Inégalités fonctionnelles, transport optimal et grandes déviations

Nathael Gozlan

Université Paris Est Marne-la-Vallée
Laboratoire d'Analyse et de Mathématiques Appliquées (UMR 8050)

Inégalités fonctionnelles, transport optimal et grandes déviations

NATHAEL GOZLAN

Synthèse des travaux en vue de l'obtention de
l'**Habilitation à Diriger des Recherches**

Habilitation à Diriger des Recherches soutenue le **11 décembre 2012** devant le jury composé de

Patrick CATTIAUX,	Université Toulouse 3
Djalil CHAFAI,	Université Marne-la-Vallée
Francis COMETS,	Université Paris 7
Arnaud GUILLIN,	Université Clermont Ferrand
Michel LEDOUX,	Université Toulouse 3
Florence MERLEVÈDE,	Université Marne-la-Vallée
Karl Theodor STURM,	Université de Bonn
Cédric VILLANI,	Université Lyon 1 et Institut Henri Poincaré

au vu des rapports de

Sergey BOBKOV,	Université du Minesota
Michel LEDOUX,	Université Toulouse 3
Karl Theodor STURM,	Université de Bonn

Remerciements

Je tiens en premier lieu à remercier Patrick Cattiaux qui a dirigé ma thèse de doctorat et qui m'a fait découvrir le monde des inégalités fonctionnelles. Je remercie également Christian Léonard avec qui j'ai eu beaucoup de plaisir à collaborer et à échanger et qui comme Patrick a contribué à ouvrir mon imaginaire mathématique.

Je remercie le « groupe de travail du jeudi à l'IHP ». Il s'est d'abord appelé IFO, puis EVOL, maintenant C'TOP ; sous ces appellations changeantes il a su fédéré avec constance un groupe de chercheurs autour de questions portant sur les inégalités géométriques et fonctionnelles, les équations d'évolution ou la courbure des espaces métriques. J'ai beaucoup appris lors de ces séances d'exposés et j'exprime ma reconnaissance à tous ceux qui ont participé activement à ce groupe.

Je remercie Sergey Bobkov, Michel Ledoux et Karl-Theodor Sturm d'avoir accepté d'être rapporteurs de ce mémoire et Patrick Cattiaux, Djalil Chafaï, Francis Comets, Arnaud Guillin, Florence Merlevède et Cédric Villani de me faire l'honneur de participer à mon jury.

Je remercie mes coauteurs Patrick Cattiaux, Arnaud Guillin, Christian Léonard, Cyril Roberto, Paul-Marie Samson et Prasad Tetali pour les travaux que nous avons menés à bien et pour ce que ces collaborations m'ont apporté sur le plan scientifique et personnel.

Je remercie mes collègues de Marne-la-Vallée qui par leurs qualités humaines font du laboratoire de mathématiques un lieu de travail et d'échanges extrêmement agréable. Je remercie Christiane Lafargue et Audrey Patout pour leur aide précieuse.

Je remercie enfin ma famille, mes amis et Laurence qui partage ma vie.

Table des matières

Bibliographie

[1] N. Gozlan. Conditional principles for random weighted measures. *ESAIM Probab. Stat.*, 9 :283–306 (electronic), 2005.

[2] N. Gozlan. Integral criteria for transportation-cost inequalities. *Electron. Comm. Probab.*, 11 :64–77 (electronic), 2006.

[3] P. Cattiaux and N. Gozlan. Deviations bounds and conditional principles for thin sets. *Stochastic Process. Appl.*, 117(2) :221–250, 2007.

[4] N. Gozlan. Characterization of Talagrand's like transportation-cost inequalities on the real line. *J. Funct. Anal.*, 250(2) :400–425, 2007.

[5] N. Gozlan and C. Léonard. A large deviation approach to some transportation cost inequalities. *Probab. Theory Related Fields*, 139(1-2) :235–283, 2007.

[6] N. Gozlan. A characterization of dimension free concentration in terms of transport inequalities. *Ann. Probab.*, 37(6) :2480–2498, 2009.

[7] P. Cattiaux, N. Gozlan, A. Guillin, and C. Roberto. Functional inequalities for heavy tailed distributions and application to isoperimetry. *Electron. J. Probab.*, 15 :no. 13, 346–385, 2010.

[8] N. Gozlan. Poincaré inequalities and dimension free concentration of measure. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(3) :708–739, 2010.

[9] N. Gozlan and C. Léonard. Transport inequalities. A survey. *Markov Process. Related Fields*, 16(4) :635–736, 2010.

[10] N. Gozlan, C. Roberto, and P. M. Samson. Isoperimetry for product of heavy tails distributions. In *Progress in analysis and its applications*, pages 470–478. World Sci. Publ., Hackensack, NJ, 2010.

[11] N. Gozlan, C. Roberto, and P-M Samson. From concentration to logarithmic Sobolev and Poincaré inequalities. *J. Funct. Anal.*, 260(5) :1491–1522, 2011.

[12] N. Gozlan, C. Roberto, and P. M. Samson. A new characterization of Talagrand's transport-entropy inequalities and applications. *Ann. Probab.*, 39(3) :857–880, 2011.

[13] N. Gozlan, C. Roberto, and P.M. Samson. Characterization of Talagrand's transport-entropy inequalities on metric spaces. A paraître dans Ann. Probab., 2012.

[14] N. Gozlan. Transport-entropy inequalities on the line. *Electron. J. Probab.*, 17(49) :1–18, 2012.

[15] N. Gozlan, C. Roberto, and P.M. Samson. Hamilton-Jacobi equations on metric spaces and transport-entropy inequalities. A paraître dans Rev. Mat. Iberoam. 2012.

Avant propos

Ce document présente les recherches que j'ai effectuées depuis mon doctorat dans le domaine des grandes déviations et des inégalités fonctionnelles. Ces travaux ont donné lieu aux publications référencées ci-dessus, dont voici un rapide descriptif thématique :

– Les articles [1], [3] et [5] font usage de techniques issues de la *théorie des grandes déviations* pour établir des lois des grands nombres conditionnelles [1, 3] et pour apporter une interprétation probabiliste des inégalités de transport [5].

– Les autres articles de la liste abordent des questions ayant trait aux inégalités fonctionnelles.

  – Les travaux [2], [4], [5], [6], [8], [12], [13], [14], [15] portent sur l'étude des *inégalités de transport*.
  Cette étude s'est déployée dans quatre directions principales :

    1. la recherche de critères (conditions nécessaires et/ou suffisantes) pour tel ou tel type d'inégalités de transport [2, 4, 8, 14].
    2. la compréhension des liens entre ces inégalités et la théorie des grandes déviations [5], déjà mentionnée ci-dessus.
    3. la caractérisation du phénomène de la concentration de la mesure adimensionnel par ces inégalités [6].
    4. l'étude des liens avec les inégalités du type Sobolev logarithmique [6, 12, 13, 15] et la perturbation de ces inégalités.

  – L'article [11] étudie les *inégalités de Poincaré et de Sobolev logarithmiques* sur des espaces métriques dont la courbure est minorée au sens de Lott-Villani-Sturm.

  – Les travaux [7] et [10] ont quant à eux pour objet l'étude des *inégalités isopérimétriques* pour des mesures de probabilité à queues lourdes.

Introduction

Ce document fait la synthèse des différents travaux de recherche que j'ai entrepris seul ou en collaboration depuis ma thèse, soutenue en 2005. Mes recherches ont commencé avec des questions relevant des grandes déviations, et se sont par la suite portées sur l'étude des inégalités fonctionnelles.

Dans le premier chapitre, j'aborde succinctement les travaux [1, 3] réalisés au début de ma thèse dans le domaine des grandes déviations. Ils portent sur l'étude des lois des grands nombres conditionnelles. Le principe de Gibbs, qui en est l'exemple le plus typique, peut s'énoncer ainsi : si N particules indépendantes et de même loi μ sont trouvées dans un état inhabituel, s'exprimant par la réalisation d'une contrainte convexe C par la mesure empirique du nuage de particules, alors la loi d'une particule typique est, quand N est très grand, proche de la loi de probabilité μ^* minimisant l'entropie relative par rapport à μ sous la contrainte considérée. Rappelons que l'entropie relative d'une probabilité ν absolument continue par rapport à μ est définie par

$$H(\nu|\mu) = \int \log \frac{d\nu}{d\mu}\, d\nu.$$

La théorie des grandes déviations permet de rendre rigoureux le principe précédent, comme l'ont montré par exemple Csiszar [Csi84] ou Stroock-Zeitouni [SZ91]. En effet, d'après le théorème de Sanov [San61], l'entropie relative par rapport à μ gouverne les grandes déviations de la mesure empirique $L_n = n^{-1}\sum_{i=1}^n \delta_{X_i}$, où X_i est une suite i.i.d de loi μ. Cela se traduit informellement par

$$\mathbb{P}(L_n \in A) \simeq e^{-n \inf_{\nu\in A} H(\nu|\mu)}, \qquad \text{lorsque } n \to \infty.$$

En restant toujours à un niveau informel, on voit par cette estimation asymptotique que $\mathbb{P}(L_n \in C) \simeq \mathbb{P}(L_n \in B)$ pour tout sous ensemble $B \subset C$ contenant le minimisant μ^* de $H(\,\cdot\,|\mu)$ sur le convexe C. On en conclue qui si L_n appartient à C alors L_n est proche de μ^*, ce qui corrobore le principe de Gibbs.

Plus généralement, on montre que si une suite de variables aléatoires Z_n suit un principe de grandes déviations de fonction de taux I, alors la loi conditionnelle de Z_n sous la contrainte $Z_n \in C$, C étant un ensemble ayant de bonnes propriétés, se concentre généralement autour des minimisants de la fonction de taux I sur C.

La question qui nous a intéressés, P. Cattiaux et moi même, dans [3] était d'étendre le principe de Gibbs aux conditionnements fins, c'est à dire aux ensembles C pour lesquels $\mathbb{P}(L_n \in C) = 0$. Notre approche consiste à grossir la contrainte C pour la rendre réalisable par L_n, en faisant décroitre la taille du grossissement à mesure que n augmente. En clair, nous étudions la loi de L_n sachant que $L_n \in C_{\varepsilon_n}$, où $(C_\varepsilon)_{\varepsilon>0}$ est une famille croissante d'ensembles d'intersection C. La question est maintenant de déterminer des vitesses de rétrécissement $\varepsilon_n \downarrow 0$ pour lesquelles la loi conditionnelle se concentre encore sur les minimisants de la fonction de taux sur C. Ce problème nous fait sortir du cadre strict des grandes déviations. En effet, pour pouvoir ajuster la taille du grossissement ε au nombre de particules n, on doit être en mesure de contrôler de manière non-asymptotique les probabilités d'événements rares $\mathbb{P}(L_n \in C_\varepsilon)$, en fonction de ε et de n. C'est du côté des inégalités de concentration de type Bernstein que nous avons trouvé l'outil adéquat permettant cette estimation non-asymptotique.

Dans l'article [1], qui aborde une variante du principe de Gibbs, le petit problème suivant s'est posé et a joué un rôle clé dans la preuve du résultat principal : contrôler un écart entre les moyennes de deux probabilités par l'entropie relative de l'une par rapport à l'autre. Plus précisément, la question était de savoir sous quelles conditions une probabilité μ sur $\mathbb{R}$ vérifiait l'inégalité

$$\alpha\left(\left|\int x\,d\nu - \int x\,d\mu\right|\right) \leqslant H(\nu|\mu),$$

pour toute probabilité ν sur $\mathbb{R}$, où α est une fonction convexe sur $\mathbb{R}^+$ nulle en 0. C'est cette question qui m'a amené à m'intéresser aux inégalités de transport, sous l'angle des grandes déviations. Si μ a son support dans l'intervalle $[-1,1]$, on peut faire appel à l'inégalité classique de Csiszar-Kullback-Pinsker (CKP) [Csi67, Kul67, Pin64] :

$$\|\nu-\mu\|_{VT} = \sup_{|f|\leqslant 1}\left\{\int f\,d\nu - \int f\,d\mu\right\} \leqslant \sqrt{2H(\nu|\mu)},$$

pour toutes probabilités μ,ν. En revanche, si le support de μ n'est pas borné, l'inégalité CKP ne suffit plus, et le problème posé relève des inégalités de transport pour la distance de Wasserstein W_1, dont l'exemple emblématique est l'inégalité $\mathbf{T}_1$ introduite par Bobkov et Götze dans [BG99]. Rappelons qu'une probabilité μ sur un espace métrique $(\mathcal{X}, d)$ vérifie l'inégalité de transport $\mathbf{T}_1(C)$ si pour toute probabilité ν

$$W_1(\nu,\mu) = \sup_{f\ 1-\mathrm{Lip}}\left\{\int f\,d\nu - \int f\,d\mu\right\} \leqslant \sqrt{CH(\nu|\mu)}.$$

L'inégalité $\mathbf{T}_1$ est vérifiée si et seulement si μ a un moment exponentiel d'ordre 2, comme l'ont montré Djellout-Guillin-Wu [DGW04].

Les chapitres 2, 3, 4 et 5 présentent les résultats que j'ai obtenus seul ou avec mes co-auteurs (C. Léonard, C. Roberto, P-M Samson) dans le domaine des inégalités de transport.

Le chapitre 2 est une introduction regroupant les principales définitions et les propriétés de base des inégalités de transport. Ces résultats (de tensorisation, de dualité, ou de concentration) sont principalement issus des articles pionniers de Marton [Mar86], Talagrand [Tal96] et Bobkov-Götze [BG99]. Le point principal développé dans ce chapitre est le fait important, découvert par Marton et Talagrand, que les inégalités de transport permettent de démontrer des inégalités de concentration de la mesure. Dans tout le manuscrit, nous accorderons une place particulière à deux inégalités de transport très étudiées ces dernières années : l'inégalité $\mathbf{T}_1$ déjà présentée et l'inégalité $\mathbf{T}_2$ introduite par Talagrand [Tal96] qui est définie comme suit : une probabilité μ sur $(\mathcal{X}, d)$ vérifie $\mathbf{T}_2(C)$ si pour toute probabilité ν sur $\mathcal{X}$, on a

$$W_2(\nu, \mu) \leqslant \sqrt{CH(\nu|\mu)}.$$

La distance W_2 est définie par la formule $W_2^2(\nu, \mu) = \inf \mathbb{E}[d(X, Y)^2]$ où l'inf porte sur tous les couples de variables aléatoires (X, Y) avec X de loi ν et Y de loi μ. Les inégalités $\mathbf{T}_1$ et $\mathbf{T}_2$ sont associées à un phénomène de concentration gaussien. On montre en effet que si μ vérifie $\mathbf{T}_1(C)$, alors pour tout $n \in \mathbb{N}^*$, la mesure produit μ^n vérifie

$$\mu^n(f > m(f) + t) \leqslant e^{-\frac{t^2}{nC}}, \qquad \forall t \geqslant 0,$$

pour toute fonction $f : \mathcal{X}^n \to \mathbb{R}$ 1-Lipschitz pour la distance $d_1(x, y) = \sum_{i=1}^n d(x_i, y_i)$. Si μ vérifie l'inégalité $\mathbf{T}_2(C)$, qui est plus forte que $\mathbf{T}_1(C)$, alors pour tout $n \in \mathbb{N}^*$, la mesure produit μ^n vérifie une propriété de concentration beaucoup plus forte :

$$\mu^n(f > m(f) + t) \leqslant e^{-\frac{t^2}{C}}, \qquad \forall t \geqslant 0, \tag{1}$$

pour toute fonction $f : \mathcal{X}^n \to \mathbb{R}$ 1-Lipschitz pour la distance $d_2(x, y) = \left[\sum_{i=1}^n d^2(x_i, y_i)\right]^{1/2}$. Ce type de concentration adimensionnel a trouvé de nombreuses applications en probabilité, en analyse, ou en statistique (voir par exemple le livre de Ledoux [Led01]).

Le chapitre 3 présente les résultats des articles [2, 4, 8, 14] dont le but est de donner des conditions simples à vérifier permettant de déterminer si une probabilité vérifie une inégalité de transport donnée. Tout d'abord, à la suite de Djellout-Guillin-Wu [DGW04] et Bolley-Villani [BV05], nous montrons qu'une grande classe d'inégalités de transport, incluant l'inégalité $\mathbf{T}_1$, est équivalente à l'existence d'un moment exponentiel (d'ordre 2 dans le cas de $\mathbf{T}_1$) fini. Cette caractérisation ne concerne pas l'inégalité $\mathbf{T}_2$ qui s'avère plus complexe. Dans [4, 14], nous donnons néanmoins une condition nécessaire et suffisante en dimension un : une probabilité sur la droite réelle vérifie l'inégalité $\mathbf{T}_2$ si et seulement si elle vérifie une inégalité de trou spectral et si elle est obtenue en déformant la mesure exponentielle symétrique $\mu_1(dx) = e^{-|x|}\,dx/2$ par une transformation croissante dont le module de continuité est en $O(\sqrt{u})$ pour u grand. En dimension quelconque, une condition suffisante portant sur le log de la densité de la probabilité de référence est donnée dans [8]. Cette condition généralise à la dimension quelconque un résultat de Cattiaux et Guillin [CG06] valable en dimension un.

