$,

$$\mathbf{Ent}_\mu\big(f^2\big) \leqslant 1/2\,\mathbf{E}_\mu\big(|\nabla f|^2\big) = 2\,\mathbf{E}_\mu(\boldsymbol{\Gamma}(f)). \tag{180}$$

A.4.3.2. *Argument de* HERBST. — Bien que (179) ne soit pas vérifiée pour le gradient discret, on peut, en appliquant l'inégalité de SOBOLEV logarithmique aux fonctions $\exp\left(\frac{\lambda}{2}f\right)$, obtenir comme précédemment une inégalité différentielle sur la transformée de LAPLACE d'une fonction f telle que $\| \, |\nabla f| \, \|_\infty \leqslant 1$. Cela repose sur le lemme suivant :

Lemme A.4.2. — *Pour toute fonction f sur $\mathbb{E}$,*

$$\mathbf{E}_\mu\Big(\big|\nabla e^{f/2}\big|^2\Big) \leqslant 1/2 \, \big\| |\nabla f|^2 \big\|_\infty \, \mathbf{E}_\mu\big(e^f\big). \tag{181}$$

Remarque A.4.3. — *On pourra trouver dans* [Led99] *une version de cette inégalité dans le cadre plus vaste des formes de* DIRICHLET.

Le lemme précédent, dont nous verrons la démonstration plus loin, conduit facilement à une inégalité de déviation :

Théorème A.4.4. — *Pour toute fonction f sur $\mathbb{E}$ telle que $\| \, |\nabla f| \, \|_\infty \leqslant 1$ et tout $r \geqslant 0$,*

$$\mu(f - \mathbf{E}_\mu(f) \geqslant r) \leqslant e^{-r^2}. \tag{182}$$

Preuve du théorème A.4.4. — Conjuguant les inégalités (180) et (181), on obtient

$$\mathbf{Ent}_\mu\big(e^f\big) \leqslant 1/4 \, \big\| |\nabla f|^2 \big\|_\infty \, \mathbf{E}_\mu\big(e^f\big). \tag{183}$$

Comme au paragraphe A.4.1, cela conduit à l'estimation sous gaussienne de la transformée de LAPLACE d'une fonction f centrée telle que $\| \, |\nabla f| \, \|_\infty \leqslant 1$

$$\mathbf{E}_\mu\big(e^{\lambda f}\big) \leqslant e^{\lambda^2/4}.$$

On conclut toujours grâce à l'inégalité de MARKOV. □

Preuve du lemme A.4.2. — Soit $x \in \mathbb{E}$; on a

$$\big|\nabla e^{f/2}\big|^2(x) = \sum_{i=1}^n \big(\nabla_i \, e^{f/2}\big)^2(x) = \sum_{y \in \mathbb{E}} \mathbb{1}_{y \sim x} \big(e^{f(y)/2} - e^{f(x)/2}\big)^2,$$

où $y \sim x$ signifie que x et y sont voisins, c'est-à-dire que toutes leurs coordonnées sauf une sont égales. Par conséquent,

$$\begin{aligned}\mathbf{E}_\mu\Big(\big|\nabla e^{f/2}\big|^2\Big) &= \sum_{x,y \in \mathbb{E}} \mu(x) \mathbb{1}_{y \sim x} \big(e^{f(y)/2} - e^{f(x)/2}\big)^2 \\ &= 2 \sum_{f(y) < f(x)} \mu(x) \mathbb{1}_{y \sim x} \big(e^{f(y)/2} - e^{f(x)/2}\big)^2.\end{aligned}$$

La dernière égalité a lieu puisque $\mu(x)\mathbb{1}_{y\sim x}$ est symétrique en x et y. Maintenant,

$$\begin{aligned}\mathbf{E}_\mu\Big(\big|\nabla e^{f/2}\big|^2\Big) &\leqslant 1/2 \sum_{f(y)<f(x)} \mu(x)\mathbb{1}_{y\sim x}(f(y)-f(x))^2 e^{f(x)} \\ &\leqslant 1/2 \sum_{x\in\mathbb{E}} \mu(x)e^{f(x)}\left(\sum_{y\in\mathbb{E}} \mathbb{1}_{y\sim x}(f(y)-f(x))^2\right) \\ &\leqslant 1/2\mathbf{E}_\mu\big(|\nabla f|^2 e^f\big) = 1/2\mathbf{E}_\mu\left(|\nabla f|^2 \frac{e^f}{\mathbf{E}_\mu(e^f)}\right)\mathbf{E}_\mu\big(e^f\big).\end{aligned}$$

Et il suffit pour conclure de remarquer que

$$\big\||\nabla f|^2\big\|_\infty = \sup_{\|g\|_1\leqslant 1} \mathbf{E}_\mu\big(|\nabla f|^2 g\big).$$

□

A.4.3.3. *Inégalité de déviation pour les fonctions* HAMMING-*lipschitziennes.* — L'espace $\mathbb{E} = \{-1,1\}^n$ peut être muni comme tout espace produit (fini) de la distance de HAMMING

$$d(x,y) = \mathbf{Card}\{i \in \{1,\dots,n\}; x_i \neq y_i\}.$$

Le théorème A.4.4 permet d'obtenir facilement une inégalité de déviation par rapport à la moyenne pour les fonctions lipschitziennes pour la distance de HAMMING. Il suffit pour cela de remarquer que, si $\|F\|_{Lip} \leqslant 1$, alors $|\nabla F|^2 \leqslant n$. D'où la proposition suivante :

Proposition A.4.5. — *Soit F une fonction lipschitzienne sur $\mathbb{E}$ muni de la distance de* HAMMING *telle que $\|F\|_{Lip} \leqslant 1$; alors, pour tout $r \geqslant 0$,*

$$\mu(F - \mathbf{E}_\mu(F) \geqslant r) \leqslant e^{-r^2/n}. \tag{184}$$

Nous venons d'établir une inégalité de déviation pour les fonctions HAMMING-lipschitziennes pour l'espace produit $\mathbb{E} = \{-1,1\}^n$ à partir de l'inégalité de SOBOLEV logarithmique sur $\{-1,1\}$.