Le chapitre 4 donne une interprétation des inégalités de transport par les grandes déviations [5] (écrit avec C. Léonard) et précisent leurs liens avec la concentration [6]. Une idée développée dans ces deux articles est que les inégalités de transport sont équivalentes à une estimation de grandes déviations. Par exemple, l'inégalité $\mathbf{T}_p(C)$ avec $p = 1$ ou 2 est satisfaite si et seulement si

$$\limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{P}(W_p(L_n, \mu) > t) \leqslant -t^2/C, \qquad \forall t \geqslant 0.$$

Cette idée joue un rôle pivot dans l'article [6] dont le résultat principal est qu'une probabilité μ vérifie le phénomène de concentration gaussien adimensionnel (1) si et seulement si elle vérifie l'inégalité $\mathbf{T}_2(C)$. Ce résultat montre que l'inégalité $\mathbf{T}_2$ est la forme fonctionnelle canonique de la concentration gaussienne. Une preuve plus analytique, obtenue dans [11] (écrit avec C. Roberto et P-M Samson) est également présentée dans ce chapitre.

Le chapitre 5 analyse les liens entre l'inégalité $\mathbf{T}_2$ et l'inégalité de Sobolev logarithmique **LSI**. Une probabilité μ (sur $\mathbb{R}^k$ pour fixer les idées) vérifie $\mathbf{LSI}(C)$ si pour toute fonction $f : \mathbb{R}^k \to \mathbb{R}$

suffisamment régulière,

$$\mathrm{Ent}_\mu(f^2) := \int \log\left(\frac{f^2}{\int f^2\, d\mu}\right) f^2\, d\mu \leqslant C \int \|\nabla f\|_2^2\, d\mu.$$

Cette inégalité fonctionnelle introduite par Gross dans [Gro75] est bien connue pour ses applications dans l'étude des semigroupes de Markov (hypercontractivité, convergence vers l'équilibre). Otto et Villani ont montré dans [OV00] que l'inégalité de Sobolev logarithmique **LSI** entraînait l'inégalité de Talagrand $\mathbf{T}_2$, avec la même constante C. Dans l'article [6], nous donnons une nouvelle preuve de cette implication fondée sur la concentration. Un argument dû à Herbst et popularisé notamment par Ledoux [Led01] montre en effet que si μ vérifie l'inégalité $\mathbf{LSI}(C)$, alors elle vérifie la propriété de concentration gaussienne adimensionnelle (1). Grâce à la caractérisation de ce phénomène par l'inégalité $\mathbf{T}_2$, nous concluons immédiatement que μ vérifie $\mathbf{T}_2$. Dans la suite du chapitre, nous présentons les résultats obtenus dans la série d'articles [12, 13, 15]. L'objectif est de raffiner le théorème d'Otto-Villani, en montrant que l'inégalité $\mathbf{T}_2$ est en réalité équivalente à l'inégalité **LSI** restreinte à une bonne classe de fonctions (dans le cas euclidien, il s'agit des fonctions semi-convexes). Cette équivalence a d'abord été obtenue dans un cadre euclidien [12], puis a été généralisée aux espaces métriques quelconques dans [13, 15]. La représentation de l'inégalité de Talagrand par une inégalité de Sobolev logarithmique nous a permis de démontrer la stabilité de $\mathbf{T}_2$ par perturbation bornée du log de la densité de la mesure de référence. Cette propriété de stabilité est bien connue dans le cas des inégalités de type log-Sobolev (Lemme de Holley-Stroock [HS87]), mais constituait un problème ouvert pour les inégalités de transport.

Le chapitre 6 est consacré aux inégalités de Sobolev logarithmiques et de Poincaré sur un espace métrique mesuré dont la courbure est minorée et présente les résultats obtenus dans [11]. La notion de courbure en question a été introduite indépendamment par Lott-Villani [LV09] et Sturm [Stu06a, Stu06b]. Elle est définie par la semi-convexité de l'entropie relative le long des géodésiques de Wasserstein pour la distance W_2. Sans entrer dans le détail, on dit qu'un espace métrique (X, d) muni d'une probabilité μ vérifie la condition de courbure $\mathrm{CD}(K, \infty)$ si pour toutes probabilités ν_0, ν_1 sur $\mathcal{X}$, il existe une géodésique $(\nu_t)_{t\in[0,1]}$ joignant l'une à l'autre telle que

$$H(\nu_t|\mu) \leqslant (1-t)H(\nu_0|\mu) + tH(\nu_1|\mu) - K\frac{t(1-t)}{2}W_2^2(\nu_0, \nu_1), \qquad \forall t \in [0, 1].$$

Si l'espace $(\mathcal{X}, d)$ est une variété riemannienne munie de sa distance géodésique et $\mu(dx) = e^{-V(x)}\, dx$, cette condition est équivalente au célèbre critère de Bakry-Emery [BE85] :

$$\mathrm{Ric} + \mathrm{Hess}\, V \geqslant K.$$

Dans le cadre riemannien, des travaux de Wang [Wan97] ou plus récemment d'E. Milman [Mil10] montrent que, sous la condition de Bakry-Emery avec $K \leqslant 0$, une propriété de concentration gaussienne assez forte entraîne l'inégalité **LSI**. Nous généralisons ce résultat dans un cadre métrique.

Le chapitre 7 présente finalement des inégalités fonctionnelles pour des mesures de probabilité μ sur $\mathbb{R}^k$ n'ayant aucun moment exponentiel : $\int e^{\delta\|x\|_2}\, \mu(dx) = \infty$ pour tout $\delta > 0$. Dans l'article [7] (écrit avec P. Cattiaux, A. Guillin et C. Roberto), nous avons obtenu des informations sur le profil isopérimétrique de telles mesures.

Chapitre 1

Lois des grands nombres conditionnelles

Dans ce chapitre, je présente succinctement les résultats des articles [1] et [3] écrits durant ma thèse de doctorat (sous la direction de Patrick Cattiaux). Celle-ci portait sur l'étude des lois des grands nombres conditionnelles. Afin d'alléger l'exposition de ce chapitre, j'ai choisi de mettre en valeurs les méthodes et les idées plutôt que les résultats (dont l'énoncé est souvent assez technique).

1.1 Conditionnement et principes de grandes déviations

Dans ce chapitre, nous nous intéressons au comportement asymptotique d'objets de la forme suivante :

$$Q_n := \mathrm{Loi}\,(Z_n | Z_n \in C)\,,$$

où Z_n est une suite de variables aléatoires à valeurs dans un espace polonais (E, d) et $C \subset E$ est un ensemble donné. La théorie des grandes déviations permet dans certains cas de déterminer la limite ou les valeurs d'adhérence d'objets de ce type.

Rappelons qu'une suite $(Z_n)_{n\in\mathbb{N}}$ à valeurs dans un espace polonais (E, d) suit une *principe de grandes déviations* (PGD) de bonne fonction de taux $I : E \to [0, +\infty]$, si les ensembles de niveaux $\{I \leqslant k\}$ sont compacts et si pour tout fermé $F \subset E$, et tout ouvert $O \subset E$, les inégalités suivantes sont vérifiées :

$$\limsup_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(Z_n \in F) \leqslant -\inf_{x\in F} I(x),$$

et

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(Z_n \in O) \geqslant -\inf_{x\in O} I(x).$$

Le résultat suivant est démontré dans [SZ91] et découle très facilement des estimation asymptotiques données par un PGD.

Proposition 1.1.1. *Si la suite* Z_n *suit un PGD de bonne fonction de taux* I *et si* C *est un ensemble fermé tel que* I *atteigne son minimum en un unique point* $z^* \in C$ *tel que* $I(z^*) = \inf\{I(x); x \in C\} = \inf\{I(x); x \in \overset{\circ}{C}\} < +\infty$, *alors*

$$Q_n \to \delta_{z^*}, \qquad \textit{lorsque } n \to +\infty.$$

Ceci peut s'appliquer par exemple lorsque $Z_n = \frac{1}{n}\sum_{i=1}^n X_i$ ou $Z_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$, avec X_i une suite i.i.d. Dans ce cas, la convergence obtenue peut se comprendre comme une loi faible des grands nombres conditionnelle.

Si l'ensemble C est trop fin pour pouvoir définir Q_n ou pour appliquer les bornes de déviations asymptotiques données par le PGD, l'approche classique, développée par Stroock et Zeitouni dans [SZ91], consiste à grossir C en considérant une famille croissante C_ε d'ensembles plus épais contenant C, et à étudier le comportement asymptotique en double limite de

$$Q_{n,\varepsilon} := \mathrm{Loi}\,(Z_n | Z_n \in C_\varepsilon)\,.$$

Là encore sous de bonnes hypothèses sur C_ε et C, on a

$$\lim_{\varepsilon\to 0}\lim_{n\to+\infty} Q_{n,\varepsilon} = \delta_{z^*}.$$

Dans les articles [3] et [1] nous proposons une autre manière d'aborder les conditionnements fins. Dans notre approche, l'ensemble fin C est approché par une suite C_{ε_n} d'ensembles épais, et on cherche à démontrer un résultat du type

$$\lim_{n\to+\infty} Q_{n,\varepsilon_n} = \delta_{z^*}. \tag{1.1.2}$$

L'objectif est, partant d'une contrainte fine C, de construire une famille croissante de grossissements C_ε et de déterminer explicitement une vitesse de rétrécissement ε_n garantissant (1.1.2). Pour des familles C_ε raisonnables, cette formulation en limite simple est plus forte que celle en double limite ; elle s'avère également plus difficile techniquement. En effet, la difficulté que présente ce point de vue est que, l'ensemble par lequel on conditionne variant en fonction n, les bornes de déviations asymptotiques fournies par le PGD sont inutilisables, et doivent être remplacées par des bornes de déviations non-asymptotiques (telles que les inégalités de concentration de type Bernstein/Hoeffding).

Nous avons mené à bien cette étude pour deux familles de variables aléatoires Z_n : les mesures empiriques de suites i.i.d (dans [3]) et les mesures à poids aléatoires i.i.d. (dans [1]).

1.2 Principe conditionnel pour les mesures empiriques

Soit $(\mathcal{X}, d)$ un espace polonais (métrique séparable complet) ; on notera $\mathcal{P}(\mathcal{X})$ l'ensemble des mesures de probabilité borelienne sur $\mathcal{X}$. L'espace $\mathcal{P}(X)$ sera muni de la topologie de la convergence étroite, qui est la moins fine rendant continue les applications $\nu \mapsto \int f\,d\nu$, avec $f \in \mathcal{C}_b(\mathcal{X})$ (fonctions continues bornées sur $\mathcal{X}$). Cette topologie est métrisable par diverses distances (distance de Fortet-Mourier, distance de Levy-Prokhorov,...).

Si $(X_i)_{i\in\mathbb{N}}$ est une suite i.i.d à valeurs dans $\mathcal{X}$ et suivant une loi μ, alors les mesures empiriques $L_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ à valeurs dans $\mathcal{P}(\mathcal{X})$ suivent, d'après le théorème de Sanov [San61] un

PGD dont la fonction de taux est l'*entropie relative* par rapport à μ définie pour toute probabilité ν par

$$H(\nu|\mu) = \int \log \frac{d\nu}{d\mu}\, d\nu,$$

si ν est absolument continue par rapport à μ et $H(\nu|\mu) = +\infty$ sinon (voir [DZ98]). Le principe conditionnel associé est appelé *principe conditionnel de Gibbs*.

L'outil principal que nous avons utilisé pour établir une formulation en limite simple du principe de Gibbs est la belle inégalité démontrée par Csiszar dans [Csi84], valable pour tout ensemble convexe A fermé :

$$H(\mu^n_{k,A}|\mu^{*\otimes k}) \leqslant -k\left[\frac{1}{n}\log \mathbb{P}_{\mu^{\otimes n}}(L_n \in A) + H(\mu^*|\mu)\right], \tag{1.2.1}$$

où $\mu^n_{k,A} = \mathrm{Loi}_{\mu^{\otimes n}}(X_1, \ldots, X_k|L_n \in A)$, et μ^* est la *I-projection* de μ sur A, c'est à dire l'unique minimisant de $H(\,\cdot\,|\mu)$ sur A.

En appliquant l'inégalité précédente avec $A = C_{\varepsilon_n}$, nous montrons qu'une condition suffisante pour que $\mu^n_{k, C_{\varepsilon_n}}$ converge vers $\mu^{*\otimes k}$ est que

$$\liminf_{n\to+\infty} \frac{1}{n}\log \mathbb{P}(L_n \in C_{\varepsilon_n}) \geqslant -H(C|\mu),$$

inégalité qui ne relève pas du théorème de Sanov classique. Grâce à un argument de recentrage classique, nous montrons que cette dernière inégalité est satisfaite dès que :

$$\lim_{n\to+\infty} \mathbb{P}_{\mu^{*\otimes n}}(L_n \in C_{\varepsilon_n}) = 1. \tag{1.2.2}$$

Notons que dans (1.2.2), la probabilité de référence est μ^* qui, par définition, appartient à C et donc à tous les C_{ε_n}. La condition (1.2.2) est donc de type loi des grands nombres. Dans le reste de l'article, nous mettons en place des techniques permettant de calculer des vitesses de rétrécissement ε_n vérifiant (1.2.2).

1.2.1 Contraintes de type moment

Lorsque C est de la forme

$$C = \left\{\nu \in \mathcal{P}(\mathcal{X}) : \int F\, d\nu \in K\right\},$$

avec F une fonction à valeurs dans un espace vectoriel normé et K un ensemble convexe, nous définissons la famille de grossissement C_ε par $C_\varepsilon = \left\{\nu : \int F\, d\nu \in K_\varepsilon\right\}$, où $K_\varepsilon = \{x : d(x,k) \leqslant \varepsilon\}$. Si $x^* = \int F\, d\mu^*$, alors (1.2.2) est satisfaite dès lors que la suite ε_n est choisie de telle sorte que :

$$\mathbb{P}_{\mu^{*\otimes n}}\left(\left\|\frac{1}{n}\sum_{i=1}^{n} F(X_i) - x^*\right\| > \varepsilon_n\right) \to 0, \tag{1.2.3}$$

lorsque $n \to +\infty$. Sous de bonnes hypothèses d'intégrabilité exponentielle de F par rapport à μ, on peut utiliser des contrôles à la Bernstein / Yurinskii. Dans les deux cas, la probabilité de déviation apparaissant dans (1.2.3) est de l'ordre de $e^{-n\varepsilon_n^2}$, ce qui conduit à des vitesses de rétrécissement ε_n de l'ordre de $\frac{1}{\sqrt{n}}$.

De plus, en dimension finie, on peut obtenir, sous des hypothèses très raisonnables, une convergence en un sens fort de $\mu^n_{k,C_{\varepsilon_n}}$ vers $\mu^{*\otimes k}$ (convergence en entropie). La preuve utilise une généralisation de l'inégalité de Pinsker obtenue par Bolley et Villani dans [BV05].

1.2.2 Contraintes plus générales

Si C est un convexe fermé de $\mathcal{P}(\mathcal{X})$ au sens de la convergence étroite, nous définissons les grossissements C_ε par $C_\varepsilon = \{\nu : d_{LP}(\nu, C) \leqslant \varepsilon\}$, où d_{LP} est la distance de Levy-Prokhorov. Pour que (1.2.2) soit vérifiée dans ce cadre, il suffit que ε_n satisfasse :

$$\mathbb{P}_{\mu^{*\otimes n}}\left(d_{LP}(L_n, \mu^*) > \varepsilon_n\right) \to 0, \tag{1.2.4}$$

lorsque $n \to +\infty$. Si l'on suppose que l'espace $\mathcal{X}$ est compact, alors $(\mathcal{P}(X), d_{LP})$ est également compact, et en utilisant une technique de recouvrement due à Kulkarni et Zeitouni (voir [KZ95]), on peut majorer la probabilité apparaissant dans (1.2.4) de la manière suivante :

$$\mathbb{P}_{\mu^{*\otimes n}}\left(d_{LP}(L_n, \mu^*) > \varepsilon_n\right) \leqslant N_{\mathcal{P}(\mathcal{X})}(\varepsilon_n/4) \cdot e^{-n\varepsilon_n^2/2},$$

où $N_{\mathcal{P}(\mathcal{X})}(\varepsilon)$ est le *nombre de recouvrement* de l'espace métrique compact $\mathcal{P}(\mathcal{X})$ de niveau ε, c'est à dire le nombre minimal de boules ouvertes de rayon ε nécessaires pour recouvrir $\mathcal{P}(\mathcal{X})$. En utilisant une estimation de $N_{\mathcal{P}(\mathcal{X})}(\varepsilon)$ en fonction de $N_{\mathcal{X}}(\varepsilon)$ (nombre de recouvrement sur $\mathcal{X}$), on démontre que si ε_n vérifie la condition

$$\frac{n\varepsilon_n^2}{2} + \log(\varepsilon_n) N_X(\varepsilon_n/8) \to +\infty, \tag{1.2.5}$$

lorsque $n \to +\infty$, alors la condition (1.2.2) est satisfaite. On montre par ailleurs, qu'il existe toujours au moins une suite ε_n vérifiant (1.2.5) ; notre méthode est donc toujours applicable. Sous des hypothèses assez fortes, il est possible d'étendre cette méthode au cas où l'espace $\mathcal{X}$ n'est pas compact. En un mot, l'idée consiste à approcher la probabilité μ^* par une suite de probabilités portées par des compacts de plus en plus gros. On observe une compétition entre la forte croissance des nombres de recouvrement et le terme de grandes déviations en $e^{-n\varepsilon_n^2/2}$. Dans certains cas, il est possible de construire une suite ε_n telle que le terme de grandes déviations l'emporte sur le terme de recouvrement.