En fait, l'inégalité (184) peut être obtenue (à un facteur 2 près) sur un produit fini quelconque

$$\left(\mathbb{E} = \prod_{i=1}^n E_i, \mu = \otimes_{i=1}^n \mu_i\right), \tag{185}$$

que les facteurs (E_i, μ_i) satisfassent à une inégalité de SOBOLEV logarithmique ou non. La technique utilisée pour obtenir ce résultat est fortement similaire. Le rôle joué précédemment par l'inégalité de SOBOLEV logarithmique est dans ce cadre

accompli par une autre inégalité qui, elle, a lieu sur tout espace produit : pour toute fonction f sur $\mathbb{E}$,

$$\mathbf{Ent}_\mu\big(e^f\big) \leqslant \frac{1}{2}\mathbf{E}_\mu\big(e^f(df)^2\big). \tag{186}$$

Ici, par définition, $(df(x))^2 = \sum_{i=1}^n (d_i f(x))^2$ avec

$$(d_i f(x))^2 \overset{\text{déf.}}{=} \int_{E_i} \mu_i(dv)(f(x) - f(x._i v))^2.$$

Pour $x \in \mathbb{E}$ et $v \in E_i$, on a noté $x._i v \overset{\text{déf.}}{=} (x_1, \ldots, x_{i-1}, v, x_{i+1}, \ldots, x_n)$. Nous renvoyons au chapitre 3 de [Led99] pour une preuve de (186). Cette inégalité donne lieu à une inégalité analogue à (183)

$$\mathbf{Ent}_\mu\big(e^f\big) \leqslant 1/2 \left\|(df)^2\right\|_\infty \mathbf{E}_\mu\big(e^f\big),$$

qui conduit de la même manière, lorsque $\|F\|_{Lip} \leqslant 1$, à

$$\mu(F - \mathbf{E}_\mu(F) \geqslant r) \leqslant e^{-r^2/(2n)}. \tag{187}$$

Exemple A.4.6. — *L'exemple fondamental de fonction lipschitzienne pour la distance de* HAMMING *est la fonction*

$$F(x) = \sum_{i=1}^n f_i(x_i)$$

où $\sup_i \|f_i\|_\infty < \infty$. *L'inégalité* (187) *permet donc d'obtenir des résultats similaires à* (178) *pour une fonction f bornée.*

A.5. L'argument de HERBST inverse

A.5.1. Théorème de WANG. — L'argument de HERBST nous assure que tout semi-groupe de diffusion satisfaisant à une inégalité de SOBOLEV logarithmique vérifie une propriété de concentration gaussienne, et donc que toutes les contractions (au sens du carré du champ, c'est-à-dire telles que $\mathbf{\Gamma}(f) \leqslant 1$) sont de carré exponentiellement intégrable. Nous nous attachons ici à une sorte de réciproque due à WANG (voir [Wan97]) : si le semi-groupe est de courbure minorée, une intégrabilité exponentielle assez forte du carré de la distance donne une inégalité de SOBOLEV logarithmique (et donc une concentration gaussienne). La constante de SOBOLEV logarithmique obtenue ne peut malheureusement pas être aisément explicitée. En outre, précisons que ce critère ne fournit pas de contrôle de cette constante stable par rapport à la dimension.

Théorème A.5.1 **(Théorème de WANG)**. — *Soit* (M, g) *une variété riemannienne connexe et complète pour la distance riemannienne d définie par la longueur*

des courbes. Soit Φ une fonction sur M de classe $\mathcal{C}^2$ telle que $e^{-\Phi}$ soit intégrable par rapport à la mesure riemannienne notée dx. On considère la mesure de probabilité

$$\mu(dx) = Z_\Phi^{-1}\, e^{-\Phi(x)}\, dx$$

sur M. Nous supposerons ici que, pour un $\rho \in \mathbb{R}$ non forcément positif,

$$\mathrm{Ric}(M) + \nabla\nabla\Phi \geqslant \rho g. \qquad \textit{(condition de courbure minorée)} \tag{188}$$

Si, de plus, pour $\tilde{\rho} = \min(\rho, 0)$, il existe $\varepsilon > 0$ tel que

$$\int\int e^{(-\tilde{\rho}+\varepsilon)d(x,y)^2}\mu(dx)\mu(dy) < \infty, \qquad \textit{(condition de } \textsc{Wang}\textit{)} \tag{189}$$

Alors μ vérifie une inégalité de Sobolev *logarithmique*

$$\forall\, f \in \mathcal{C}_c^\infty(M)\,, \qquad \mathbf{Ent}_\mu(f^2) \leqslant c\mathbf{E}_\mu(|\nabla f|^2). \tag{190}$$

Remarque A.5.2. — *La condition de* Wang *est vérifiée dès que, pour un point x_0 de M,*

$$\int e^{-2(\tilde{\rho}-\varepsilon)d(x,x_0)^2}\mu(dx)$$

est fini.

Rappelons ici que le semi-groupe de diffusion $(\mathbf{P}_t)_{t\geqslant 0}$ de générateur (sur les fonctions $\mathcal{C}^\infty$ à support compact sur M)

$$\mathbf{L} = \Delta - \nabla\Phi \tag{191}$$

admet μ comme mesure réversible. La condition (188) revient à dire que l'opérateur $\mathbf{L}$ vérifie le critère de courbure (5.9) dans le *livre*.

Pour des raisons de simplicité, nous présenterons, au paragraphe A.5.4, la preuve de ce résultat lorsque $M = \mathbb{R}^n$. La condition de courbure minorée se résume alors simplement à

$$\mathrm{Hess}\ \Phi \geqslant \rho \mathrm{Id}.$$

Les arguments sont aisément transposables au cadre général (on pourra se référer à l'article de Wang précédemment cité).

Une extension intéressante du théorème A.5.1 consiste à amoindrir la condition de courbure minorée en une condition *à l'infini*. Remarquons en effet que la condition de Wang (189) est invariante à une constante près par perturbation bornée, tout comme l'inégalité de Sobolev logarithmique (voir le théorème 3.2.2 dans le *livre*). Cela permet d'obtenir le corollaire suivant.

Corollaire A.5.3. — *Soient U une fonction bornée de $\mathbb{R}^n$ dans $\mathbb{R}$ et Φ une fonction sur $\mathbb{R}^n$ de classe $\mathcal{C}^2$ telle que $e^{-\Phi}$ soit intégrable par rapport à la mesure de* Lebesgue *et, pour $\rho \in \mathbb{R}$,*

$$\mathrm{Hess}\ \Phi \geqslant \rho \mathrm{Id}.$$

On considère la mesure $\mu_H(dx) = Z_H^{-1} e^{-H(x)} dx$ dont la phase $H = \Phi + U$ est obtenue par perturbation bornée de Φ. Si, pour un ε assez petit, μ_H vérifie la condition de

WANG (189) *pour la distance euclidienne, alors* μ_H *vérifie une inégalité de* SOBOLEV *logarithmique.*

A.5.2. Inégalité de type HARNACK. — Lors de la preuve du théorème A.5.1, WANG a établi une inégalité de type HARNACK pour la solution de l'équation de la chaleur $\partial_t = \mathbf{L}$, où $\mathbf{L}$ est l'opérateur défini en (191).

***Lemme A.5.4* (Lemme de WANG).** — *Sous les hypothèses du théorème A.5.1, on a, pour toute fonction f borélienne bornée sur M,*

$$\forall\, x, y \in M,\ \forall\, t > 0, \qquad (\mathbf{P}_t f(x))^2 \leqslant \mathbf{P}_t(f^2)(y) e^{\rho\left(e^{2\rho t}-1\right)^{-1} d(x,y)^2}. \tag{192}$$

Nous donnerons la preuve de ce résultat dans le cas $M = \mathbb{R}^n$ au paragraphe suivant.

Une inégalité de HARNACK permet la comparaison des valeurs, en différents points, de la solution d'une équation aux dérivées partielles. Le lemme A.5.4 n'en exprime qu'une version affaiblie, puisqu'on ne compare pas exactement les valeurs de $\mathbf{P}_t f$. Néanmoins, un trait remarquable de ce lemme est son indépendance vis-à-vis de la dimension.