1.2.3 Applications

Ces méthodes nous permettent de donner une interprétation statistique des ponts de Schrödinger et des processus de Nelson. On s'intéresse aux comportements inhabituels de grands nuages de particules browniennes. Si $X_1, \ldots, X_n$ sont n particules browniennes indépendantes, le problème est de déterminer l'évolution la plus probable du nuage sachant que celui-ci a été trouvé avec une distribution approximativement égale à ν_t aux instants $t \in I$ (I étant un sous ensemble de $[0,1]$). Posant

$$C = \left\{\mathbb{V} \in \mathcal{P}\left(C([0,1], \mathbb{R}^q)\right) : \forall t \in I, \mathbb{V}_t = \nu_t\right\},$$

il s'agit d'estimer (le nombre de particules étant supposé très grand)

$$\lim_{n\to+\infty} \mathrm{Loi}(L_n | L_n \in C).$$

Ceci reste bien sûr formel, puisque la contrainte C est une contrainte (convexe) fine. Pour de bons flots de marginales $(\nu_t)_{t\in I}$, le problème de l'existence de la I-projection $\mathbb{W}^*$ de $\mathbb{W}$ (mesure de Wiener sur $C([0,1],\mathbb{R}^q)$) sur C a été étudié par différents auteurs [CL94, CL95]. Dans le cas où $I = \{0,1\}$, on parle de ponts de Schrödinger et pour $I = [0,1]$, de processus de Nelson. Dans les deux cas, nous montrons comment construire des suites ε_n explicites telles que

$$\lim_{n\to+\infty} \mathrm{Loi}\,(L_n\,| L_n \in C_{\varepsilon_n}) = \delta_{\mathbb{W}^*}.$$

Cette approche des conditionnements fins nous a également permis de donner une justification théorique à la méthode de calibration d'Avellaneda (voir [AFHS97]) utilisée en mathématiques financières.

1.3 Mesures à poids aléatoires

Dans l'article [1] nous démontrons des principes conditionnels pour des mesures à poids aléatoires de la forme suivante :

$$N_n = \frac{1}{n}\sum_{i=1}^{n} Z_i \delta_{x_i^n},$$

où Z_i est une suite de variables aléatoires réelles i.i.d dont la loi sera notée μ, et x_i^n une famille de points d'un espace métrique $\mathcal{X}$ telle que la mesure $\frac{1}{n}\sum_{i=1}^{n} \delta_{x_i^n}$ converge étroitement vers une probabilité $R \in \mathcal{P}(\mathcal{X})$ donnée.

Les grandes déviations de ces mesures aléatoires ont été étudiées par différents auteurs (voir par exemple [Naj02]). Sous de bonnes hypothèses, elles satisfont un PGD de bonne fonction de taux notée $I_\mu(P|R)$ et définie par

$$I_\mu(P|R) = \int \Lambda_\mu^*\left(\frac{dP}{dR}\right)\, dR,$$

si $P \ll R$, et $I_\mu(P|R) = +\infty$ sinon, où Λ_μ^* est la transformée de Cramer de μ.

Dans [1], nous cherchons des suites ε_n telles que

$$R_{n,\varepsilon_n} := \mathbb{E}\,[N_n | N_n \in C_{\varepsilon_n}] \to R^*, \tag{1.3.1}$$

lorsque $n \to +\infty$, où pour tout $\varepsilon > 0$, $C_\varepsilon = \left\{Q \in \mathcal{M}(\mathcal{X}), \int_{\mathcal{X}} F\, dQ \in K^\varepsilon\right\}$, F est une application de $\mathcal{X}$ dans $\mathbb{R}^q$, K est un convexe fermé de $\mathbb{R}^q$ et R^* est l'unique minimisant de $I_\mu(\,\cdot\,|R)$ sur C.

Sous de bonnes hypothèses, nous montrons que (1.3.1) a lieu pour toute suite $\varepsilon_n = 1/n^a$, avec $a < 1/2$. La raison pour laquelle nous ne considérons que des conditionnements par des contraintes de type moment est qu'ici, contrairement au principe conditionnel de Gibbs, la forme algébrique particulière de R^* est utilisée dans la preuve et cette forme n'est connue que dans ce cas précis.

L'étude de $R_{n,\,\varepsilon_n}$ suit dans ses grandes lignes celle du principe conditionnel de Gibbs dans le cas d'un conditionnement par une contrainte de type moment. Notre preuve repose sur une généralisation de l'inégalité de Csiszar (1.2.1) qui est la suivante :

$$\alpha\left(\left\|R_{n,\,\varepsilon}-R^*_{n,\,\varepsilon}\right\|_{VT}\right)\leqslant -\left[\frac{1}{n}\log\mathbb{P}\left(L_n\in C_\varepsilon\right)+I_\mu(R^*_{n,\,\varepsilon}|\overline{R}_n)\right], \tag{1.3.2}$$

où

- α est une fonction convexe strictement croissante et nulle en 0,
- $\overline{R}_n=\frac{1}{n}\sum_{i=1}^n\delta_{x_i^n}$,
- pour tout $\varepsilon>0$, $R^*_{n,\varepsilon}$ est l'unique minimisant sur l'ensemble C_ε de la fonctionnelle $I_\mu(\,\cdot\,|\overline{R}_n)$ définie par
$$I_\mu(P|\overline{R}_n):=\int\Lambda^*_\mu\left(\frac{dP}{d\overline{R}_n}\right)d\overline{R}_n.$$

Les mesures $R^*_{n,\,\varepsilon}$ sont appelées *minimisants de l'entropie en moyenne*. Ces mesures ont été introduites et étudiées par Gamboa et Gassiat dans [GG97].

Nous montrons en utilisant des techniques de [GG97] que la suite de fonctions $I_\mu(\,\cdot\,|\overline{R}_n)$ converge en un sens suffisamment fort vers $I_\mu(\,\cdot\,|R)$ pour que pour toute suite ε_n de limite nulle $R^*_{n,\,\varepsilon_n}$ converge vers R^*. Ainsi, pour montrer (1.3.1), il suffit de contrôler le membre de droite de (1.3.2). Cette dernière étape fait intervenir des outils déjà utilisés pour le principe de Gibbs : recentrage et bornes à la Bernstein.

La preuve de l'inégalité (1.3.2) est assez proche de celle de l'inégalité de Csiszar (1.2.1). La principale nouveauté est le recours à une inégalité de type transport, inspirée des travaux de Bobkov et Götze [BG99] sur l'inégalité de transport $\mathbf{T}_1$ (voir chapitre 2). Nous démontrons, en effet que pour toute mesure de probabilité μ sur $\mathbb{R}$, on peut construire une fonction α convexe strictement croissante nulle en 0 telle que

$$\alpha\left(\left|\int x\,d\nu-\int x\,d\mu\right|\right)\leqslant H(\nu|\mu),\qquad\forall\nu\in\mathcal{P}(\mathbb{R}).$$

Ce dernier résultat est ce qui a orienté mes recherches vers une étude des inégalités de transport et de leurs liens avec les grandes déviations.

Chapitre 2

Inégalités de transport - une brève introduction

Dans ce chapitre, nous présentons brièvement quelques propriétés de base des inégalités de transport. Les principaux points abordés sont d'une part la forme duale et la tensorisation des inégalités de transport et d'autre part leur lien avec le phénomène de concentration de la mesure. On pourra consulter [9] ou [Vil09] pour de plus amples informations sur le sujet.

2.1 Définitions

Dans tout ce qui suit, $(\mathcal{X}, d)$ sera un espace polonais (métrique séparable et complet) et $\mathcal{P}(\mathcal{X})$ désignera l'ensemble des mesures de probabilités boreliennes sur $\mathcal{X}$.

Les inégalités de transport sont des inégalités comparant un coût de transport optimal au sens de Kantorovich à la fonctionnelle d'entropie relative déjà rencontrée au chapitre 1. Commençons par rappeler la définition des coûts de transport optimaux (voir [Vil03, Vil09] ou [RR98] pour une étude détaillée).

Définition 2.1.1. *Soit* $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$ *une fonction de coût mesurable. Pour tout couple* $(\nu_1, \nu_2) \in \mathcal{P}(\mathcal{X})^2$*, le coût de transport optimal* $\mathcal{T}_c(\nu_1, \nu_2)$ *de* ν_1 *sur* ν_2 *est défini par*

$$\mathcal{T}_c(\nu_1, \nu_2) = \inf \iint c(x, y)\, \pi(dxdy),$$

où l'infimum porte sur les mesures de probabilités π *sur l'espace produit* $\mathcal{X} \times \mathcal{X}$ *ayant pour première marginale* ν_1 *et pour seconde* ν_2.

Pour tout $p \geqslant 1$, nous utiliserons la notation $\mathcal{T}_p$ pour désigner le coût de transport $\mathcal{T}_{d^p}$ associé à la fonction de coût d^p. Par ailleurs, rappelons qu'en posant

$$W_p(\nu_1, \nu_2) = \mathcal{T}_p(\nu_1, \nu_2)^{1/p},$$

on obtient une distance appelée distance de Wasserstein d'ordre p (voir [Vil09]).

Définition 2.1.2. *Considérons une fonction* $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ *; nous dirons que* $\mu \in \mathcal{P}(\mathcal{X})$ *vérifie l'inégalité de transport* $\alpha(\mathcal{T}_c) \leqslant H$ *si pour toute mesure de probabilité* $\nu \in \mathcal{P}(\mathcal{X})$*, on a*

$$\alpha\left(\mathcal{T}_c(\mu,\nu)\right) \leqslant H(\nu|\mu). \tag{2.1.3}$$

Il est intéressant de noter que la célèbre inégalité de Csiszar-Kullback-Pinsker [Csi67, Kul67, Pin64],

$$\|\nu-\mu\|_{VT} \leqslant \sqrt{2H(\nu|\mu)}, \qquad \forall \nu,\mu \in \mathcal{P}(\mathcal{X}),$$

où $\|\nu-\mu\|_{VT} = \sup_{f\in[-1,1]} \int f\,d\nu - \int f d\mu$ désigne la norme en variation totale, rentre dans le cadre de cette définition, puisqu'il est facile de montrer que $\|\nu-\mu\|_{VT} = \mathcal{T}_c(\nu,\mu)$, avec $c(x,y) = \mathbf{1}_{x\neq y}$.

Signalons deux classes particulières d'inégalités de transport apparaissant fréquemment dans la littérature.

Définition 2.1.4 (Inégalités $\mathbf{T}_1$ et $\mathbf{T}_2$)**.** *On dit que* μ *vérifie l'inégalité* $\mathbf{T}_1(a)$ *si elle vérifie l'inégalité*

$$W_1(\mu,\nu) \leqslant \sqrt{aH(\nu|\mu)}, \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

On dit que μ *vérifie l'inégalité* $\mathbf{T}_2(a)$ *si elle vérifie l'inégalité*

$$W_2(\mu,\nu) \leqslant \sqrt{aH(\nu|\mu)}, \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

L'inégalité de Jensen montre immédiatement que $\mathbf{T}_2(a)$ entraîne $\mathbf{T}_1(a)$.

L'étude des inégalités de transport a réellement été lancée par les travaux de Marton [Mar86, Mar96a, Mar96b, Mar98, Mar04] et Talagrand [Tal96] dans les années 90. La motivation principale de ces travaux et de ceux qui ont suivi était les diverses inégalités de concentration de la mesure qui, comme nous allons le voir, peuvent être tirées d'inégalités du type (2.1.3). A ce titre, les premiers travaux de Marton [Mar86, Mar96a, Mar96b] portent sur des raffinements multidimensionnels de l'inégalité de Csiszar-Kullback-Pinsker qui permettent de retrouver certaines inégalités de concentration universelles pour les mesures produits obtenues par Talagrand [Tal95]. L'article de Talagrand [Tal96] introduit quant à lui l'inégalité de transport $\mathbf{T}_2$ et prouve que la mesure gaussienne standard $\gamma(dx) = e^{-x^2/2}\,dx/\sqrt{2\pi}$ sur $\mathbb{R}$ la vérifie avec la constante optimale $a = 2$. Une autre inégalité de transport adaptée cette fois à la mesure exponentielle symétrique $\mu_1(dx) = e^{-|x|}\,dx/2$ est également démontrée dans ce papier. Nous aurons l'occasion de rappeler ce résultat dans le chapitre suivant.

2.2 Forme duale et tensorisation

Le premier résultat que nous allons rappeler est dû à Bobkov et Götze [BG99]. Il fournit une formulation duale des inégalités de transport. En fait le résultat de [BG99] concernait les inégalités $\mathbf{T}_1$ et $\mathbf{T}_2$, mais il a par la suite été étendu au cas général, par exemple dans [5].

Théorème 2.2.1. *Soit* $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ *une fonction convexe telle que* $\alpha(0) = 0$*. Une probabilité* $\mu \in \mathcal{P}(\mathcal{X})$ *vérifie l'inégalité de transport* $\alpha(\mathcal{T}_c) \leqslant H$ *si et seulement si elle vérifie, pour toute fonction* $\varphi \in \mathcal{C}_b(\mathcal{X})$ *(ensemble des fonctions continues bornées sur* $\mathcal{X}$*),*

$$\int e^{sQ_c\varphi(x)}\,\mu(dx) \leqslant e^{s\int \varphi(x)\,\mu(dx)+\alpha^*(s)}, \qquad \forall s \geqslant 0,$$

où l'opérateur d'inf-convolution Q_c *est défini comme suit*

$$Q_c\varphi(x) = \inf_{y\in\mathcal{X}}\{\varphi(y) + c(x,y)\}, \qquad \forall x \in \mathcal{X}.$$

La preuve (voir [9]) se fait par la combinaison des résultats de dualité classiques suivants :
- Dualité de Kantorovich (voir par exemple [Vil09]) :

$$\mathcal{T}_c(\mu,\nu) = \sup_{\varphi\in\mathcal{C}_b(\mathcal{X})} \left\{ \int Q_c\varphi(x)\,\mu(dx) - \int \varphi(y)\,\nu(dy) \right\}.$$

- Dualité pour l'entropie relative (voir, par exemple, [9] pour une preuve simple) :

$$\log\left(\int e^h\,d\mu\right) = \sup_{\nu\in\mathcal{P}(\mathcal{X})} \left\{ \int h\,d\nu - H(\nu|\mu) \right\}, \qquad \forall h \in \mathcal{C}_b(\mathcal{X}).$$

- Dualité de Fenchel pour la fonction α : $\alpha(s) = \sup_{t\in\mathbb{R}^+}\{st - \alpha^*(t)\}$, $s \geqslant 0$.

Remarque 2.2.2. *Les inégalités de transport sont fortement apparentées à la propriété* (τ) *introduite par Maurey dans [Mau91]. Si* $c : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ *est une fonction de coût sur un espace* $\mathcal{X}$ *et* μ *une probabilité sur* $\mathcal{X}$*, on dit que le couple* (c,μ) *vérifie la propriété* τ *si pour toute fonction* f *bornée sur* $\mathcal{X}$*,*

$$\int e^{Q_c f}\,d\mu \int e^{-f}\,d\mu \leqslant 1,$$

avec $Qf(x) = \inf_{y\in\mathcal{X}}\{f(y) + c(x,y)\}$*. D'après la forme duale de Bobkov-Götze, il est clair que cette propriété entraîne que* μ *vérifie l'inégalité* $\mathcal{T}_c \leqslant H$*. La réciproque est vraie si* $c(x,y) = \theta\left(d(x,y)\right)$*, avec* $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ *une fonction convexe nulle en* 0*, et* d *une distance sur* $\mathcal{X}$*. On montre dans [9] que si* μ *vérifie l'inégalité* $\mathcal{T}_c \leqslant H$*, alors le couple* $(\tilde{c},\mu)$ *vérifie la propriété* (τ)*, avec* $\tilde{c}(x,y) = 2\theta\left(d(x,y)/2\right)$*.*

Le second résultat concerne le passage au produit tensoriel dans les inégalités de transport. Le résultat suivant, issu de [5], généralise (en utilisant les mêmes ingrédients) des cas particuliers précédemment établis par Marton [Mar86, Mar96a, Mar96b] et Talagrand [Tal96].