Le semi-groupe $(\mathbf{P}_t)_{t\geqslant 0}$ étant une diffusion de courbure ρ, on a, d'après la remarque 5.4.6 du *livre*,

$$|\nabla \mathbf{P}_t f| \leqslant e^{-\rho t} \mathbf{P}_t |\nabla f|. \tag{193}$$

En appliquant le lemme précédent à la fonction $|\nabla f|$ au lieu de f et en utilisant la relation de commutation (193), on obtient aisément le corollaire suivant.

Corollaire A.5.5. — *Sous les hypothèses du théorème A.5.1, pour toute fonction f de classe $\mathcal{C}^1$ sur M de gradient borné et tous x et y dans M,*

$$|\nabla(\mathbf{P}_t f)|(x) \leqslant e^{-\rho t} e^{1/2\,\rho\left(e^{2\rho t}-1\right)^{-1} d(x,y)^2} \sqrt{\mathbf{P}_t\big(|\nabla f|^2\big)(y)}. \tag{194}$$

Ce corollaire exprime un contrôle de la norme de LIPSCHITZ de $\mathbf{P}_t f$ en fonction de celle de f.

A.5.3. Preuve du lemme de WANG. — Rappelons que nous nous restreignons ici au cas $M = \mathbb{R}^n$. Dans ce cas, $\mathbf{L} = \Delta - \nabla\Phi \cdot \nabla$, où Δ est le laplacien usuel.

Soit f une fonction que l'on peut supposer $\mathcal{C}^\infty$ à support compact. Nous avons vu dans le chapitre 5 du *livre* qu'à la base de l'utilisation du critère de courbure intervient le calcul de la dérivée en s de $\psi(s,x) = \mathbf{P}_s\big((\mathbf{P}_{t-s} f)^2\big)(x)$ (voir la proposition 5.4.1). On a

$$\frac{\partial}{\partial s}\psi(s,x) = 2\mathbf{P}_s(\mathbf{\Gamma}(\mathbf{P}_{t-s} f))(x).$$

La fonction $\psi(s,x)$ intervient dans la preuve de l'inégalité de POINCARÉ locale puisque $\psi(t,x) = \mathbf{P}_t(f^2)(x)$ et $\psi(0,x) = (\mathbf{P}_t f)^2(x)$.

Ici, il s'agit de composer $\psi(s,\cdot)$ par un chemin partant de x en 0 et atteignant y en t. Soit h une paramétrisation $\mathcal{C}^1$ du segment $[0,t]$, c'est-à-dire une fonction $\mathcal{C}^1$ de $[0,t]$ sur lui-même telle que $h(0)=0$ et $h(t)=t$. Notons

$$\tau(s)=\left(1-\frac{h(s)}{t}\right)x+\frac{h(s)}{t}y,$$

le chemin défini sur $[0,t]$ parcourant le segment $[x,y]$ à la vitesse $h'(\cdot)/t$.

Remarque A.5.6. — *On pourrait considérer plus simplement le chemin parcourant $[x,y]$ à vitesse uniforme, mais nous verrons qu'une optimisation sur la paramétrisation de $[0,t]$ conduit à un meilleur contrôle.*

On note $\varphi(s)=\psi(s,\tau(s))$. Il vient, notant $\langle\cdot,\cdot\rangle$ le produit scalaire de $\mathbb{R}^n$,

$$\begin{aligned}\varphi'(s)&=2\mathbf{P}_s(\mathbf{\Gamma}(\mathbf{P}_{t-s}f))(\tau(s))+\big\langle\nabla\mathbf{P}_s\big((\mathbf{P}_{t-s}f)^2\big)(\tau(s)),\tau'(s)\big\rangle\\&=2\mathbf{P}_s(\mathbf{\Gamma}(\mathbf{P}_{t-s}f))(\tau(s))+t^{-1}h'(s)\big\langle\nabla\mathbf{P}_s\big((\mathbf{P}_{t-s}f)^2\big)(\tau(s)),y-x\big\rangle\\&\geqslant 2\mathbf{P}_s(\mathbf{\Gamma}(\mathbf{P}_{t-s}f))(\tau(s))-t^{-1}|h'(s)||x-y|\big|\nabla\mathbf{P}_s\big((\mathbf{P}_{t-s}f)^2\big)(\tau(s))\big|.\end{aligned}$$

Remarquons que $\mathbf{\Gamma}(f)=|\nabla f|^2$ et appliquons la relation de commutation (193) à $(\mathbf{P}_{t-s}f)^2$. On obtient

$$\big|\nabla\mathbf{P}_s\big((\mathbf{P}_{t-s}f)^2\big)(\tau(s))\big|\leqslant e^{-\rho s}\mathbf{P}_s\big|\nabla\big[(\mathbf{P}_{t-s}f)^2\big]\big|=2e^{-\rho s}\mathbf{P}_s(|\mathbf{P}_{t-s}f||\nabla\mathbf{P}_{t-s}f|).$$

Vu que $b^2-2ab\geqslant -a^2$, il s'ensuit

$$\begin{aligned}\varphi'(s)&\geqslant 2\mathbf{P}_s\big(|\nabla\mathbf{P}_{t-s}f|^2-t^{-1}|h'(s)||x-y|e^{-\rho s}|\mathbf{P}_{t-s}f||\nabla\mathbf{P}_{t-s}f|\big)(\tau(s))\\&\geqslant-\frac{t^{-2}}{2}|x-y|^2e^{-2\rho s}(h'(s))^2\varphi(s).\end{aligned}$$

Le lemme de Gronwall entraîne alors la majoration

$$(\mathbf{P}_tf(x))^2\leqslant\big(\mathbf{P}_t(f^2)\big)(y)\exp\left(\frac{t^{-2}}{2}|x-y|^2\int_0^t e^{-2\rho s}(h'(s))^2ds\right).$$

Il ne reste plus qu'à minimiser la fonctionnelle $U(h)=\int_0^t G(s,h(s),h'(s))ds$, où $G(s,x,y)=e^{-2\rho s}y^2$, sur l'ensemble des chemins h $\mathcal{C}^1$ tels que $h(0)=0$ et $h(t)=t$. La paramétrisation optimale est

$$h(s)=t\big(e^{2\rho t}-1\big)^{-1}\big(e^{2\rho s}-1\big).$$

On l'obtient par un calcul de variations élémentaire ou grâce à l'équation d'Euler-Lagrange

$$\frac{d}{ds}\frac{\partial G}{\partial y}(s,h(s),h'(s))=\frac{\partial G}{\partial x}(s,h(s),h'(s)),$$

associée aux conditions de bord. Le résultat en découle.

A.5.4. Preuve du théorème de Wang. — Comme au paragraphe précédent, nous nous restreindrons pour la preuve au cas $M = \mathbb{R}^n$. Avant toute chose, commençons par simplifier le problème.