Théorème 2.2.3. *Soit* $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ *une fonction convexe telle que* $\alpha(0) = 0$*. Si* $\mu \in \mathcal{P}(\mathcal{X})$ *vérifie l'inégalité* $\alpha(\mathcal{T}_c) \leqslant H$*, alors pour tout* $n \geqslant 1$*, la mesure produit* $\mu^n \in \mathcal{P}(\mathcal{X}^n)$ *vérifie l'inégalité*

$$n\alpha\left(\frac{\mathcal{T}_{c^{\oplus n}}(\mu^n,\nu)}{n}\right) \leqslant H(\nu|\mu^n), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}^n),$$

où la fonction de coût produit $c^{\oplus n}$ *est définie par*

$$c^{\oplus n}(x,y) = \sum_{i=1}^{n} c(x_i,y_i), \qquad \forall x = (x_1,\ldots,x_n) \in \mathcal{X}^n,\ y = (y_1,\ldots,y_n) \in \mathcal{X}^n.$$

Remarque 2.2.4. *Dans le cas où la fonction* α *est linéaire, on observe que la tensorisation se fait de manière indépendante de la dimension.*

La preuve du théorème 2.2.3 repose sur les formules de sous-additivité suivantes pour le coût de transport et pour l'entropie relative (écrites pour $n = 2$ pour simplifier) : si $\nu \in \mathcal{P}(\mathcal{X}^2)$ admet la désintégration suivante

$$\nu(dy_1\,dy_2) = \nu_1(dy_1)\nu_2(dy_2|y_1),$$

où ν_1 désigne sa première marginale, alors

$$H(\nu|\mu^2) = H(\nu_1|\mu) + \int H(\nu_2(\,\cdot\,|y_1)|\mu)\,\nu_1(dy_1)$$
$$\mathcal{T}_{c^{\oplus 2}}(\mu^2,\nu) \leqslant \mathcal{T}_c(\mu,\nu_1) + \int \mathcal{T}_c(\mu,\nu_2(\,\cdot\,|y_1))\,\nu_1(dy_1).$$

La formule concernant l'entropie est bien connue. Celle portant sur le coût de transport se démontre en exhibant un couplage π^* entre μ^2 et ν tel que

$$\iint c^{\oplus 2}(x,y)\,\pi^*(dxdy) = \mathcal{T}_c(\mu,\nu_1) + \int \mathcal{T}_c(\mu,\nu_2(\,\cdot\,|y_1))\,\nu_1(dy_1).$$

On obtient π^* en faisant appel à une construction classique due à Knott et Rosenblatt :

$$\pi^*(dxdy) = \pi_1(dx_1dy_1)\pi_2(dx_2dy_2|y_1),$$

où π_1 est un couplage optimal entre μ et ν_1 et, pour tout y_1, $\pi_2(\,\cdot\,|y_1)$ un couplage optimal entre μ et $\nu_2(\,\cdot\,|y_1)$.

Notons qu'il est également possible de démontrer le théorème 2.2.3 en utilisant la forme duale donnée par le théorème 2.2.1 (voir par exemple [9]).

2.3 Lien avec la concentration

Nous allons maintenant aborder le point le plus important de ce chapitre qui est qu'une inégalité de transport contient une information sur le profil de concentration de la mesure de référence.

Commençons par rappeler quelques définitions.
Si $A \subset \mathcal{X}$, l'épaissi A_r de A, $r \geqslant 0$ est défini de la manière suivante :

$$A_r = \{x \in \mathcal{X};\, d(x,A) \leqslant r\}.$$

Définition 2.3.1. *On dit qu'une mesure de probabilité μ sur $(\mathcal{X}, d)$ vérifie une inégalité de concentration s'il existe une fonction croissante $\beta : \mathbb{R}^+ \to \mathbb{R}^+$ telle que pour tout $A \subset \mathcal{X}$ vérifiant $\mu(A) \geqslant 1/2$,*

$$\mu(A_r) \geqslant 1 - \exp(-\beta(r)), \qquad \forall r \geqslant 0.$$

De manière équivalente, pour toute fonction 1-Lipschitz $f : \mathcal{X} \to \mathbb{R}$,

$$\mu(f > m(f) + r) \leqslant \exp(-\beta(r)), \qquad \forall r \geqslant 0,$$

où $m(f)$ désigne une valeur médiane de f pour la probabilité μ.

L'équivalence énoncée dans la définition précédente est très classique, on en trouvera la preuve dans [Led01] par exemple.

Tout comme les inégalités de Sobolev logarithmiques et de Poincaré (et leurs différentes variantes), les inégalités de transport permettent de démontrer des estimées de concentration. Cette propriété, énoncée dans le théorème ci-dessous, a été mise au jour pour la première fois par Marton dans [Mar86]. La preuve étant très courte, nous la rappelons à la suite.

Théorème 2.3.2. *Soient* $p \geqslant 1$ *et* μ *une mesure de probabilité sur* $\mathcal{X}$ *vérifiant l'inégalité de transport*

$$\alpha\left(\mathcal{T}_p(\nu, \mu)\right) \leqslant H(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}),$$

où $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ *est une bijection continue et strictement croissante. Sous ces hypothèses, la probabilité* μ *vérifie l'inégalité de concentration suivante :*

$$\mu(A_r) \geqslant 1 - \exp(-\alpha\left((r - r_o)^p\right)), \qquad \forall r \geqslant r_o = \alpha^{-1}(\log(2))^{1/p},$$

pour tout $A \subset \mathcal{X}$ *tel que* $\mu(A) \geqslant 1/2$.

Démonstration. Prenons $A \subset \mathcal{X}$ vérifiant $\mu(A) \geqslant 1/2$, posons $B = \mathcal{X} \setminus A_r$, $r \geqslant 0$ et définissons les mesures de probabilité $\mu_A(dx) = \frac{1}{\mu(A)}\mathbf{1}_A(x)\,dx$ et $\mu_B(dx) = \frac{1}{\mu(B)}\mathbf{1}_B(x)\,dx$. En appliquant l'inégalité triangulaire pour la distance W_p à la première ligne, puis l'inégalité de transport à la deuxième, on voit que

$$\begin{aligned}
W_p(\mu_A, \mu_B) &\leqslant W_p(\mu_A, \mu) + W_p(\mu_B, \mu) \\
&\leqslant \left(\alpha^{-1}(H(\mu_A|\mu))\right)^{1/p} + \left(\alpha^{-1}(H(\mu_B|\mu))\right)^{1/p} \\
&= \left(\alpha^{-1}(-\log(\mu(A)))\right)^{1/p} + \left(\alpha^{-1}(-\log(\mu(B)))\right)^{1/p} \\
&\leqslant r_o + \left(\alpha^{-1}(-\log(\mu(B)))\right)^{1/p}.
\end{aligned}$$

Pour tout couplage π entre $\mu(A)$ et $\mu(B)$, on a $\pi(A \times B) = 1$. Comme $d(x, y) \geqslant r$ pour tout $(x, y) \in A \times B$, il s'ensuit que

$$\iint d^p(x, y)\, \pi(dxdy) \geqslant r^p.$$

Par conséquent, $W_p(\mu_A, \mu_B) \geqslant r$ et l'inégalité de concentration en découle facilement. □

La propriété de tensorisation associée au théorème 2.3.2 permet d'obtenir immédiatement des bornes de concentration pour les mesures produits. Avant d'énoncer ce résultat, nous devons spécifier de quelle distance sera muni l'espace produit $\mathcal{X}^n$. Pour tout $p \geqslant 1$, nous noterons d_p la distance sur $\mathcal{X}^n$ définie par

$$d_p(x, y) = \left[\sum_{i=1}^{n} d^p(x_i, y_i)\right]^{1/p}, \qquad \forall x = (x_1, x_2, \dots, x_n),\ y = (y_1, y_2, \dots, y_n).$$

La dimension n est omise dans la notation de d_p pour alléger les écritures.

Corollaire 2.3.3. *Soient* $p \geqslant 1$ *et* μ *une mesure de probabilité sur* $\mathcal{X}$ *vérifiant l'inégalité de transport*

$$\alpha\left(\mathcal{T}_p(\nu, \mu)\right) \leqslant H(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}),$$

où $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ *est une fonction convexe telle que* $\alpha(0) = 0$. *Sous ces hypothèses, pour tout* $n \geqslant 1$ *et toute fonction* $f : \mathcal{X}^n \to \mathbb{R}$ 1*-Lipschitz pour la distance* d_p, *on a*

$$\mu^n(f > m(f) + r) \leqslant \exp\left(-n\alpha\left(\frac{(r - r_o)^p}{n}\right)\right), \qquad \forall r \geqslant r_o = \left(n\alpha^{-1}(\log(2)/n)\right)^{1/p}.$$

Remarque 2.3.4. *Là encore, la situation où α est linéaire est très particulière, puisque dans ce cas la dimension n n'apparaît plus dans l'inégalité de concentration pour la mesure produit μ^n. On parle de* concentration adimensionnelle. *Par exemple, si μ vérifie l'inégalité $\mathbf{T}_2(a)$, on obtient que*

$$\mu^n(f > m(f) + r) \leqslant e^{-\frac{1}{a}(r-r_o)^2}, \qquad \forall r \geqslant r_o = \sqrt{a \log(2)},$$

pour tout n et pour toute fonction $f : \mathcal{X}^n \to \mathbb{R}$ qui est 1-Lipschitz pour la distance d_2. Nous reviendrons plus longuement sur ce phénomène dans le chapitre 3.
Si μ vérifie seulement $\mathbf{T}_1(a)$, alors on a le résultat plus faible suivant : pour tout $n \geqslant 1$,

$$\mu^n(f \geqslant m(f) + r) \leqslant e^{-\frac{1}{na}(r-r_o)^2}, \qquad \forall r \geqslant r_o = \sqrt{na \log(2)},$$

pour toute $f : \mathcal{X}^n \to \mathbb{R}$ 1-Lipschitz par rapport à la distance d_1.

Grâce à une technique de changement de métrique introduite dans [13], il est possible d'étendre facilement le résultat précédent à des inégalités de transport de la forme $\alpha(\mathcal{T}_c) \leqslant H$ où $c(x,y) = \theta(d(x,y))$, à condition que θ vérifie les conditions suivantes.

Hypothèse 2.3.5. *La fonction $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ est une fonction strictement croissante, convexe, telle que $\theta(0) = 0$ et vérifiant la condition Δ_2,* i.e.

$$\sup_{x>0} \frac{\theta(2x)}{\theta(x)} < +\infty.$$

Le lemme suivant est démontré dans [13].

Lemme 2.3.6. *Si $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ est une fonction vérifiant l'hypothèse 2.3.5, alors*

$$r_\theta := \inf_{x>0} \frac{x\theta'_-(x)}{\theta(x)} \geqslant 1 \qquad \textit{et} \qquad 1 \leqslant p_\theta := \sup_{x>0} \frac{x\theta'_-(x)}{\theta(x)} < +\infty,$$

où θ'_- et θ'_+ sont les dérivées de θ à gauche et à droite. Par ailleurs, $p_\theta = 1$ si et seulement si θ est linéaire. Enfin, la fonction, θ^{1/p_θ} est sous-additive :

$$\theta^{1/p_\theta}(x+y) \leqslant \theta^{1/p_\theta}(x) + \theta^{1/p_\theta}(y), \qquad \forall x, y \geqslant 0.$$

En vertu de ce lemme, si d est une distance sur $\mathcal{X}$, alors la fonction d_θ définie par

$$d_\theta(x,y) = \theta^{1/p_\theta}(d(x,y)), \qquad \forall x, y \in \mathcal{X},$$

est encore une distance sur $\mathcal{X}$.

Grâce à ce changement de métrique, l'inégalité $\alpha(\mathcal{T}_c) \leqslant H$, avec $c(x,y) = \theta(d(x,y))$ est équivalente à l'inégalité $\alpha\left(\mathcal{T}_{p_\theta}\right) \leqslant H$, étant entendu que, dans la deuxième inégalité, la distance équipant $\mathcal{X}$ est cette fois d_θ. On en déduit immédiatement le résultat suivant (qui synthétise la proposition 3.3 de [6] et la proposition 4.2 de [9]).

Théorème 2.3.7. *Soit μ une mesure de probabilité sur $\mathcal{X}$ vérifiant l'inégalité de transport*

$$\alpha\left(\mathcal{T}_c(\nu,\mu)\right) \leqslant H(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}),$$

où $c(x,y) = \theta(d(x,y))$ *avec* θ *une fonction vérifiant l'hypothèse 2.3.5 et* $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ *une fonction convexe telle que* $\alpha(0) = 0$. *Alors pour tout* $n \geqslant 1$, *on a*

$$\mu^n(f > m(f) + r) \leqslant \exp\left(-n\alpha\left(\frac{(r-r_o)^{p_\theta}}{n}\right)\right), \qquad \forall r \geqslant r_o = \left(n\alpha^{-1}(\log(2)/n)\right)^{1/p_\theta},$$

pour toute fonction $f : \mathcal{X}^n \to \mathbb{R}$ 1*-Lipschitz pour la distance* d_{θ,p_θ} *sur* $\mathcal{X}^n$ *définie par*

$$d_{\theta,p_\theta}(x,y) = \left[\sum_{i=1}^n d_\theta^{p_\theta}(x_i,y_i)\right]^{1/p_\theta} = \left[\sum_{i=1}^n \theta(d(x_i,y_i))\right]^{1/p_\theta}.$$

2.4 Quelques résultats sur les inégalités T_1 et T_2

Pour finir cette introduction, nous allons revenir sur les inégalités $\mathbf{T}_1$ et $\mathbf{T}_2$ et rappeler quelques résultats dûs à Djellout, Guillin et Wu [DGW04] et Otto et Villani [OV00]. Comme nous allons le voir, ces inégalités sont de natures très différentes, bien que toutes deux associées à un phénomène de concentration gaussien.

2.4.1 T_1 et l'intégrabilité

Djellout, Guillin et Wu [DGW04] ont caractérisé l'inégalité $\mathbf{T}_1$ par une condition d'intégrabilité (sur laquelle nous reviendrons plus longuement dans le chapitre 3) :

Théorème 2.4.1. *Soit* $\mu \in \mathcal{P}(\mathcal{X})$ *; les propositions suivantes sont équivalentes :*

1. *La probabilité* μ *vérifie* $\mathbf{T}_1(a)$, *pour un certain* $a > 0$.
2. *La probabilité* μ *vérifie* $\iint e^{\varepsilon d^2(x,y)}\,\mu(dx)\mu(dy) < +\infty$.

De plus, 2. entraîne 1. avec $a = \frac{4}{\varepsilon^2} \sup_{k\geqslant 1} \left(\frac{(k!)^2}{(2k)!}\right)^{1/k} \left[\iint e^{\varepsilon d^2(x,y)}\,\mu(dx)\mu(dy)\right]^{1/k} < +\infty.$

La classe des inégalités vérifiant $\mathbf{T}_1$ est donc identique à celle des probabilités ayant un moment exponentiel d'ordre 2 fini.

2.4.2 T_2 entre Poincaré et log-Sobolev

L'inégalité $\mathbf{T}_2$ est nettement plus structurée que l'inégalité $\mathbf{T}_1$. Tout d'abord, comme l'ont montré Otto et Villani dans [OV00], l'inégalité $\mathbf{T}_2$ entraîne une inégalité de trou spectral, également appelée inégalité de Poincaré :

Théorème 2.4.2. *Soient* $(\mathcal{X}, g)$ *une variété riemannienne complète et connexe équipée de sa distance géodésique* d *et* μ *une mesure de probabilité absolument continue par rapport à la mesure de Lebesgue sur* $\mathcal{X}$. *Si* μ *vérifie l'inégalité* $\mathbf{T}_2(a)$ *pour un certain* $a > 0$, *alors* μ *vérifie l'inégalité de Poincaré suivante :*

$$\frac{2}{a}\mathrm{Var}(f) \leqslant \int \|\nabla f\|_2^2\, d\mu,$$

pour toute fonction $f : \mathcal{X} \to \mathbb{R}$ *Lipschitz.*

Une mesure de probabilité vérifiant l'inégalité de Poincaré a en particulier un support connexe, ce qui impose d'emblée une restriction géométrique à la classe des mesures de probabilités vérifiant $\mathbf{T}_2$.

L'inégalité $\mathbf{T}_2$ est aussi apparentée à l'inégalité de Sobolev logarithmique introduite par Gross [Gro75] (voir [ABC+00] pour un panorama d'applications : hypercontractivité, convergence à l'équilibre de semi-groupes,...).

Rappelons que la fonctionnelle d'entropie est définie de la manière suivante :

$$\mathrm{Ent}_\mu(h) := \int h \log\left(\frac{h}{\int h\,d\mu}\right) d\mu, \qquad \forall h > 0.$$

Définition 2.4.3. *On dit qu'une mesure de probabilité μ absolument continue vérifie l'inégalité de Sobolev logarithmique* $\mathbf{LSI}(C)$, $C > 0$, *si*

$$\mathrm{Ent}_\mu(f^2) \leqslant C \int \|\nabla f\|_2^2 \, d\mu,$$

pour toute fonction $f : \mathcal{X} \to \mathbb{R}$ Lipschitz.

Le théorème suivant est également dû à Otto et Villani [OV00].

Théorème 2.4.4. *Soient $(\mathcal{X}, g)$ une variété riemannienne complète et connexe équipée de sa distance géodésique d et μ une mesure de probabilité absolument continue par rapport à la mesure de Lebesgue sur $\mathcal{X}$. Si μ vérifie l'inégalité de Sobolev logarithmique* $\mathbf{LSI}(C)$ *alors elle vérifie aussi* $\mathbf{T}_2(C)$.

On peut grâce à ce théorème dégager une première classe d'exemples de mesures de probabilité vérifiant l'inégalité $\mathbf{T}_2$.

Rappelons qu'une mesure de probabilité $\mu(dx) = e^{-V(x)}\,dx$ sur $\mathcal{X}$ associée à une fonction $V : \mathcal{X} \to \mathbb{R}$ de classe $\mathcal{C}^2$ vérifie le critère de Bakry-Emery [BE85] $\mathrm{CD}(K, \infty)$, $K \in \mathbb{R}$, si

$$\mathrm{Ric}_x + \mathrm{Hess}_x V \geqslant K g_x, \qquad \forall x \in \mathcal{X}.$$

Il est bien connu que pour $K > 0$, la condition $\mathrm{CD}(K, \infty)$ entraîne $\mathbf{LSI}(2/K)$, et donc aussi $\mathbf{T}_2(2/K)$ en vertu du théorème précédent. On retrouve entre autre le fait que la mesure gaussienne standard sur $\mathbb{R}^k$ vérifie $\mathbf{T}_2(2)$.

Chapitre 3

Inégalités de transport - conditions nécessaires et/ou suffisantes

Ce chapitre présente les résultats des articles [2], [4], [8] et [14]. Le point commun de ces travaux est qu'ils cherchent à mettre en évidence des conditions simples permettant de déterminer si une mesure de probabilité vérifie telle ou telle inégalité de transport. L'article [2] a pour objet l'étude des inégalités de transport de la forme $\alpha(\mathcal{T}_c) \leqslant H$, pour une fonction α *non-linéaire*. A cette classe appartient notamment l'inégalité $\mathbf{T}_1$. Les articles [4], [8] et [14] sont consacrés aux inégalités de la forme $\mathcal{T}_c \leqslant H$, classe qui comprend par exemple l'inégalité $\mathbf{T}_2$ et dont l'étude est beaucoup plus délicate.

3.1 Le rôle de l'intégrabilité

L'article [2] fait suite à l'article [DGW04] de Djellout, Guillin et Wu et à [BV05] de Bolley et Villani. L'objet de [2] est de montrer qu'une grande classe d'inégalités de transport de la forme $\alpha(\mathcal{T}_c) \leqslant H$, faisant intervenir une fonction α non-linéaire, est équivalente à une condition d'intégrabilité exponentielle de la fonction de coût du type

$$\exists \varepsilon > 0,\ \exists y \in \mathcal{X}, \qquad \int e^{\alpha(\varepsilon c(x,y))}\,\mu(dx) < +\infty.$$

Ce critère sera valable entre autre sous la condition que $\alpha(t) = O(t^2)$ pour t proche de 0, ce qui inclut par exemple l'inégalité $\mathbf{T}_1$ mais exclut l'inégalité $\mathbf{T}_2$.

3.1.1 Résultat principal

Soit $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ une fonction vérifiant l'hypothèse 2.3.5 et introduisons l'espace d'Orlicz $\mathbb{L}_{\tau_\alpha}(\mathcal{X}, \mu)$ associé à la fonction $\tau_\alpha(u) = e^{\alpha(|u|)} - 1$, $u \in \mathbb{R}$. Rappelons que

$$\mathbb{L}_{\tau_\alpha}(\mathcal{X}, \mu) = \left\{ f : \mathcal{X} \to \mathbb{R};\ \exists \lambda > 0,\ \int \tau_\alpha(f(x)/\lambda)\,\mu(dx) < \infty \right\}.$$

Cet espace vectoriel est muni de la norme de Luxembourg $\|\cdot\|_{\tau_\alpha}$ défini comme suit

$$\|f\|_{\tau_\alpha} = \inf\left\{\lambda > 0;\ \int \tau_\alpha(f(x)/\lambda)\,\mu(dx) \leqslant 1\right\}.$$

Nous renvoyons à [RR91] pour ces définitions.

Le résultat principal de [2] est le suivant

Théorème 3.1.1. *Soient* $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ *et* α *des fonctions vérifiant l'hypothèse 2.3.5 ; supposons par ailleurs que* $\alpha(t) = O(t^2)$ *lorsque* $t \to 0$ *et que* $\sup\{\alpha^*(t); t > 0$ *tel que* $\alpha^*(t) < +\infty\} = +\infty$ *et posons* $c(x,y) = \theta(d(x,y))$, $x, y \in \mathcal{X}$. *Sous ces hypothèses, les propositions suivantes sont équivalentes :*

1. *Il existe* $a > 0$, *tel que* μ *vérifie l'inégalité de transport suivante*

$$\alpha\left(\frac{\mathcal{T}_c(\nu,\mu)}{a}\right) \leqslant H(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

2. *Il existe* $x_o \in \mathcal{X}$ *tel que* $c(x_o, \cdot\,) \in \mathbb{L}_{\tau_\alpha}(\mathcal{X},\mu)$.

Par ailleurs, dans ces conditions, on peut prendre

$$a = K_\theta m_\alpha \inf_{x_o \in \mathcal{X}} \|c(x_o, \cdot\,)\|_{\tau_\alpha},$$

où $K_\theta = \sup_{t>0} \theta(2t)/\theta(t)$ *et où* m_α *est une constante ne dépendant que de* α.

De plus, notons que la norme de Luxembourg $\|\cdot\|_{\tau_\alpha}$ peut être estimée grâce au lemme suivant, légèrement adapté de [2].

Lemme 3.1.2. *Pour toute fonction* $f \in \mathbb{L}_{\tau_\alpha}(\mathcal{X},\mu)$, *on a*

$$\|f\|_{\tau_\alpha} \leqslant \inf_{\varepsilon>0} \frac{1}{\varepsilon}\left[1 + \frac{\log \int e^{\alpha(\varepsilon f(x))}\,\mu(dx)}{\log(2)}\right]^{1/r_\alpha} \leqslant 2^{1/r_\alpha}\|f\|_{\tau_\alpha},$$

où l'exposant r_α *est défini dans le lemme 2.3.6.*

Avant de dire quelques mots sur la preuve du théorème 3.1.1, faisons quelques commentaires.

– Le théorème s'applique notamment pour l'inégalité $\mathbf{T}_1$, laquelle est donc équivalente à la condition $\int e^{\varepsilon d^2(x_o,y)}\,\mu(dy) < +\infty$. Ce résultat avait préalablement été démontré par Djellout, Guillin et Wu dans [DGW04]. Le théorème 3.1.1 permet également de récupérer les résultats de Bolley et Villani [BV05]. Plus précisément l'article [BV05] traitait uniquement le cas $c(x,y) = d^p(x,y)$, $p \geqslant 1$ avec $\alpha(t) = t^2$ ou $\alpha(t) = \min(t^2, t)$, $t \geqslant 0$.
– En revanche, le résultat précédent ne s'applique pas à l'inégalité $\mathbf{T}_2$, ou à toute autre inégalité de transport $\alpha(\mathcal{T}_c) \leqslant H$ associée à une fonction α linéaire près de 0. Cette limitation est structurelle, car l'inégalité $\mathbf{T}_2$ entraîne l'inégalité de Poincaré laquelle n'est pas réductible à une condition d'intégrabilité.
– Le théorème 3.1.1 donne une condition simple à vérifier pour déterminer si la mesure de référence μ satisfait une inégalité de transport. Malheureusement, la constante a fournie par ce théorème est en général assez mauvaise. Elle peut notamment dépendre fortement de la dimension de l'espace. Considérons, en effet la mesure gaussienne standard γ^n sur $(\mathbb{R}^n, \|\cdot\|_2)$. Pour tout $\varepsilon \in (0, 1/2)$, on a $\int e^{\varepsilon\|y-x_o\|_2^2}\,\gamma^n(dy) = M^n$, avec

$M = \int e^{\varepsilon(y-x_o)^2}\gamma(dy) < +\infty$. Le théorème précédent indique alors que γ^n vérifie $\mathbf{T}_1(a)$ avec $a = \kappa\sqrt{1+n\log(M)}$, où κ est un facteur numérique, alors qu'on sait que la constante optimale dans l'inégalité $\mathbf{T}_1$ pour la gaussienne vaut 2 en toute dimension.

3.1.2 Quelques mots sur la preuve

Le premier ingrédient est contenu dans la proposition suivante qui permet de majorer un coût de transport $\mathcal{T}_c$ associé à une fonction de coût $c(x,y) = \theta(d(x,y))$ par une norme en variation totale pondérée. Cette technique apparaît également dans [BV05] et remonte à Rachev-Rüschendorff [RR98, Lemma 10.2.3].

Proposition 3.1.3. *Si $c(x,y) = \theta(d(x,y))$ avec θ vérifiant l'hypothèse 2.3.5, alors, pour tout $x_o \in \mathcal{X}$, on a*

$$\mathcal{T}_c(\mu,\nu) \leqslant \|\chi_{x_o}\cdot(\mu-\nu)\|, \qquad \forall \mu,\nu \in \mathcal{P}(\mathcal{X}),$$

où χ_{x_o} est la fonction définie par $\chi_{x_o}(y) = \frac{1}{2}\theta(2d(x_o,y))$, $y \in Y$.

Cette majoration assez élémentaire ramène l'étude à la question de savoir sous quelles conditions une probabilité μ sur $\mathcal{X}$ vérifie une inégalité de Csiszar-Kullback-Pinsker pondérée de la forme

$$\alpha(\|\chi\cdot(\mu-\nu)\|) \leqslant H(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}), \tag{3.1.4}$$

où $\chi : \mathcal{X} \to \mathbb{R}^+$ est une fonction mesurable positive fixée. Une réponse générale est donnée dans le théorème qui suit, issu de [2].

Théorème 3.1.5. *Soit Φ une famille de fonctions mesurables bornées sur $\mathcal{X}$ telle que $\Phi = -\Phi$. Posons $\|\nu-\mu\|_\Phi^* = \sup_{\varphi\in\Phi}\{\int \varphi\, d\nu - \int \varphi\, d\mu\}$ et considérons une fonction α vérifiant l'hypothèse 2.3.5 et telle que $\alpha(t) = O(t^2)$ lorsque $t \to 0$ et $\sup\{\alpha^*(t); t > 0$ tel que $\alpha^*(t) < +\infty\} = +\infty$. Pour une mesure de probabilité $\mu \in \mathcal{P}(\mathcal{X})$, les propositions suivantes sont équivalentes :*

1. *Il existe $a > 0$ tel que μ vérifie*

$$\alpha\left(\frac{\|\nu-\mu\|_\Phi^*}{a}\right) \leqslant H(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}), \tag{3.1.6}$$

2. *Il existe $M > 0$ tel que pour tout $\varphi \in \Phi$, $\|\varphi - \int \varphi\, d\mu\|_{\tau_\alpha} \leqslant M$.*

De plus 1. entraîne 2. avec $M = 3a$ et 2. entraîne 1. avec $a = m_\alpha M$, où m_α est une constante ne dépendant que de la fonction α.

En particulier, notons que la constante optimale a_{opt} dans l'inégalité (3.1.6) vérifie l'encadrement

$$\frac{m_\alpha}{3}\sup_{\varphi\in\Phi}\left\|\varphi - \int\varphi\, d\mu\right\|_{\tau_\alpha} \leqslant a_{opt} \leqslant m_\alpha \sup_{\varphi\in\Phi}\left\|\varphi - \int\varphi\, d\mu\right\|_{\tau_\alpha}.$$

Le théorème précédent englobe bien sûr le cas des inégalités de Csiszar-Kullback-Pinsker pondérées (3.1.4), puisque $\|\chi\cdot(\mu-\nu)\|_{VT} = \sup_{|\varphi|\leqslant\chi}\left\{\int\varphi\, d\nu - \int\varphi\, d\mu\right\}$. Par ailleurs, $\sup_{|\varphi|\leqslant\chi}\|\varphi - \int\varphi\, d\mu\|_{\tau_\alpha} \leqslant 2\|\chi\|_{\tau_\alpha}$ et il est facile de voir que le théorème 3.1.5 associé à la proposition 3.1.3 entraîne le théorème 3.1.1.

La démonstration du théorème 3.1.5 passe tout d'abord par la réécriture de l'inégalité (3.1.6) sous une forme duale à la Bobkov-Götze : on montre par la même preuve que pour le théorème 2.2.1 que μ vérifie (3.1.6) si et seulement si, pour toute $\varphi \in \Phi$, on a

$$\int e^{s\varphi}\, d\mu \leqslant e^{s\int \varphi\, d\mu + \alpha^*(as)}, \qquad \forall s \geqslant 0.$$

Pour montrer que 2. implique 1. dans le théorème 3.1.5, on utilise le résultat suivant dû à Kozachenko et Ostrovskii [KO85] (voir aussi [BK00] p. 63-68).

Théorème 3.1.7. *Soit α une fonction vérifiant les hypothèses du théorème précédent ; si $\varphi \in \mathbb{L}_{\tau_\alpha}(\mathcal{X}, \mu)$, alors*

$$\int e^{s\varphi}\, d\mu \leqslant e^{s\int \varphi\, d\mu + \alpha^*(as)}, \qquad \forall s \geqslant 0,$$

avec $a = m_\alpha \|\varphi - \int \varphi\, d\mu\|_{\tau_\alpha}$, où la constante m_α dépend uniquement de la fonction α.

La preuve de ce théorème, qu'on peut lire dans ma thèse de doctorat, est assez technique. Néanmoins, il est facile d'obtenir la conclusion du théorème 3.1.7 dans des cas simples, $\alpha(t) = t^2$ notamment (voir [2]).

3.2 Une condition nécessaire et suffisante en dimension un

Les articles [4] et [14] portent sur les inégalités de transport $\mathcal{T}_c \leqslant H$ en dimension un.

De nombreuses inégalités fonctionnelles ont été complètement caractérisées en dimension un. C'est le cas notamment de l'inégalité de Sobolev logarithmique [BG99, BR03], de l'inégalité de Poincaré [Muc72], de l'inégalité de Cheeger [BH97a] ou de l'inégalité isopérimétrique Gaussienne [BL96]. Pour les inégalités de transport, le premier résultat allant dans ce sens est dû à Cattiaux et Guillin [CG06] ; il fournit une condition suffisante assez large pour qu'une mesure de probabilité μ sur $\mathbb{R}$ vérifie l'inégalité $\mathbf{T}_2$. Le résultat en question est le suivant.

Théorème 3.2.1. *Soit $\mu(dx) = e^{-V(x)}\, dx$ une mesure de probabilité sur $\mathbb{R}$ associée à une fonction V de classe $\mathcal{C}^2$, paire et telle que $V''(x)/(V'^2(x)) \to 0$ lorsque $x \to \infty$. Si la fonction V vérifie en plus la condition*

$$\liminf_{x \to +\infty} \frac{V'(x)}{x} > 0, \tag{3.2.2}$$

alors μ vérifie l'inégalité $\mathbf{T}_2(a)$, pour une certaine constante $a > 0$.

Grâce à cette condition, Cattiaux et Guillin ont réussi à construire le premier exemple d'une mesure de probabilité vérifiant $\mathbf{T}_2$ mais ne vérifiant pas l'inégalité de Sobolev logarithmique [CG06], répondant ainsi à une question restée ouverte pendant plusieurs années.

L'objectif principal de l'article [4] était de retrouver la condition (3.2.2) par une méthode complètement différente, utilisant notamment un argument de transport de masse, et plus généralement d'obtenir un critère analogue pour les inégalités de transport associées à des coûts $c(x, y) = \theta(|x - y|)$.

L'article [14], écrit cinq ans plus tard, reprend la méthode développée dans [4] et lui apporte une amélioration permettant d'aboutir à une vraie condition nécessaire et suffisante pour cette classe d'inégalités de transport. C'est ce résultat que nous allons mettre en avant dans cette section.

3.2.1 Contraction des inégalités de transport

Les articles [4] et [14] (et également [8]) utilisent une technique très simple permettant d'obtenir de nouvelles inégalités de transport à partir d'anciennes grâce à un transport de mesure. Le principe est expliqué dans la proposition qui suit, issue de [4] (voir aussi [Mau91]) :

Proposition 3.2.3. *Soient $\mathcal{X}$ et $\mathcal{Y}$ deux espaces polonais, $\mu_o \in \mathcal{P}(\mathcal{X})$ et $\mu \in \mathcal{P}(Y)$. On suppose que $T : \mathcal{X} \to \mathcal{Y}$ est une application mesurable transportant μ_o sur μ : $\mu = T_\# \mu_o$. Si μ_o vérifie l'inégalité de transport $\mathcal{T}_{c_o} \leqslant H$ pour une certaine fonction de coût c_o sur $\mathcal{X}$, alors μ vérifie l'inégalité de transport $\mathcal{T}_{c_o^T} \leqslant H$ avec*

$$c_o^T(y_1, y_2) = \inf\{c_o(x_1, x_2); T(x_1) = y_1,\ T(x_2) = y_2\}, \qquad \forall y_1, y_2 \in \mathcal{Y}.$$

En particulier, avec les notations de la proposition ci-dessus, si l'on souhaite montrer que μ vérifie l'inégalité $\mathcal{T}_c \leqslant H$, pour une fonction de coût c sur Y donnée, il suffit pour cela que $c \leqslant c_o^T$, ce qui se traduit par la propriété de contraction suivante pour T :

$$c(T(x_1), T(x_2)) \leqslant c_o(x_1, x_2), \qquad \forall x_1, x_2 \in \mathcal{X}.$$

Dans [4] et [14], l'argument précédent est utilisé en choisissant la mesure exponentielle symétrique comme mesure de référence $\mu_o(dx) = e^{-|x|}\,dx/2$ qui, d'après un résultat de Talagrand [Tal96], vérifie une inégalité de transport avec un coût proportionnel à $c(x,y) = \min(|x-y|^2, |x-y|)$. Un résultat élémentaire de transport optimal, donne par ailleurs l'existence d'une application de transport T, obtenue par réarrangement croissant, transportant μ_o sur la mesure étudiée μ. Les deux articles [4] et [14] exploitent l'idée que plus cette application de transport est contractante et plus l'inégalité de transport vérifiée par μ est forte.