Remarque A.5.7. — *Nous pouvons nous restreindre au cas où la courbure ρ est strictement négative. Seule la détermination de la constante c en sera affectée. En effet, si $\rho > 0$, le critère de courbure est vérifié et assure l'existence d'une inégalité de* Sobolev *logarithmique de constante $2/\rho$ sans aucune hypothèse d'intégrabilité exponentielle (voir le théorème 5.4.7 et la discussion faisant suite à la définition 5.5.1 dans le* livre*). Enfin, si les conditions de courbure minorée et de* Wang *sont satisfaites pour $\rho = 0$ et $\varepsilon = \eta$, elles le sont trivialement pour $\rho = -\eta/2$ et $\varepsilon = \eta/2$.*

La preuve du théorème A.5.1 se décompose en deux parties distinctes que nous traiterons sous forme de paragraphes. Tout d'abord, nous établissons une inégalité de Poincaré, puis, ceci étant acquis, nous en déduisons une inégalité de Sobolev logarithmique via l'hypercontractivité du semi-groupe associé à la mesure.

A.5.4.1. *Intégrabilité exponentielle et existence d'un trou spectral.* — Nous nous attachons dans un premier temps à déduire de la condition de Wang (189) une inégalité de Poincaré.

Idée générale de la preuve :
Soit f une fonction $\mathcal{C}^\infty$ à support compact que l'on peut supposer centrée. Nous devons contrôler $\mathbf{Var}_\mu(f) = \mathbf{E}_\mu(f^2)$ par l'énergie $\mathbf{E}_\mu\big(|\nabla f|^2\big)$. Pour ce faire, considérons le semi-groupe de diffusion $(\mathbf{P}_t)_{t\geqslant 0}$ défini en (191). On se ramène à étudier $\mathbf{E}_\mu\big((\mathbf{P}_t f)^2\big)$ grâce au lemme suivant.

Lemme A.5.8. — *Soit f une fonction $\mathcal{C}^\infty$ bornée. Pour tout $t \geqslant 0$,*

$$\mathbf{E}_\mu\big(f^2\big) \leqslant \mathbf{E}_\mu\big((\mathbf{P}_t f)^2\big) + 2t\mathbf{E}_\mu\big(|\nabla f|^2\big). \tag{195}$$

Nous verrons un peu plus loin la preuve de ce lemme. Comme $\mathbf{E}_\mu(f) = 0$, la variance $\mathbf{E}_\mu\big((\mathbf{P}_t f)^2\big)$ tend vers 0 quand t tend vers l'infini. C'est la condition de Wang qui nous permettra de fixer une borne t pour laquelle $\mathbf{E}_\mu\big((\mathbf{P}_t f)^2\big)$ est contrôlé par $\mathbf{E}_\mu\big(|\nabla f|^2\big)$ uniformément sur toutes les fonctions $\mathcal{C}^\infty$ bornées. Cette dernière étape constitue en fait le coeur de la preuve et nécessite une large utilisation de méthodes de semi-groupes.

Preuve du lemme A.5.8. — Posons

$$\varphi(t) \overset{\text{déf.}}{=} \mathbf{E}_\mu\big((\mathbf{P}_t f)^2\big).$$

Il vient $\varphi'(t) = 2\mathbf{E}_\mu(\mathbf{P}_t f \mathbf{L}\mathbf{P}_t f)$ et

$$\varphi''(t) = 2\big[\mathbf{E}_\mu\big((\mathbf{L}\mathbf{P}_t f)^2\big) + \mathbf{E}_\mu(\mathbf{P}_t f \mathbf{L}\mathbf{L}(\mathbf{P}_t f))\big] = 4\mathbf{E}_\mu\big((\mathbf{L}\mathbf{P}_t f)^2\big) \geqslant 0,$$

puisque le semi-groupe est réversible. Vu que

$$\varphi'(0) = 2\mathbf{E}_\mu(f\mathbf{L}f) = -2\mathbf{E}_\mu(\mathbf{\Gamma}(f)) = -2\mathbf{E}_\mu\big(|\nabla f|^2\big),$$

la formule de TAYLOR permet de conclure. □

Preuve proprement dite :
Soit comme précédemment f une fonction $\mathcal{C}^\infty$ à support compact de moyenne nulle. Afin de contrôler la variance de $\mathbf{P}_t f$, fixons dans un premier temps un point $x \in \mathbb{R}^n$ et étudions $|\mathbf{P}_t f(x)|$. Par invariance de la mesure,

$$|\mathbf{P}_t f(x)| = \left| \int (\mathbf{P}_t f(x) - \mathbf{P}_t f(y))\mu(dy) \right| \leqslant \int |\mathbf{P}_t f(x) - \mathbf{P}_t f(y)|\mu(dy).$$

Or, $|\mathbf{P}_t f(x) - \mathbf{P}_t f(y)| \leqslant |x-y||\nabla \mathbf{P}_t f(z)|$ pour un z sur le segment $[x, y]$.

Le corollaire A.5.5 permet alors d'éliminer l'inconnue z en fournissant un contrôle de $|\nabla \mathbf{P}_t f(z)|$ dans lequel on majore la distance de y à z par celle séparant y de x. On obtient ainsi

$$(\mathbf{P}_t f(x))^2 \leqslant \left(\int |x-y| e^{-\rho t} e^{\frac{1}{2}\rho\left(e^{2\rho t}-1\right)^{-1}|x-y|^2} \sqrt{\mathbf{P}_t(|\nabla f|^2)(y)}\mu(dy) \right)^2.$$

L'inégalité de CAUCHY-SCHWARZ appliquée à ce dernier terme conduit, après intégration en x et utilisation de la propriété d'invariance, à la majoration

$$\mathbf{Var}_\mu(\mathbf{P}_t f) \leqslant e^{-2\rho t} \mathbf{E}_\mu\big(|\nabla f|^2\big) \int\int |x-y|^2 e^{\rho\left(e^{2\rho t}-1\right)^{-1}|x-y|^2} \mu(dx)\mu(dy).$$

Puisque l'on a supposé $\rho < 0$, il vient $\rho(e^{2\rho t}-1)^{-1} < -\rho + \varepsilon$ pour t assez grand, et l'intégrale double ci-dessus est alors finie. Il est temps de faire appel au lemme A.5.8 appliqué pour un tel t. On obtient

$$\mathbf{Var}_\mu(f) \leqslant \left(2t + e^{-2\rho t} \int\int |x-y|^2 e^{\rho\left(e^{2\rho t}-1\right)^{-1}|x-y|^2} \mu(dx)\mu(dy) \right) \mathbf{E}_\mu\big(|\nabla f|^2\big),$$

c'est-à-dire une inégalité de trou spectral pour la mesure μ.