Remarque 3.2.4. *Soulignons que la technique de contraction des inégalités fonctionnelles est assez classique et a été utilisée par de nombreux auteurs. Citons quelques exemples d'applications proches de notre sujet. Dans [Mau91], Maurey énonce la proposition 3.2.3 dans le cadre des inégalités d'inf-convolutions. Talagrand utilise dans [Tal94] un argument de transport pour déduire les propriétés de concentration des mesures produit $\mu_p^n(dx) = \frac{1}{Z_p^n} e^{-\|x\|_p^p}\,dx$, $p > 1$ de celles de la mesure exponentielle produit $\mu_1^n(dx) = \frac{1}{2^n} e^{-\|x\|_1}\,dx$. Bobkov et Houdré utilisent également un argument de transport (avec la mesure exponentielle comme mesure source) pour caractériser l'inégalité de Cheeger sur $\mathbb{R}$ [BH97a] et pour délimiter la classe des mesures de probabilité possédant la propriété de concentration adimensionnelle en norme infinie [BH00].*

3.2.2 Résultat principal

Le résultat principal de [14] est le théorème 3.2.6 suivant, qui caractérise une grande classe d'inégalités de transport de la forme $\mathcal{T}_c \leqslant H$ sur la droite réelle. Ce résultat s'applique notamment à l'inégalité $\mathbf{T}_2$.

Quelques notations seront utiles pour l'énoncer. Pour toute mesure de probabilité μ sur $\mathbb{R}$, on notera F_μ sa fonction de répartition, définie par $F_\mu(x) = \mu(]-\infty, x])$, $x \in \mathbb{R}$. Rappelons que si $\mu \in \mathcal{P}(\mathbb{R})$ est sans atome et $\nu \in \mathcal{P}(\mathbb{R})$ est une mesure de probabilité arbitraire, l'application $T_{\mu,\nu}$ définie par

$$T_{\mu,\nu}(x) = F_\nu^{-1} \circ F_\mu(x), \qquad \forall x \in \mathbb{R}, \tag{3.2.5}$$

avec $F_\nu^{-1}(t) = \inf\{x \in \mathbb{R}; F_\nu(x) \geqslant t\}$, $t \in [0,1]$, transporte μ sur ν : $\nu = T_{\mu,\nu\,\#}\mu$.

Théorème 3.2.6. *Soient $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ une fonction convexe telle qu'il existe $h > 0$ pour lequel $\theta(t) = t^2$ pour tout $0 \leqslant t \leqslant h$.*
Pour une probabilité $\mu \in \mathcal{P}(\mathbb{R})$, les propositions suivantes sont équivalentes :

1. *Il existe $a > 0$ tel que μ vérifie l'inégalité de transport $\mathcal{T}_{c_a} \leqslant H$, avec la fonction de coût*
$$c_a(x,y) = \theta(a|x-y|).$$
2. *Il existe $\lambda, d > 0$ tels que*
   (a) *La probabilité μ vérifie l'inégalité de Poincaré avec la constante λ*
   $$\lambda \mathrm{Var}_\mu(f) \leqslant \int |\nabla f|^2 \, d\mu, \qquad \forall f \text{ Lipschitz},$$
   où $|\nabla f|(x) := \limsup_{y \to x} \frac{|f(y)-f(x)|}{|y-x|}$.
   (b) *L'application $T = T_{\mu_1,\mu}$ envoyant la mesure exponentielle symétrique $\mu_1(dx) = e^{-|x|}\,dx/2$ sur μ vérifie*
   $$|T(x) - T(y)| \leqslant \frac{1}{d}\theta^{-1}(h^2 + |x-y|), \qquad \forall x, y \in \mathbb{R}.$$

Par ailleurs, les constantes optimales $a_{opt}, \lambda_{opt}, d_{opt}$ dans les inégalités ci-dessus sont reliées entre elles par l'inégalité

$$\kappa_1 \min(\sqrt{\lambda_{opt}}, d_{opt}) \leqslant a_{opt} \leqslant \kappa_2 \min(\sqrt{\lambda_{opt}}, d_{opt}),$$

où κ_1, κ_2 sont des constantes qui ne dépendent que de h.

En particulier, l'inégalité $\mathbf{T}_2$ est équivalente à l'existence d'un trou spectral et au caractère Hölder d'ordre 1/2 (pour les grandes valeurs) de l'application de transport T.

Le résultat précédent est de nature assez différente des conditions nécessaires et suffisantes pour l'inégalité de Sobolev logarithmique [BG99] ou l'inégalité de Poincaré [Muc72], obtenues par le biais d'inégalités de type Hardy (voir plus loin). En revanche, cette caractérisation rappelle les travaux de Bobkov [Bob94] et de Bobkov et Houdré [BH97a, BH00] caractérisant diverses inégalités fonctionnelles sur $\mathbb{R}$ au travers de propriétés lisibles sur l'application de transport T. Par exemple, il est établi dans [BH97a] qu'une mesure de probabilité μ vérifie l'inégalité de Cheeger sur $\mathbb{R}$:

$$\exists \beta > 0, \qquad \beta \int |f - m(f)| \, d\mu \leqslant \int |\nabla f| \, d\mu, \qquad \forall f \text{ Lipschitz},$$

$m(f)$ étant une médiane de f, si et seulement si l'application de transport T envoyant la mesure exponentielle symétrique μ_1 sur μ est Lipschitz. De plus, la meilleure constante β dans l'inégalité précédente est égale à l'inverse de la constante de Lipschitz de T.

3.2.3 Quelques mots sur la preuve

La raison principale pour laquelle une caractérisation des inégalités de transport est accessible en dimension un est que, dans ce cadre, le transport optimal est particulièrement simple, comme le montre le résultat suivant, qui remonte aux travaux d'Hoeffding, Frechet et Dall'Aglio [Hoe40, Fré60, Dal56] (voir la preuve dans [Vil03] ou [RR98]).

Théorème 3.2.7. *Soit* $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ *une fonction convexe telle que* $\theta(0) = 0$ *et* $\mu \in \mathcal{P}(\mathbb{R})$ *une probabilité ne possédant pas d'atome. Pour toute* $\nu \in \mathcal{P}(\mathbb{R})$ *telle que* $\iint \theta(|x-y|)\,\mu(dx)\nu(dy) < +\infty$, *l'application* $T_{\mu,\nu}$ *définie par* (3.2.5) *réalise le transport optimal de* μ *sur* ν *pour le coût* $c(x,y) = \theta(|x-y|)$. *En d'autres termes, le couplage déterministe* $\pi^*(dxdy) = \delta_{T_{\mu,\nu}(x)}(dy)\mu(dx)$ *réalise l'infimum dans* $\mathcal{T}_c(\mu,\nu)$ *:*

$$\mathcal{T}_c(\mu,\nu) = \int \alpha(|x - T_{\mu,\nu}(x)|)\,\mu(dx).$$

Notons que ce couplage π^* reste optimal pour une plus grande famille de coûts [CSS76].

Voyons comment cette caractéristique du transport optimal unidimensionnel intervient dans la preuve du théorème 3.2.6. Plaçons nous pour simplifier dans le cas $\theta(t) = t^2$, et supposons que μ est absolument continue par rapport à Lebesgue.

On remarque tout d'abord qu'on peut écrire

$$\theta = \theta_1 + \theta_2,$$

avec $\theta_1(t) = t^2$ si $t \leqslant 1$ et $\theta_1(t) = 2t - 1$ si $t \geqslant 1$ et $\theta_2(t) = [t-1]_+^2$. Les fonctions θ_1 et θ_2 étant convexes, l'égalité ci-dessus s'étend aux coûts de transport optimaux, grâce au théorème 3.2.7

$$\mathcal{T}_2(\mu,\nu) = \mathcal{T}_{c^1}(\mu,\nu) + \mathcal{T}_{c^2}(\mu,\nu), \qquad \forall \nu \in \mathcal{P}(\mathbb{R}),$$

en notant $c^i(x,y) = \theta_i(|x-y|)$, $i = 1,2$. Cette décomposition du coût de transport quadratique induit une décomposition de l'inégalité $\mathbf{T}_2$: μ vérifie $\mathbf{T}_2(a)$, pour une certaine constante a si et seulement si μ vérifie $\mathcal{T}_{c^1_{a_1}} \leqslant H$ et $\mathcal{T}_{c^2_{a_2}} \leqslant H$, pour certaines constantes a_1 et a_2.

D'après un résultat de Bobkov, Gentil et Ledoux [BGL01] (voir plus loin le théorème 3.3.3), l'inégalité $\mathcal{T}_{c^1_{a_1}} \leqslant H$ est équivalente à l'inégalité de Poincaré, et la meilleure constante a_1 est égale à $\sqrt{\lambda_{opt}}$ à un facteur universel près.

On montre ensuite que l'inégalité $\mathcal{T}_{c^2_{a_2}}$ est équivalente à la propriété de type Hölder suivante pour l'application de transport $T = T_{\mu_1,\mu}$:

$$\exists d > 0,\ \forall x,y \in \mathbb{R}, \qquad |T(x) - T(y)| \leqslant \frac{1}{d}(1 + \sqrt{|x-y|}). \tag{3.2.8}$$

Justifions le caractère suffisant de cette propriété. La probabilité μ_1 vérifie l'inégalité de transport $\mathcal{T}_{c^1_\kappa} \leqslant H$ pour une certaine constante $\kappa > 0$. Cela résulte d'un résultat de Talagrand [Tal96], ou d'un autre de Maurey [Mau91]. On peut également invoqué le fait que μ_1 vérifie une inégalité de Poincaré et utiliser le théorème de Bobkov, Gentil, Ledoux. D'après la proposition 3.2.3, μ vérifie l'inégalité de transport $\mathcal{T}_c \leqslant H$, avec $c(x,y) = \theta_1(\kappa|T^{-1}(x) - T^{-1}(y)|)$. La propriété (3.2.8), entraîne alors facilement que $c(x,y) \geqslant [|x-y|/a_2 - 1]_+^2$, pour une certaine constante $a_2 > 0$, ce qui achève la preuve.

3.2.4 Exemples et contrexemples

Pour utiliser le théorème 3.2.6, il faut être en mesure de déterminer si une probabilité μ vérifie l'inégalité de Poincaré et d'étudier les propriétés de contraction de l'application transportant μ_1 sur μ.

Une condition sur la densité

Un résultat de Muckenhoupt [Muc72] (voir aussi [ABC+00]), montre que si $\mu \in \mathcal{P}(\mathbb{R})$ est absolument continue par rapport à la mesure de Lebesgue, et si

$$A_+ = \sup_{x \geqslant m} \mu[x, +\infty[\int_m^x \frac{1}{p(t)}\, dt < +\infty,$$

et

$$A_- = \sup_{x \leqslant m} \mu] - \infty, x[\int_x^m \frac{1}{p(t)}\, dt < +\infty,$$

où m est la médiane de μ et p sa densité, alors la probabilité μ vérifie l'inégalité de Poincaré avec une constante λ_{opt} optimale vérifiant $\max(A^-, A^+) \leqslant 1/\lambda_{\mathrm{opt}} \leqslant 4\max(A^-, A^+)$.

Une étude asymptotique élémentaire des intégrales ci-dessus permet de montrer (voir [ABC+00]) que si $\mu(dx) = e^{-V(x)}\, dx$, avec une fonction V paire (pour simplifier), de classe $\mathcal{C}^2$ telle que $V''(x)/(V'^2(x)) \to 0$, lorsque $x \to +\infty$ et telle que

$$\liminf_{x \to \infty} V'(x) > 0, \tag{3.2.9}$$

alors $A^{\pm} < \infty$ et μ vérifie l'inégalité de Poincaré.

Dans le même ordre d'idée, on montre dans [4] que la condition de contraction de l'application de transport T dans le théorème 3.2.6 peut se lire sur le comportement asymptotique de V' :

Théorème 3.2.10. *Soient* $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ *une fonction convexe de classe* $\mathcal{C}^2$ *telle que* $\theta(t) = t^2$ *sur* $[0, h]$, $h > 0$ *et* $\theta''(t)/(\theta'^2(t)) \to 0$, *quand* $t \to +\infty$ *et* μ *une mesure de probabilité sur* $\mathbb{R}$ *s'écrivant* $\mu(dx) = e^{-V(x)}\, dx$, *avec une fonction* V *paire (pour simplifier), de classe* $\mathcal{C}^2$ *telle que* $V''(x)/(V'^2(x)) \to 0$, *lorsque* $x \to +\infty$. *Si la fonction* V *vérifie la condition*

$$\liminf_{x \to +\infty} \frac{V'(x)}{\theta'(x)} > 0, \tag{3.2.11}$$

alors $T = T_{\mu_1, \mu}$ *vérifie la condition*

$$|T(x) - T(y)| \leqslant \frac{1}{d}\theta^{-1}(h^2 + |x - y|), \qquad \forall x, y \in \mathbb{R}$$

pour un certain $d > 0$.

La condition (3.2.11) impliquant (3.2.9), on conclut que (3.2.11) est une condition suffisante pour que μ vérifie l'inégalité de transport $\mathcal{T}_{c_a} \leqslant H$, avec $c_a(x, y) = \theta(a|x - y|)$, pour un certain $a > 0$. Dans le cas où $\theta(t) = t^2$, la condition (3.2.11) redonne la condition (3.2.2) de Cattiaux et Guillin.

Une probabilité vérifiant $\mathbf{T}_2$ et non log-Sobolev

Dans [CG06], Cattiaux et Guillin ont construit une probabilité vérifiant la condition (3.2.2) mais pas le critère de Bobkov-Götze pour l'inégalité **LSI**. L'exemple donné dans [CG06] présente des queues de distribution décroissant en $O(e^{-t^3})$. Dans [14], nous construisons un nouvel exemple possédant des queues de distributions gaussiennes.

3.3 Une condition suffisante en dimension quelconque

L'article [8] a pour but de généraliser les conditions suffisantes (3.2.11) à la Cattiaux et Guillin en dimension quelconque. Dans le cas de l'inégalité $\mathbf{T}_2$, on peut par exemple dégager la condition multidimensionnelle suivante [1] [8], qui redonne (3.2.2) en dimension un.

Théorème 3.3.1. *Soit* $\mu(dx) = e^{-V(x)}\,dx$ *une mesure de probabilité sur* $\mathbb{R}^k$ *avec* $V : \mathbb{R}^k \to \mathbb{R}$ *une fonction de classe* $\mathcal{C}^2$*. Il existe une constante numérique* c*, telle que si la fonction* V *vérifie pour un certain* $\varepsilon \in]0,1[$ *et* $u \in]0,1[$

$$\liminf_{x\to\infty} \sum_{i=1}^{d} \left[\varepsilon \left(\frac{\partial V}{\partial x_i}\right)^2 (x) - \frac{\partial^2 V}{\partial x_i^2}(x)\right] \frac{1}{1+ux_i^2} > ukc, \tag{3.3.2}$$

alors μ *vérifie l'inégalité* $\mathbf{T}_2(a)$ *pour un certain* $a > 0$.

L'intérêt de ce critère est qu'il peut être vérifié par des potentiels V qui ne sont pas uniformément convexes. Plusieurs variantes de ce résultat ont été développées par la suite par Cattiaux, Guillin et Wu dans [CGW10].

Le théorème 3.3.1 est obtenu par une technique de changement de métrique dans l'inégalité de Poincaré que nous allons maintenant exposer.

3.3.1 Inégalités de Poincaré pour des métriques non-euclidiennes

Le point de départ de [8], est le résultat suivant de Bobkov, Gentil et Ledoux [BGL01] qui encode l'inégalité de Poincaré par une inégalité de transport.

Théorème 3.3.3. *Soit* $\mu \in \mathcal{P}(\mathbb{R}^k)$ *une mesure de probabilité absolument continue par rapport à la mesure de Lebesgue. Les propriétés suivantes sont équivalentes.*

1. *La probabilité* μ *vérifie l'inégalité de Poincaré avec une constante* $\lambda > 0$ *;*
2. *La probabilité* μ *vérifie l'inégalité de transport* $\mathcal{T}_{c_a} \leqslant H$ *avec un coût* $c_a(x,y) = \theta_1(a\|x-y\|_2)$*, pour un certain* $a > 0$*, où* $\theta_1(t) = t^2$ *si* $0 \leqslant t \leqslant 1$ *et* $\theta_1(t) = 2t-1$*, si* $t \geqslant 1$.

De plus, les constantes sont reliées de la manière suivante : 1. entraîne 2. avec $a = \kappa\sqrt{\lambda}$*, où* κ *est une constante universelle, et 2. entraîne 1. avec* $\lambda = 2a^2$.

L'idée est ensuite de généraliser le théorème précédent à d'autres métriques que la métrique euclidienne. Considérons pour cela une fonction $\omega : \mathbb{R} \to \mathbb{R}$ telle que $\omega(x)/x$ est strictement

1. La condition donnée dans le théorème 3.3.1 se déduit facilement de la Proposition 3.5 de [8].

croissante sur $]0,+\infty[$, $\omega(x) \geqslant 0$, pour tout $x \geqslant 0$ et ω est impaire. Il sera pratique d'étendre ω à $\mathbb{R}^k$, $k \geqslant 1$, en posant $\omega(x) = (\omega(x_1), \ldots, \omega(x_k))$, pour tout $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$. La distance [2] δ_ω sur $\mathbb{R}^k$ est alors définie par

$$\delta_\omega(x,y) = \|\omega(x) - \omega(y)\|_2 = \sqrt{\sum_{i=1}^{k} |\omega(x_i) - \omega(y_i)|^2}, \qquad \forall x, y \in \mathbb{R}^k.$$

Définition 3.3.4. *Nous dirons que $\mu \in \mathcal{P}(\mathbb{R}^k)$ vérifie l'inégalité de Poincaré, avec la constante $\lambda > 0$, relativement à la distance δ_ω, si*

$$\lambda \mathrm{Var}_\mu(f) \leqslant \int |\nabla f|_\omega^2 \, d\mu,$$

pour toute fonction $f : \mathbb{R}^k \to \mathbb{R}$ Lipschitz pour la distance δ_ω, où la longueur du gradient $|\nabla f|_\omega$ est définie par

$$|\nabla f|_\omega(x) = \limsup_{y \to x} \frac{|f(y) - f(x)|}{\delta_\omega(x,y)}, \qquad \forall x \in \mathbb{R}^k.$$

Dans ce cas, nous écrirons que μ vérifie l'inégalité $\mathbf{SG}(\omega, \lambda)$.

Il est très facile de ramener l'inégalité $\mathbf{SG}(\omega, \lambda)$ à une inégalité de Poincaré pondérée ou à l'inégalité de Poincaré standard au travers d'un changement de variable.

Proposition 3.3.5. *Soit $\mu \in \mathcal{P}(\mathbb{R}^k)$ une probabilité absolument continue par rapport à la mesure de Lebesgue. Les propositions suivantes sont équivalentes :*

1. *La probabilité μ vérifie l'inégalité $\mathbf{SG}(\omega, \lambda)$, pour un certain $\lambda_1 > 0$.*
2. *La probabilité μ vérifie l'inégalité*

$$\lambda \mathrm{Var}_\mu(f) \leqslant \int \sum_{i=1}^{k} \frac{1}{\omega'(x_i)^2} \left(\frac{\partial f}{\partial x_i}(x) \right)^2 \mu(dx),$$

pour toute fonction $f : \mathbb{R}^k \to \mathbb{R}$, Lipschitz pour δ_ω.

3. *La probabilité $\bar{\mu} := \omega_\# \mu$ vérifie l'inégalité de Poincaré standard*

$$\lambda \mathrm{Var}_{\bar{\mu}}(f) \leqslant \int \|\nabla f\|_2^2 \, \bar{\mu}(dx),$$

pour toute fonction f Lipschitz.

Tout comme l'inégalité de Poincaré standard, l'inégalité $\mathbf{SG}(\omega, \cdot\,)$ est équivalente à une inégalité de transport, comme le montre le résultat suivant prouvé dans [8].

Théorème 3.3.6. *Soit $\mu \in \mathcal{P}(\mathbb{R}^k)$ une mesure de probabilité absolument continue par rapport à la mesure de Lebesgue. Les propriétés suivantes sont équivalentes.*

1. *La probabilité μ vérifie l'inégalité $\mathbf{SG}(\omega, \lambda)$ avec une constante $\lambda > 0$;*
2. *La probabilité μ vérifie l'inégalité de transport $\mathcal{T}_{c_a} \leqslant H$ avec un coût $c_a(x,y) = \theta_1(a\delta_\omega(x,y))$, pour un certain $a > 0$, où $\theta_1(t) = t^2$ si $0 \leqslant t \leqslant 1$ et $\theta_1(t) = 2t - 1$, si $t \geqslant 1$.*

2. à ne pas confondre avec la distance d_θ introduite au chapitre 2

De plus, les constantes sont reliées de la manière suivante : 1. entraîne 2. avec $a = \kappa\sqrt{\lambda}$*, où* κ *est une constante universelle, et 2. entraîne 1. avec* $\lambda = 2a^2$*.*

Ce résultat se déduit très facilement du théorème 3.3.3 de Bobkov, Gentil et Ledoux en utilisant le point 3. de la proposition précédente et le résultat de contraction 3.2.3

3.3.2 Liens avec les inégalités classiques

Comme nous l'avons vu, l'inégalité $\mathbf{SG}(\omega, \cdot)$ est équivalente à une inégalité de transport $\mathcal{T}_c \leqslant H$, mais la fonction c en question est assez inhabituelle. La proposition technique suivante, permet de la comparer à des coûts plus classiques.

Lemme 3.3.7. *Pour toute fonction* ω *comme ci-dessus, on a l'inégalité suivante*

$$\theta_1(a\delta_\omega(x,y)) \geqslant \theta_1\left(\frac{a}{\sqrt{k}}\right)\sum_{i=1}^{k}\theta_1 \circ \omega\left(\frac{|x_i - y_i|}{2}\right), \qquad \forall a > 0,\ \forall x, y \in \mathbb{R}^k.$$

En particulier, si on applique ce résultat à la fonction impaire ω_2 définie sur $\mathbb{R}^+$ par $\omega_2(x) = \max(x, x^2)$, $x \geqslant 0$, on obtient le corollaire suivant.

Corollaire 3.3.8. *Si une mesure de probabilité* μ *sur* $\mathbb{R}^k$ *absolument continue par rapport à la mesure de Lebesgue vérifie l'inégalité* $\mathbf{SG}(\omega_2, \lambda)$ *pour un certain* $\lambda > 0$*, alors elle vérifie l'inégalité* $\mathbf{T}_2(a)$*, avec* $a = \tau\omega_2(\sqrt{k/\lambda})$*, où* τ *est une constante universelle.*

Voyons maintenant comment obtenir la condition explicite (3.3.2). On dispose de plusieurs conditions explicites pour déterminer si une mesure de probabilité $\nu(dx) = e^{-W(x)}\,dx$ sur $\mathbb{R}^k$, associée à un potentiel W de classe $\mathcal{C}^2$ vérifie l'inégalité de Poincaré standard. Par exemple, si W vérifie la condition

$$\liminf_{x\to\infty} \varepsilon\|\nabla W(x)\|_2^2 - \Delta W(x) > 0, \tag{3.3.9}$$

pour un certain $\varepsilon \in]0, 1[$, alors ν vérifie l'inégalité de Poincaré (voir par exemple [BBCG08] pour une preuve élémentaire). Comme μ vérifie $\mathbf{SG}(\omega, \cdot)$ si et seulement si $\tilde{\mu} = \omega_{\#}\mu$ vérifie l'inégalité de Poincaré, il est assez facile de déduire de (3.3.9) (par changement de variable) une condition analogue pour $\mathbf{SG}(\omega, \cdot)$ (voir la proposition 3.5 de [8] pour une condition générale). En particulier, en travaillant avec la fonction ω_2, on peut voir que la condition (3.3.2) est suffisante pour $\mathbf{SG}(\omega_2, \cdot)$ et donc, d'après le corollaire ci-dessus, pour $\mathbf{T}_2$.

L'article [8] contient également un résultat établissant la hiérarchie entre certaines inégalités de type Sobolev et les inégalités $\mathbf{SG}(\omega, \cdot)$. En particulier, on dispose du résultat suivant, renforçant la conclusion du théorème d'Otto-Villani présenté au chapitre 2 :

Théorème 3.3.10. *Si une probabilité* μ *sur* $\mathbb{R}^k$ *absolument continue par rapport à la mesure de Lebesgue vérifie l'inégalité de Sobolev logarithmique pour une certaine constante* C*, alors elle vérifie l'inégalité* $\mathbf{SG}(\omega_2, \lambda$ *pour une certaine constante* λ*.*

Les constantes C et λ peuvent être reliées entre elles de façon précise à un facteur dépendant de la dimension près. On dispose en fait d'un résultat beaucoup plus général (théorème 5.4 de

[8]) montrant que si μ vérifie une inégalité de type Nash (également appelée inégalité Super Poincaré [Wan00]) de la forme suivante : pour toute fonction f localement Lipschitz,

$$\forall s \geqslant 1, \qquad \int f^2 \, d\mu \leqslant \beta(s) \int \|\nabla f\|_2^2 \, d\mu + s \left(\int |f| \, d\mu \right)^2 ,$$

où $\beta : [1, +\infty[\to \mathbb{R}^+$ est une fonction décroissante, alors elle vérifie une inégalité $\mathbf{SG}(\omega_\beta, \lambda_\beta)$, pour une certaine constante λ_β et une certaine fonction ω_β naturellement associées à la fonction β. La preuve de ces résultats repose sur les inégalités de type capacité/ mesure développées par Barthe, Cattiaux et Roberto dans une série de papiers [BCR06, BCR07].

Chapitre 4

Inégalités de transport - liens avec les grandes déviations et la concentration

Nous présentons dans ce chapitre des résultats développés dans [5], [6] et [11]. L'article [5] cherche à interpréter les inégalités de transport sous l'angle des grandes déviations. Cette interprétation est motivée par le fait que l'entropie relative est la fonction de taux dans le théorème de Sanov. L'article [6] reprend cette réflexion, axée sur la comparaison entre inégalités de déviation exactes et inégalités asymptotiques, pour montrer que les inégalités de transport $\mathcal{T}_c \leqslant H$ sont la forme fonctionnelle canonique de la concentration adimensionnelle. Si les articles [5, 6] privilégient les arguments probabilistes, une approche purement analytique est également possible. Celle-ci est mise en place dans la première partie de l'article [11].

4.1 Heuristique

Commençons par présenter l'heuristique commune aux articles [5] et [6].

Si $(X_n)_{n\in\mathbb{N}}$ est une famille de variables aléatoires réelles indépendantes et identiquement distribuées, telle que $\mathbb{E}[|X_1|] < +\infty$, la loi forte des grands nombres indique que $M_n = \frac{1}{n}\sum_{i=1}^n X_i$ converge presque sûrement vers $m := \mathbb{E}[X_1]$. En particulier, pour tout $r > 0$,

$$\mathbb{P}(M_n \geqslant m + r) \to 0, \qquad \text{lorsque } n \to +\infty.$$

Ces écarts à la loi des grands nombres peuvent être quantifiés de plusieurs manières.

Tout d'abord, si les variables X_n sont bornées, $|X_n| \leqslant 1$ pour fixer les idées, on dispose par exemple de la très classique inégalité de Hoeffding [Hoe63] :

$$\mathbb{P}(M_n \geqslant m + r) \leqslant e^{-n\frac{t^2}{2}}, \qquad \forall r \geqslant 0,\ \forall n \in \mathbb{N}^*.$$

On renvoie par exemple à [Mas07] pour une présentation d'autres inégalités classiques dues à Bernstein et Bennett.

D'un autre côté, la théorie des grandes déviations, et dans le cas présent le théorème de Cramér [Cra38], fournit une estimation asymptotique de la probabilité d'un grand écart à la moyenne :

Théorème 4.1.1. *Soit* $(X_n)_{n\in\mathbb{N}^*}$, X *une famille de variables aléatoires i.i.d telle que* $\mathbb{E}[e^{sX}] < +\infty$, *pour tout* $|s| \leqslant s_o$, *pour un certain* $s_o > 0$. *Alors pour tout,* $A \subset \mathbb{R}$,

$$-\inf_{a\in\mathring{A}} I_X(a) \leqslant \liminf_{n\to\infty} \frac{1}{n}\log\mathbb{P}(M_n\in A) \leqslant \limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(M_n\in A) \leqslant -\inf_{a\in\overline{A}} I_X(a),$$

où I_X *est la* transformée de Cramér *de la loi de* X *définie par*

$$I_X(a) = \sup_{s\in\mathbb{R}}\{sa - \log\mathbb{E}[e^{sx}]\}, \qquad \forall a\in\mathbb{R}.$$

Voyons comment la comparaison de ces deux résultats permet de démontrer (de manière très compliquée) l'inégalité de Csiszar-Kullback-Pinsker

$$\frac{1}{2}\|\nu-\mu\|_{VT}^2 \leqslant H(\nu|\mu), \qquad \forall \nu,\mu\in\mathcal{P}(\mathcal{X}).$$

Le petit raisonnement qui suit illustre un fait qui sera développé dans les sections suivantes : une inégalité de transport permet de démontrer des inégalités de concentration, qui en retour redonnent l'inégalité de transport initiale.

Démonstration. Appliquons le théorème de Cramér à un ensemble A de la forme $A = [a,+\infty[$

$$-I_X(a) \leqslant \liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(M_n\geqslant a).$$

Si les X_n vérifie $|X_n|\leqslant 1$, alors la borne d'Hoeffding donne en comparaison

$$\limsup_{n\to+\infty}\frac{1}{n}\log\mathbb{P}(M_n\geqslant a) \leqslant -\frac{[a-\mathbb{E}[X]]_+^2}{2}.$$

Par suite, en notant $p = \mathrm{Loi}(X)$, on obtient

$$\frac{1}{2}\left[a - \int x\,p(dx)\right]_+^2 \leqslant \sup_{s\in\mathbb{R}}\left\{sa - \log\int e^{sx}\,p(dx)\right\}, \qquad \forall a\in\mathbb{R}.$$

Si μ,ν sont deux probabilités sur un espace $\mathcal{X}$ et $f:\mathcal{X}\to[-1;1]$, alors en appliquant l'inégalité ci-dessus à $a = \int f\,d\nu$ et $p = f_\#\mu$, on obtient

$$\frac{1}{2}\left[\int f\,d\nu - \int f\,d\mu\right]_+^2 \leqslant \sup_{s\in\mathbb{R}}\left\{s\int f\,d\nu - \log\int e^{sf}\,d\mu\right\}$$

En prenant le supremum en f, le membre de gauche devient $\frac{1}{2}\|\nu-\mu\|_{VT}^2$ et le membre de droite devient $\sup_{h\in\mathcal{C}_b(\mathcal{X})}\{\int h\,d\nu - \log\int e^h\,d\mu\} = H(\nu|\mu)$. □

4.2 De la concentration aux inégalités de transport

Nous rappelons que $(\mathcal{X}, d)$ est un espace polonais.

4.2.1 Inégalités de transport et déviations pour les moyennes empiriques

Le résultat principal de [5] est le suivant. Il généralise le raisonnement fait dans la section précédente et montre l'équivalence entre inégalités de transport et bornes de déviations exactes (resp. asymptotiques).

Théorème 4.2.1. *Soit* $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ *une fonction de coût semi-continue inférieurement telle que* $c(x, x) = 0$*, et* $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ *une fonction convexe croissante telle que* $\alpha(0) = 0$*. Pour une probabilité* $\mu \in \mathcal{P}(\mathcal{X})$*, les propositions suivantes sont équivalentes :*

1. *La probabilité* μ *vérifie l'inégalité de transport* $\alpha(\mathcal{T}_c) \leqslant H$*.*
2. *Pour toute suite* $(X_i)_{i\in\mathbb{N}}$ *i.i.d de loi* μ *et toute fonction* $\varphi \in \mathcal{C}_b(\mathcal{X})$*,*
$$\frac{1}{n}\log \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} Q_c\varphi(X_n) > \int \varphi\, d\mu + r\right) \leqslant -\alpha(r), \qquad \forall n \in \mathbb{N},\ \forall r \geqslant 0,$$
où $Q_c\varphi(x) = \inf_{y\in\mathcal{X}}\{\varphi(y) + c(x, y)\}$*,* $x \in \mathcal{X}$*.*
3. *Pour toute suite* $(X_i)_{i\in\mathbb{N}}$ *i.i.d de loi* μ *et toute fonction* $\varphi \in \mathcal{C}_b(\mathcal{X})$*,*
$$\limsup_{n\to+\infty} \frac{1}{n}\log \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} Q_c\varphi(X_n) > \int \varphi\, d\mu + r\right) \leqslant -\alpha(r), \qquad \forall r \geqslant 0.$$

Dans le cas, où $c(x, y) = \rho(x, y)$ *est une distance semi-continue inférieurement, il y a équivalence entre*

1. *La probabilité* μ *vérifie l'inégalité de transport* $\alpha(\mathcal{T}_\rho) \leqslant H$*.*
2. *Pour toute suite* $(X_i)_{i\in\mathbb{N}}$ *i.i.d de loi* μ *et toute fonction* φ *mesurable et* 1*-Lipschitz pour* ρ*,*
$$\frac{1}{n}\log \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} \varphi(X_n) > \int \varphi\, d\mu + r\right) \leqslant -\alpha(r), \qquad \forall n \in \mathbb{N},\ \forall r \geqslant 0.$$
3. *Pour toute suite* $(X_i)_{i\in\mathbb{N}}$ *i.i.d de loi* μ *et toute fonction* φ *mesurable et* 1*-Lipschitz pour* ρ*,*
$$\limsup_{n\to+\infty} \frac{1}{n}\log \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} \varphi(X_n) > \int \varphi\, d\mu + r\right) \leqslant -\alpha(r), \qquad \forall r \geqslant 0.$$

L'exemple de l'inégalité de Csiszar-Kullback-Pinsker traité dans la section précédente correspond ici à la distance $\rho(x, y) = \mathbf{1}_{x\neq y}$ et $\alpha(t) = t^2/2$.