A.5.4.2. *Inégalité de* SOBOLEV *logarithmique.* — Nous venons de montrer que la condition de WANG entraîne une inégalité de POINCARÉ. D'après le théorème 4.3.9 du *livre*, cette inégalité permet de « tendre » une inégalité de SOBOLEV logarithmique non tendue (voir (2.6.4) dans le *livre*). Notre objectif est donc de déduire de la condition de WANG une telle inégalité. On peut alors utiliser le théorème de HOEGH-KROHN et SIMON (théorème 2.8.6 dans le *livre*) : pour obtenir une inégalité de SOBOLEV logarithmique non tendue, il suffit de s'assurer que le semi-groupe réversible pour μ, engendré par $\mathbf{L} = \Delta - \nabla\Phi \cdot \nabla$, est hypercontractif au sens affaibli, c'est-à-dire qu'il existe $1 < p < q < \infty$ tels que $\|\mathbf{P}_t\|_{p\to q} < \infty$ pour un certain $t > 0$. C'est ce que nous allons montrer maintenant.

L'argument est basé sur l'inégalité de HÖLDER. Soient $1 < \theta < 2$ et $t > 0$. Appliquant le lemme A.5.4,

$$\begin{aligned}\mathbf{E}_\mu\Big(|\mathbf{P}_t f|^{2\theta}\Big) &= \mathbf{E}_\mu\Big(|\mathbf{P}_t f|^\theta \big(|\mathbf{P}_t f|^2\big)^{\frac{\theta}{2}}\Big)\\ &\leqslant \int\int |\mathbf{P}_t f(x)|^\theta e^{\frac{\theta}{2}\rho\left(e^{2\rho t}-1\right)^{-1}|x-y|^2}\big(\mathbf{P}_t\big(f^2\big)(y)\big)^{\frac{\theta}{2}}\mu(dx)\mu(dy).\end{aligned}$$

Détaillons l'inégalité de HÖLDER sur le produit $\mu\otimes\mu(dx,dy)$

$$\|hg\|_1 \leqslant \|h\|_u \|g\|_v$$

pour $v = 2/(2-\theta)$, $u = 2/\theta$. Les fonctions h et g sont données ici par

$$h(x,y) = |\mathbf{P}_t f(x)|^\theta \big(\mathbf{P}_t\big(f^2\big)(y)\big)^{\frac{\theta}{2}}$$

et

$$g(x,y) = e^{\frac{\theta}{2}\rho\left(e^{2\rho t}-1\right)^{-1}|x-y|^2}.$$

On a $\|g\|_v = N_{\theta,t}^{\frac{2-\theta}{2}}$, où

$$N_{\theta,t} = \int\int e^{\frac{\theta}{2-\theta}\rho\left(e^{2\rho t}-1\right)^{-1}|x-y|^2}\mu(dx)\mu(dy).$$

Pour θ assez proche de 1 et t assez grand, la condition de WANG assure que $N_{\theta,t}$ est fini. Fixons dorénavant $t \geqslant 0$ et $\theta > 1$ vérifiant cette propriété. D'après l'inégalité de JENSEN sur les mesures $\mathbf{P}_t(\cdot)(x)$,

$$\begin{aligned}\|h\|_{u,\mu\otimes\mu} &= \left(\int (\mathbf{P}_t f(x))^2\mu(dx)\right)^{\frac{\theta}{2}}\left(\int \mathbf{P}_t\big(f^2\big)(y)\mu(dy)\right)^{\frac{\theta}{2}}\\ &\leqslant \left(\int \mathbf{P}_t\big(f^2\big)(y)\mu(dy)\right)^{\theta} = \|f\|_{2,\mu}^{2\theta},\end{aligned}$$

cette dernière égalité étant due à l'invariance de la mesure.

Finalement,

$$\|\mathbf{P}_t f\|_{2\theta,\mu}^{2\theta} = \mathbf{E}_\mu\Big(|\mathbf{P}_t f|^{2\theta}\Big) \leqslant \|hg\|_{1,\mu\otimes\mu} \leqslant N_{\theta,t}^{\frac{2-\theta}{2}}\|f\|_{2,\mu}^{2\theta}$$

et donc

$$\|\mathbf{P}_t f\|_{2\theta,\mu} \leqslant N_{\theta,t}^{\frac{2-\theta}{4\theta}}\|f\|_{2,\mu}. \tag{196}$$

Comme $\theta > 1$, le semi-groupe est faiblement hypercontractif. Et ceci achève la preuve d'après la discussion du début du paragraphe.

A.6. Notes

La propriété de concentration de la mesure intervient à l'heure actuelle dans de nombreux domaines des probabilités et de l'analyse. Son importance a été révélée par MILMAN dans la théorie locale des espaces de BANACH (voir [Mil71, Mil92, MS86]). Dans [Mil71], MILMAN donne une preuve nouvelle, fondée sur l'inégalité isopérimétrique des sphères et les inégalités de concentration qui en découlent, du théorème de DVORETZKY. Cet important théorème établit le caractère presque elliptique, en grande dimension, des sections euclidiennes d'un corps convexe [Dvo61]. Puis GROMOV et MILMAN [GM83] ont utilisé la concentration pour déduire des applications topologiques d'inégalités isopérimétriques. TALAGRAND s'est plus récemment attaché à l'étude d'inégalités de concentration pour des mesures produit (voir [Tal88, Tal95, Tal96b]). Il en découle en particulier des estimations précises pour les processus empiriques (voir [Tal96a]) utilisées par BIRGÉ et MASSART [BM98, BM97] dans le traitement statistique de la sélection de modèles. Une autre approche du phénomène de concentration de la mesure via des inégalités de transport [Mar96, Tal96c] sera abordée dans le chapitre 8 du *livre*.

La discussion qui fait suite au lemme A.2.6 établit l'équivalence aux constantes près entre les inégalités de concentration autour d'une médiane et autour de la moyenne. Cette équivalence est initialement mentionnée dans [MS86]. La notion de fonction de concentration remonte également à ce livre.

La méthode employée pour déduire une concentration gaussienne d'une inégalité de SOBOLEV logarithmique (voir théorème A.4.1) est originellement due à HERBST. Elle est mentionnée pour la première fois dans [DS84] et a été réactualisée par AIDA, MASUDA et SHIGEKAWA dans [AMS94]. Depuis, de nombreux travaux, dont ceux de AIDA et STROOCK [AS94], BOBKOV et GÖTZE [BG99], GROSS et ROTHAUS [GR98], et LEDOUX [Led99] ont développé la méthode pour l'étendre à des cadres plus généraux. En particulier, LEDOUX [Led97] retrouve par cette méthode l'essentiel des résultats de concentration pour les mesures produits de TALAGRAND.

Le lien entre inégalité de POINCARÉ et concentration en e^{-r} remonte à GROMOV et MILMAN [GM83] et a été étudié depuis par différentes méthodes : en utilisant des bornes sur les moments [AMS94] et parallèlement via une inégalité différentielle sur la transformée de LAPLACE [Sch98].

L'inégalité de style HARNACK du lemme A.5.4 a été établie dans [Wan97]. Le théorème de WANG remonte à cet article. D'autres travaux traitent de questions similaires. Citons en particulier les résultats de AIDA [Aid98] et, pour les mesures log-concaves sur $\mathbb{R}^n$, l'article de BOBKOV [Bob97].

Signalons pour finir qu'un bon ouvrage de référence sur nombre de points abordés ici est [Led99], notes d'un cours de LEDOUX donné au « Graduierten-kolleg » de Berlin.

BIBLIOGRAPHIE

[ABC+00] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto & G. Scheffer – *Sur les inégalités de sobolev logarithmiques*, Société Mathématique de France, Paris, 2000.

[Aid98] S. Aida – « Uniform positivity improving property, Sobolev inequalities, and spectral gaps », *J. Funct. Anal.* **158** (1998), no. 1, p. 152–185.

[AMS94] S. Aida, T. Masuda & I. Shigekawa – « Logarithmic Sobolev inequalities and exponential integrability », *J. Funct. Anal.* **126** (1994), no. 1, p. 83–101.

[AS94] S. Aida & D. Stroock – « Moment estimates derived from Poincaré and logarithmic Sobolev inequalities », *Math. Res. Lett.* **1** (1994), no. 1, p. 75–86.

[Aub82] T. Aubin – *Nonlinear analysis on manifolds. Monge-Ampère equations*, Springer-Verlag, New York, 1982.

[Bak94] D. Bakry – « L'hypercontractivité et son utilisation en théorie des semigroupes », Lectures on probability theory (Saint-Flour, 1992), Springer, Berlin, 1994, p. 1–114.

[Bak96] ———, « Remarques sur les semigroupes de Jacobi », *Astérisque* (1996), no. 236, p. 23–39, Hommage à P. A. Meyer et J. Neveu.

[Bak97] ———, « On Sobolev and logarithmic Sobolev inequalities for Markov semigroups », New trends in stochastic analysis (Charingworth, 1994), World Sci. Publishing, River Edge, NJ, 1997, p. 43–75.

[BCL97] D. Bakry, D. Concordet & M. Ledoux – « Optimal heat kernel bounds under logarithmic Sobolev inequalities », *ESAIM Probab. Statist.* **1** (1995/97), p. 391–407 (electronic).

[BÉ84] D. Bakry & M. Émery – « Hypercontractivité de semi-groupes de diffusion », *C. R. Acad. Sci. Paris Sér. I Math.* **299** (1984), no. 15, p. 775–778.

[BÉ85a] ———, « Diffusions hypercontractives », Séminaire de probabilités, XIX, 1983/84, Springer, Berlin, 1985, p. 177–206.

[BÉ85b] ———, « Inégalités de Sobolev pour un semi-groupe symétrique », *C. R. Acad. Sci. Paris Sér. I Math.* **301** (1985), no. 8, p. 411–413.

[BG99] S. G. Bobkov & F. Götze – « Exponential integrability and transportation cost related to logarithmic Sobolev inequalities », *J. Funct. Anal.* **163** (1999), no. 1, p. 1–28.

[BH97] S. G. Bobkov & C. Houdré – « Isoperimetric constants for product probability measures », *Ann. Probab.* **25** (1997), no. 1, p. 184–205.

[BH99] T. Bodineau & B. Helffer – « The log-Sobolev inequality for unbounded spin systems », *J. Funct. Anal.* **166** (1999), no. 1, p. 168–178.

[BL96a] D. Bakry & M. Ledoux – « Lévy-Gromov's isoperimetric inequality for an infinite-dimensional diffusion generator », *Invent. Math.* **123** (1996), no. 2, p. 259–281.

[BL96b] ———, « Myer's theorem and sobolev inequalities », *Duke Math. J.* **85** (1996), no. 1, p. 253–270.

[BM92] D. Bakry & D. Michel – « Sur les inégalités FKG », Séminaire de Probabilités, XXVI, Springer, Berlin, 1992, p. 170–188.

[BM97] L. Birgé & P. Massart – « From model selection to adaptive estimation », Festschrift for Lucien Le Cam, Springer, New York, 1997, p. 55–87.

[BM98] ———, « Minimum contrast estimators on sieves : exponential bounds and rates of convergence », *Bernoulli* **4** (1998), no. 3, p. 329–375.

[BM00] F. Barthe & B. Maurey – « Some remarks on isoperimetry of Gaussian type », *Ann. Inst. H. Poincaré Probab. Statist.* **36** (2000), no. 4, p. 419–434.

[Bob96a] S. G. Bobkov – « Extremal properties of half-spaces for log-concave distributions », *Ann. Probab.* **24** (1996), no. 1, p. 35–48.

[Bob96b] ———, « A functional form of the isoperimetric inequality for the Gaussian measure », *J. Funct. Anal.* **135** (1996), no. 1, p. 39–49.

[Bob97] ———, « An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space », *Ann. Probab.* **25** (1997), no. 1, p. 206–214.

[Bob99] ———, « Isoperimetric and analytic inequalities for log-concave probability measures », *Ann. Probab.* **27** (1999), no. 4, p. 1903–1921.

[Bré83] H. Brézis – *Analyse fonctionnelle*, Masson, Paris, 1983, Théorie et applications.

[Che70] J. Cheeger – « A lower bound for the smallest eigenvalue of the Laplacian », Problems in analysis (Papers dedicated to Salomon Bochner, 1969), Princeton Univ. Press, Princeton, N. J., 1970, p. 195–199.

[CS86] E. A. Carlen & D. W. Stroock – « An application of the Bakry-Emery criterion to infinite-dimensional diffusions », Séminaire de Probabilités, XX, 1984/85, Springer, Berlin, 1986, p. 341–348.

[DS84] E. B. Davies & B. Simon – « Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians », *J. Funct. Anal.* **59** (1984), no. 2, p. 335–395.

[DSC96] P. Diaconis & L. Saloff-Coste – « Logarithmic Sobolev inequalities for finite Markov chains », *Ann. Appl. Probab.* **6** (1996), no. 3, p. 695–750.

[Dvo61] A. Dvoretzky – « Some results on convex bodies and Banach spaces », *Proc. Symp. on Linear Spaces* (Jerusalem), 1961, p. 123–160.

[EK86] S. N. Ethier & T. G. Kurtz – *Markov processes*, John Wiley & Sons Inc., New York, 1986, Characterization and convergence.

[Ell85] R. S. Ellis – *Entropy, large deviations, and statistical mechanics*, Springer-Verlag, New York, 1985.

[Fou00] P. Fougères – « Hypercontractivité et isopérimétrie gaussienne. Applications aux systèmes de spins », *Ann. Inst. H. Poincaré Probab. Statist.* **36** (2000), no. 5, p. 647–689.

[Geo88] H.-O. Georgii – *Gibbs measures and phase transitions*, Walter de Gruyter & Co., Berlin, 1988.

[GM83] M. Gromov & V. D. Milman – « A topological application of the isoperimetric inequality », *Amer. J. Math.* **105** (1983), no. 4, p. 843–854.

[GR98] L. Gross & O. Rothaus – « Herbst inequalities for supercontractive semigroups », *J. Math. Kyoto Univ.* **38** (1998), no. 2, p. 295–318.

[Gro75] L. Gross – « Logarithmic Sobolev inequalities », *Amer. J. Math.* **97** (1975), no. 4, p. 1061–1083.

[GW00] F. Gong & L. Wu – « Spectral gap of positive operators and applications », *C. R. Acad. Sci. Paris Sér. I Math.* **331** (2000), no. 12, p. 983–988.

[GZ00] A. Guionnet & B. Zegarliński – « Lectures on logarithmic Sobolev inequalities », A paraître aux Séminaires de Probabilités, 2000.

[Heb99] E. Hebey – « Meilleures constantes et inégalités de Sobolev optimales sur les variétés riemanniennes compactes », Séminaire de Théorie Spectrale et Géométrie, Vol. 16, Année 1997–1998, Univ. Grenoble I, Saint, 199 ?, p. 175–210.

[Heb97] ———, *Introduction à l'analyse non linéaire sur les variétés*, Diderot Editeur, Arts et Sciences, Paris, 1997.

[Hin00] M. Hino – « Exponential decay of positivity preserving semigroups on L^p », *Osaka J. Math.* **37** (2000), no. 3, p. 603–624.

[HS87] R. A. Holley & D. W. Stroock – « Logarithmic Sobolev inequalities and stochastic Ising models », *J. Statist. Phys.* **46** (1987), no. 5-6, p. 1159–1194.

[HS89] ———, « Uniform and L^2 convergence in one-dimensional stochastic Ising models », *Comm. Math. Phys.* **123** (1989), no. 1, p. 85–93.

[Lar95] É. Laroche – « Hypercontractivité pour des systèmes de spins de portée infinie », *Probab. Theory Related Fields* **101** (1995), no. 1, p. 89–132.

[Led96] M. Ledoux – « Isoperimetry and Gaussian analysis », Lectures on probability theory and statistics (Saint-Flour, 1994), Springer, Berlin, 1996, p. 165–294.

[Led99] ———, « Concentration of measure and logarithmic Sobolev inequalities », Séminaire de Probabilités, XXXIII, Springer, Berlin, 1999, p. 120–216.

[Led00] ———, « The geometry of markov diffusion generators », *-Ann. Fac. Sci. Toulouse* **IX** (2000), p. 305–366.

[Led01] ———, « Logarithmic Sobolev inequalities for unbounded spin systems revisited », Séminaire de Probabilités, XXXV, Springer, Berlin, 2001, p. 167–194.

[Led97] ———, « On Talagrand's deviation inequalities for product measures », *ESAIM Probab. Statist.* **1** (1995/97), p. 63–87 (electronic).

[LO00] R. Latala & K. Oleszkiewicz – « Between Sobolev and Poincaré », Geometric aspects of functional analysis, Springer, Berlin, 2000, p. 147–168.

[LT91] M. Ledoux & M. Talagrand – *Probability in Banach spaces, isoperimetry and processes*, Springer-Verlag, Berlin, 1991.

[LY93] S. L. Lu & H.-T. Yau – « Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics », *Comm. Math. Phys.* **156** (1993), no. 2, p. 399–433.

[Mal01] F. Malrieu – « Logarithmic Sobolev inequalities for some nonlinear PDE's », *Stochastic Process. Appl.* **95** (2001), no. 1, p. 109–132.

[Mar96] K. Marton – « Bounding $\overline{d}$-distance by informational divergence : a method to prove measure concentration », *Ann. Probab.* **24** (1996), no. 2, p. 857–866.

[Maz02] O. Mazet – « A characterization of Markov property for semigroups with invariant measure », *Potential Anal.* **16** (2002), no. 3, p. 279–287.

[Mic99a] L. Miclo – « An example of application of discrete Hardy's inequalities », *Markov Process. Related Fields* **5** (1999), p. 319–330.

[Mic99b] ———, « Relations entre isopérimetrie et trou spectral pour les chaînes de Markov finies », *Probab. Theor. Relat. Fields* **114** (1999), p. 431–485.

[Mil71] V. D. Milman – « A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies », *Funkcional. Anal. i Priložen.* **5** (1971), no. 4, p. 28–37.

[Mil92] ———, « Dvoretzky's theorem—thirty years later », *Geom. Funct. Anal.* **2** (1992), no. 4, p. 455–479.

[MO94a] F. Martinelli & E. Olivieri – « Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case », *Comm. Math. Phys.* **161** (1994), no. 3, p. 447–486.

[MO94b] ———, « Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case », *Comm. Math. Phys.* **161** (1994), no. 3, p. 487–514.

[Mos64] J. Moser – « A Harnack inequality for parabolic differential equations », *Comm. Pure Appl. Math.* **17** (1964), p. 101–134.

[MS86] V. D. Milman & G. Schechtman – *Asymptotic theory of finite-dimensional normed spaces*, Springer-Verlag, Berlin, 1986, With an appendix by M. Gromov.

[Muc72] B. Muckenhoupt – « Hardy's inequality with weights », *Studia Math.* **44** (1972), p. 31–38, collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, **I**.

[Nas58] J. Nash – « Continuity of solutions of parabolic and elliptic equations », *Amer. J. Math.* **80** (1958), p. 931–954.

[Roy99] G. Royer – *Une initiation aux inégalités de Sobolev logarithmiques*, Société Mathématique de France, Paris, 1999.

[RS78] M. Reed & B. Simon – *Methods of modern mathematical physics. IV. Analysis of operators*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1978.

[Sch98] M. Schmuckenschläger – « Martingales, Poincaré type inequalities, and deviation inequalities », *J. Funct. Anal.* **155** (1998), no. 2, p. 303–323.

[Sch01] G. Scheffer – « Inégalités fonctionnelles, géométrie conforme et noyaux markoviens », Thèse soutenue sous la direction de D.Bakry à l'université Paul Sabatier, 2001.

[SZ92a] D. W. Stroock & B. Zegarliński – « The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition », *Comm. Math. Phys.* **144** (1992), no. 2, p. 303–323.

[SZ92b] ———, « The logarithmic Sobolev inequality for continuous spin systems on a lattice », *J. Funct. Anal.* **104** (1992), no. 2, p. 299–326.

[SZ95] ———, « On the ergodic properties of Glauber dynamics », *J. Statist. Phys.* **81** (1995), no. 5-6, p. 1007–1019.

[Tal88] M. Talagrand – « An isoperimetric theorem on the cube and the Kintchine-Kahane inequalities », *Proc. Amer. Math. Soc.* **104** (1988), no. 3, p. 905–909.

[Tal91] ———, « A new isoperimetric inequality and the concentration of measure phenomenon », Geometric aspects of functional analysis (1989–90), Springer, Berlin, 1991, p. 94–124.

[Tal95] ———, « Concentration of measure and isoperimetric inequalities in product spaces », *Inst. Hautes Études Sci. Publ. Math.* (1995), no. 81, p. 73–205.

[Tal96a] ———, « New concentration inequalities in product spaces », *Invent. Math.* **126** (1996), no. 3, p. 505–563.

[Tal96b] ———, « A new look at independence », *Ann. Probab.* **24** (1996), no. 1, p. 1–34.

[Tal96c] ———, « Transportation cost for Gaussian and other product measures », *Geom. Funct. Anal.* **6** (1996), no. 3, p. 587–600.

[Var84] N. T. Varopoulos – « Une généralisation du théorème de Hardy-Littlewood-Sobolev pour les espaces de Dirichlet », *C. R. Acad. Sci. Paris Sér. I Math.* **299** (1984), no. 14, p. 651–654.

[Var85] ———, « Hardy-Littlewood theory for semigroups », *J. Funct. Anal.* **63** (1985), no. 2, p. 240–260.

[Wan97] F.-Y. Wang – « Logarithmic Sobolev inequalities on noncompact Riemannian manifolds », *Probab. Theor. Relat. Fields* **109** (1997), no. 3, p. 417–424.

[Yos80] K. Yosida – *Functional analysis*, 6ème éd., Springer-Verlag, Berlin, 1980.

[Yos01] N. Yoshida – « The equivalence of the log-Sobolev inequality and a mixing condition for unbounded spin systems on the lattice », *Ann. Inst. H. Poincaré Probab. Statist.* **37** (2001), no. 2, p. 223–243.

[Zeg90] B. Zegarliński – « Log-Sobolev inequalities for infinite one-dimensional lattice systems », *Comm. Math. Phys.* **133** (1990), no. 1, p. 147–162.

[Zeg92] ———, « Dobrushin uniqueness theorem and logarithmic Sobolev inequalities », *J. Funct. Anal.* **105** (1992), no. 1, p. 77–111.

[Zeg96] ———, « The strong decay to equilibrium for the stochastic dynamics of unbounded spin systems on a lattice », *Comm. Math. Phys.* **175** (1996), no. 2, p. 401–432.

[Zeg01] ———, « Isoperimetry for Gibbs measures », *Ann. Probab.* **29** (2001), no. 2, p. 802–819.

Functional inequalities involving DIRICHLET forms.
From isoperimetry to SOBOLEV inequalities.

Abstract

Ergodic MARKOV semigroups provide approximations of probability measures by means of some functional inequalities. The aim of the thesis is to investigate some of these inequalities, from isoperimetry to SOBOLEV inequalities. What we are essentially interested in is to establish some links between these inequalities, to determine their optimal constants and to get some criteria which ensure their existence. Three main issues were investigated.

Firstly it is shown that for diffusion semigroups whose curvature is bounded from below (by a possibly negative number) a logarithmic SOBOLEV inequality implies a BOBKOV Gaussian isoperimetric inequality. We further establish a weakened functional form of the BOBKOV inequality which is satisfied by any hypercontractive jump semigroup on a finite set. As an application, some conditions on the interaction potential of a spins system to control the curvature of the associated GLAUBER dynamic are also described. In this way, we get isoperimetric inequalities for spins systems satisfying the DOBRUSHIN and SCHLOSMAN mixing condition.

In the second part, we propose a simple parameter to describe the exact order of the POINCARÉ constant (or the inverse of the spectral gap) for a log-concave probability measure on the real line. This parameter is the square of the mean value of the distance to the median. BOBKOV (1999) derives a similar result in terms of the variance of the measure. His approach was based on the study of the CHEEGER constant. Our viewpoint is quite different and makes use of the MUCKENHOUPT functional and of a variation computation in the set of convex functions.

The final part of the work is devoted to establishing new criteria for SOBOLEV inequalities when the Bakry-Emery (or curvature-dimension) criterion fails. We investigate the behaviour (or stability properties) of the curvature-dimension inequality for diffusion operators under two different notions of projection. Contrary to the classical operator projection, we may always project a DIRICHLET structure. However, though the curvature-dimension inequality is preserved by an operator projection, this is shown to be wrong for DIRICHLET structure. We describe in detail an example in spherical geometry in which these projection arguments give rise to a SOBOLEV inequality even if no curvature-dimension inequality holds. We then present a new way to prove SOBOLEV inequalities based on the construction of a DIRICHLET structure in a higher dimension by means of conformal changes of metric. This latter lifting must project onto the initial structure in the lower dimension.

Keywords : SOBOLEV inequality. Logarithmic SOBOLEV inequality. Functional isoperimetric inequalities. POINCARÉ inequality. Spectral gap. MARKOV semigroups. DIRICHLET forms. Statistical mechanics. Log-concave measures.

Field : Mathematics, Probability Theory.

Auteur : Pierre Fougères.
Titre : Inégalités fonctionnelles liées aux formes de DIRICHLET.
De l'isopérimétrie aux inégalités de SOBOLEV.

Thèse soutenue le 18 Octobre 2002 à l'Université Paul SABATIER (Toulouse III).

Résumé : Les semi-groupes de MARKOV ergodiques permettent d'approcher des mesures de probabilités au moyen d'inégalités fonctionnelles. L'objectif de la thèse est l'étude de certaines de ces inégalités, de l'isopérimétrie gaussienne aux inégalités de SOBOLEV. Nous cherchons essentiellement à établir des liens entre elles, à déterminer leurs constantes optimales et à obtenir des critères assurant leur existence.

Dans une première partie, nous montrons que, pour un semi-groupe diffusif de courbure minorée (éventuellement négative), une inégalité de SOBOLEV logarithmique induit une inégalité isopérimétrique gaussienne de BOBKOV. En outre, nous établissons une forme fonctionnelle affaiblie de l'inégalité de BOBKOV pour tout semi-groupe hypercontractif à sauts sur un ensemble fini. Enfin, nous donnons des critères sur le potentiel d'interaction d'un système de spins permettant de contrôler la courbure de la dynamique de GLAUBER. Nous obtenons ainsi les inégalités isopérimétriques correspondantes sous la condition de mélange de DOBRUSHIN et SCHLOSMAN.

La seconde partie est consacrée à la détermination de l'ordre de grandeur exact du trou spectral d'une mesure log-concave sur la droite réelle. Le bon paramètre décrivant cette constante est le carré de la distance moyenne à la médiane. S.BOBKOV a obtenu en 1999 un résultat similaire en termes de la variance de la mesure par une approche basée sur l'étude de la constante de CHEEGER. Notre point de vue repose sur l'utilisation de la fonctionnelle de MUCKENHOUPT et sur un calcul de variations dans l'ensemble des fonctions convexes.

Dans la dernière partie, nous établissons de nouveaux critères conduisant aux inégalités de SOBOLEV lorsque le critère de courbure-dimension de BAKRY et EMERY est mis en défaut. Contrairement aux opérateurs de diffusion, il est toujours possible de projeter une structure de DIRICHLET. La remarque fondamentale à l'origine de ce travail est que certaines de ces structures vérifiant le critère de courbure-dimension se projettent sur des structures qui ne le vérifient pas. Après avoir décrit en détails un exemple simple en géométrie sphérique qui servira de guide, nous donnons un critère général assurant l'existence d'une structure de DIRICHLET satisfaisant le critère de BAKRY et EMERY qui se projette sur la structure de départ.

Mots clés : Inégalité de SOBOLEV. Inégalité de SOBOLEV logarithmique. Inégalités isopérimétriques fonctionnelles. Inégalité de POINCARÉ. Trou spectral. Semi-groupes de MARKOV. Formes de DIRICHLET. Mécanique statistique. Mesures log-concaves.
Discipline : Mathématiques, Probabilités.

Laboratoire de Statistique et Probabilités
UFR MIG - UMR CNRS 5883
Université Paul Sabatier, Toulouse III