4.2.2 Inégalités de transport et bornes de déviations pour la mesure empirique

Le résultat précédent reposait sur le théorème de Cramér. Il est également possible de travailler un étage au dessus en faisant jouer un rôle analogue au théorème de Sanov [San61] dont nous rappelons l'énoncé ci-dessous. On équipera l'espace $\mathcal{P}(\mathcal{X})$ de la topologie de la convergence étroite, c'est-à-dire la moins fine rendant continues les applications $\nu \mapsto \int \varphi \, d\nu$, $\varphi \in \mathcal{C}_b(\mathcal{X})$.

Théorème 4.2.2. *Soit* $(X_i)_{i\in\mathbb{N}^*}$ *une suite i.i.d à valeurs dans* $\mathcal{X}$ *de loi* $\mu \in \mathcal{P}(\mathcal{X})$. *La mesure empirique* $L_n := \frac{1}{n}\sum_{i=1}^n \delta_{X_i} \in \mathcal{P}(\mathcal{X})$ *vérifie un principe de grandes déviations gouverné par la fonction de taux Entropie Relative* $H(\,\cdot\,|\mu)$. *En d'autres termes, pour tout ensemble* $F \subset \mathcal{P}(\mathcal{X})$ *fermé,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(L_n \in F) \leqslant -\inf_{\nu\in F} H(\nu|\mu),$$

et pour tout ensemble $O \subset \mathcal{P}(\mathcal{X})$ *ouvert,*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(L_n \in O) \geqslant -\inf_{\nu\in O} H(\nu|\mu).$$

Sur le même modèle que l'implication 3. $\Rightarrow$ 1. dans le théorème 4.2.1, on peut montrer le résultat suivant [5] :

Proposition 4.2.3. *Soit* $c : \mathcal{X}\times\mathcal{X} \to \mathbb{R}^+$ *une fonction de coût semi-continue inférieurement,* $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ *une fonction continue croissante quelconque, et* $(X_i)_{i\in\mathbb{N}^*}$ *une suite i.i.d de loi* $\mu \in \mathcal{P}(\mathcal{X})$. *Si la probabilité* μ *vérifie*

$$\limsup_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\mathcal{T}_c(L_n,\mu) > r) \leqslant -\alpha(r), \qquad \forall r \geqslant 0,$$

alors μ *vérifie l'inégalité de transport* $\alpha(\mathcal{T}_c) \leqslant H$.

D'autres liens entre les inégalités de transport et des contrôles quantitatifs dans le théorème de Sanov sont également apparus dans [BGV07].

Dans [6], la proposition 4.2.3 a été utilisée pour montrer que les inégalités de transport étaient équivalentes à des propriétés de concentration pour les mesures produit.

Théorème 4.2.4. *Soient* $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ *une fonction vérifiant l'hypothèse 2.3.5 et* $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ *une fonction convexe et strictement croissante telle que* $\alpha(0) = 0$. *Pour une probabilité* $\mu \in \mathcal{P}(\mathcal{X})$, *les propositions suivantes sont équivalentes :*

1. *La probabilité* μ *vérifie l'inégalité* $\alpha(\mathcal{T}_c) \leqslant H$, *pour la fonction de coût* $c(x,y) = \theta(d(x,y))$, $x, y \in \mathcal{X}$.
2. *Il existe une suite* $(r_o^k)_{k\in\mathbb{N}^*}$ *de nombres positifs vérifiant* $r_o^k/(k^{1/p_\theta}) \to 0$ *lorsque* $k \to +\infty$, *telle que pour tout* $n \in \mathbb{N}^*$,

$$\mu^n\left(f > m(f) + r\right) \leqslant \exp\left(-n\alpha\left(\frac{(r - r_o^n)^{p_\theta}}{n}\right)\right), \qquad \forall r \geqslant r_o^n, \tag{4.2.5}$$

pour toute fonction $f : \mathcal{X}^n \to \mathbb{R}$ 1-*Lipschitz pour la distance* d_{θ,p_θ} *définie par*

$$d_{\theta,p_\theta}(x,y) = \left[\sum_{i=1}^n d_\theta^{p_\theta}(x_i,y_i)\right]^{1/p_\theta} = \left[\sum_{i=1}^n \theta(d(x_i,y_i))\right]^{1/p_\theta},$$

où l'exposant $p_\theta = \sup_{x>0} \frac{x\theta'_-(x)}{\theta(x)} \geqslant 1$ *a été introduit au lemme 2.3.6.*

Rappelons que l'implication 1. $\Rightarrow$ 2. est adaptée de l'argument de Marton et fait l'objet du théorème 2.3.7. Dans ce cas, on a $r_o^k = \left(k\alpha^{-1}(\log(2)/k)\right)^{1/p_\theta}$. Signalons un cas particulier important : l'inégalité $\mathbf{T}_2$ est équivalente à la concentration gaussienne adimensionnelle.

Corollaire 4.2.6. *Pour une probabilité μ sur $\mathcal{X}$, les propositions suivantes sont équivalentes :*

1. *La probabilité μ vérifie l'inégalité $\mathbf{T}_2(a)$ pour une certaine constante $a > 0$.*
2. *Il existe une suite $r_o^k \geqslant 0$, vérifiant $r_o^k/\sqrt{k} \to 0$ lorsque $k \to \infty$, telle que pour tout $n \in \mathbb{N}^*$,*

$$\mu^n(f \geqslant m(f) + r) \leqslant \exp\left(-\frac{1}{a}(r - r_o^n)^2\right), \qquad \forall r \geqslant r_o^n,$$

pour toute fonction $f : \mathcal{X}^n \to \mathbb{R}$ 1-Lipschitz pour la distance

$$d_2(x,y) = \left[\sum_{i=1}^n d^2(x_i,y_i)\right]^{1/2}, \qquad x,y \in \mathcal{X}^n.$$

Pour mettre ces résultats en perspective, rappelons quelques unes des nombreuses inégalités fonctionnelles introduites ces dernières années pour démontrer des résultats de concentration adimensionnels. On peut citer l'inégalité de Poincaré [BL97, GM83] associée à de la concentration de type exponentiel, l'inégalité de Sobolev logarithmique donnant de la concentration gaussienne [Led96, BG99] ainsi que ses versions modifiées balayant un large spectre de concentration [BR08, BL97, BZ05, GGM05]. On peut citer également les inégalités de Beckner-Latała-Oleszkiewicz [BCR06, BR03, Bec89, LO00] qui interpolent entre Poincaré et log-Sobolev, ou les inégalités d'inf-convolutions [Mau91, LW08] (qui sont très proches des inégalités de transport). L'intérêt théorique du théorème 4.2.4 est qu'il montre que les inégalités de transport sont le bon point de vue sur la question. Nous verrons par ailleurs dans le chapitre suivant qu'on peut utiliser le corollaire 4.2.6 pour donner une nouvelle preuve du théorème 2.4.4 d'Otto-villani.

Pour terminer, disons un mot de la preuve du théorème 4.2.4. Par le même argument de changement de métrique que celui utilisé dans la preuve du théorème 2.3.7, on peut se ramener au cas des inégalités de la forme $\alpha(\mathcal{T}_p) \leqslant H$, $p \geqslant 1$. L'idée est d'appliquer l'inégalité de concentration (4.2.5) à la fonction $F_n(x) = W_p(L_n^x, \mu)$, où $L_n^x = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$, $x \in \mathcal{X}^n$. Cette fonction est $1/n^{1/p}$-Lipschitz pour la distance $d_p(x,y) = \left[\sum_{i=1}^n d^p(x_i,y_i)\right]^{1/p}$ $x,y \in \mathcal{X}^n$. En effet, d'après l'inégalité triangulaire,

$$|F_n(x) - F_n(y)| \leqslant W_p(L_n^x, L_n^y) \leqslant \left[\frac{1}{n}\sum_{i=1}^n d^p(x_i,y_i)\right]^{1/p} = \frac{1}{n^{1/p}} d_p(x,y),$$

la deuxième inégalité venant du fait que le coût de transport $\mathcal{T}_p$ entre L_n^x et L_n^y est certainement inférieur au coût obtenu via l'application de transport envoyant x_i sur y_i, pour chaque i. Grâce à cela, on aboutit, après quelques calculs, à l'inégalité

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\mathcal{T}_p(L_n, \mu) > r) \leqslant -\alpha(r), \qquad \forall r \geqslant 0, \tag{4.2.7}$$

qui entraîne, d'après la proposition 4.2.3, l'inégalité $\alpha(\mathcal{T}_p) \leqslant H$.

Remarque 4.2.8. *La preuve montre au passage que l'estimation de grande déviation* (4.2.7) *est équivalente à l'inégalité de transport $\alpha(\mathcal{T}_c) \leqslant H$.*

4.3 Approche analytique

Dans l'article [11], nous montrons que les propriétés de concentration d'une mesure peuvent être encodées par une inégalité de transport non-tendue. Ces résultats sont très généraux, mais, par souci de simplicité, nous allons restreindre la présentation au cas gaussien.

4.3.1 Inégalité T_2 non tendue

Définition 4.3.1. *Nous dirons qu'une mesure de probabilité* $\mu \in \mathcal{P}(X)$ *vérifie l'inégalité* $\mathbf{T}_2(c_1, c_2)$ $c_1 > 0, c_2 \geqslant 0$ *si*

$$\mathcal{T}_2(\nu, \mu) \leqslant c_1 H(\nu|\mu) + c_2, \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Lorsque $c_2 \neq 0$, on parle d'inégalités $\mathbf{T}_2$ *non-tendues*. Ces inégalités sont équivalentes à la concentration gaussienne.

Théorème 4.3.2. *Soit* $\mu \in \mathcal{P}(\mathcal{X})$ *; les propositions suivantes sont équivalentes :*

1. *La probabilité* μ *vérifie l'inégalité de concentration gaussienne suivante :*

$$\mu(A_r) \geqslant 1 - Me^{-ar^2}, \qquad \forall r \geqslant 0, \tag{4.3.3}$$

pour tout $A \subset \mathcal{X}$ *tel que* $\mu(A) \geqslant 1/2$*, pour certaines constantes* $a, M \geqslant 0$.

2. *La probabilité* μ *vérifie l'inégalité* $\mathbf{T}_2(c_1, c_2)$*, pour certaines constantes* $c_1, c_2 \geqslant 0$.

De plus, les constantes sont reliées de la manière suivante :
1. entraîne 2. avec $c_1 = u/a$ *et* $c_2 = c(u)M/a$*, pour tout* $u > 1$*, avec* $c(u) = 4u/(u-1)$.
2. entraîne 1. avec $a = \varepsilon/c_1$ *et* $M = \left(\sqrt{\log(2)} + 2\sqrt{c_2/c_1}\right)^{\frac{2\varepsilon}{1-\varepsilon}}$*, pour tout* $\varepsilon \in [0, 1[$.

Faisons quelques commentaires sur ce résultat.

- L'implication $2 \Rightarrow 1$ est une simple adaptation de l'argument de Marton habituel.
- L'autre implication est plus originale. Remarquons en particulier que les constantes a, M, c_1, c_2 sont reliées en circuit fermé. Ce point sera crucial pour la suite.
- Ce théorème est à rapprocher du théorème 3.1.1, mais sa conclusion est beaucoup plus forte. En effet, on voit facilement que, sous l'hypothèse de concentration gaussienne (4.3.3), la constante
$$I_\varepsilon := \iint e^{\varepsilon d^2(x,y)} \, \mu(dx)\mu(dy) < +\infty,$$
pour $\varepsilon > 0$ assez petit. Le théorème 3.1.1 entraîne alors facilement l'inégalité $\mathbf{T}_2(c_1, c_2)$, mais avec des constantes c_1, c_2 s'exprimant en fonction de K (prendre, par exemple, $\alpha(t) = \min(t^2, t)$, $t \geqslant 0$). Or la constante I_ε ne peut pas s'exprimer uniquement en fonction de a et M ; comme nous l'avons vu dans la remarque suivant le théorème 3.1.1, elle peut notamment faire intervenir la dimension de l'espace (si $\mathcal{X} = \mathbb{R}^n$).

Avant de présenter quelques éléments de la preuve du théorème 4.3.2, voyons comment ce résultat permet de retrouver le corollaire 4.2.6 caractérisant la concentration gaussienne adimensionnelle.

Seconde preuve du corollaire 4.2.6. On suppose que μ vérifie la propriété de concentration gaussienne adimensionnelle : pour tout $n \geqslant 1$,

$$\mu^n(A_r) \geqslant 1 - Me^{-ar^2}, \qquad \forall r \geqslant 0,$$

pour tout $A \subset \mathcal{X}^n$, tel que $\mu^n(A) \geqslant 1/2$, où $A_r = \{x \in \mathcal{X}^n; d_2(x, A) \leqslant r\}$. D'après le théorème 4.3.2, cette propriété implique que, pour tout $n \in \mathbb{N}^*$, μ^n vérifie l'inégalité $\mathbf{T}_2(c_1, c_2)$, pour des constantes c_1, c_2 ne dépendant que de a et M, et donc, en particulier, indépendantes de la dimension n. Remarquons que si $\nu \in \mathcal{P}(X)$, alors $\mathcal{T}_2(\nu^n, \mu^n) = n\mathcal{T}_2(\nu, \mu)$ et $H(\nu^n|\mu^n) = nH(\nu|\mu)$. En appliquant à la mesure ν^n l'inégalité $\mathbf{T}_2(c_1, c_2)$ vérifiée par μ^n, on obtient par conséquent,

$$\mathcal{T}_2(\nu, \mu) \leqslant c_1 H(\nu|\mu) + \frac{c_2}{n}.$$

En faisant tendre n vers $+\infty$, on obtient $\mathbf{T}_2(c_1)$, ce qui achève la preuve. □

La preuve du théorème 4.3.2 (1. ⇒ 2.), fait d'abord appel à une écriture duale de l'inégalité $\mathbf{T}_2(c_1, c_2)$, facilement adaptée du cas classique (theorème 2.2.1) :

$$\mu \text{ vérifie } \mathbf{T}_2(c_1, c_2) \quad \Leftrightarrow \quad \int e^{\frac{1}{c_1} Qf}\, d\mu \leqslant e^{\frac{1}{c_1}\int f\, d\mu + \frac{c_2}{c_1}}, \qquad \forall f \in \mathcal{C}_b(\mathcal{X}),$$

où $Qf(x) = \inf_{y \in \mathcal{X}}\{f(y) + d^2(x, y)\}$, $x \in \mathcal{X}$. Le deuxième ingrédient consiste à traduire la propriété de concentration par un contrôle uniforme des queues de distribution des fonctions Qf :

$$\mu \text{ vérifie (4.3.3)} \quad \Leftrightarrow \quad \mu(Qf \geqslant m(f) + r) \leqslant Me^{-ar}, \quad \forall r \geqslant 0, \forall f \text{ minorée.}$$

Une intégration par partie montre alors que $\int e^{\frac{a}{u}(Qf - m(f))}\, d\mu \leqslant C(u, a, M) < +\infty$ pour tout $u > 1$. On conclut grâce à un petit argument technique permettant de remplacer la médiane par la moyenne.

4.3.2 Généralisation à des inégalités de transport non entropiques

Comme nous l'avons écrit plus haut, le cadre de [11] est plus large que ce que nous venons de présenter. Nous introduisons une famille d'inégalités de transport faisant intervenir d'autres fonctionnelles convexes sur l'espace $\mathcal{P}(\mathcal{X})$ et étudions leurs liens avec le phénomène de concentration de la mesure.

Plus précisément, à toute fonction $U : [0, +\infty[$ strictement convexe, continue en 0, telle que $U(1) \geqslant 0$ et vérifiant $U(x)/x \to +\infty$, on associe la fonctionnelle U_μ définie sur $\mathcal{P}(\mathcal{X})$ par

$$U_\mu(\nu) = \int U\left(\frac{d\nu}{d\mu}\right) d\mu, \qquad \forall \nu \in \mathcal{P}(\mathcal{X}),$$

généralisant l'entropie relative.

On dit que μ vérifie l'inégalité $\mathbf{T}_2U(c_1, c_2)$ si pour toute $\nu \in \mathcal{P}(\mathcal{X})$, on a

$$\mathcal{T}_2(\nu, \mu) \leqslant c_1 U_\mu(\nu) + c_2.$$

Le théorème de dualité de Bobkov et Götze peut se généraliser dans ce cadre (voir le théorème 2.6 et la proposition 2.9 de [11]), ainsi que l'argument de Marton (voir la proposition 2.3 de [11]) et on dispose également d'un résultat montrant que tel ou tel type de concentration entraîne telle ou telle inégalité de transport (voir le théorème 2.19 de [11]).

Chapitre 5

Inégalités de transport - liens avec les inégalités de type log-Sobolev

Ce chapitre tourne autour du théorème 2.4.4 d'Otto-Villani [OV00] montrant que l'inégalité de Sobolev logarithmique **LSI** entraîne l'inégalité $\mathbf{T}_2$. Les résultats présentés ici sont issus de la série d'articles [6, 12, 13, 15]. Dans [6], une nouvelle méthode de preuve fondée sur la concentration et sur la stabilité des inégalités par passage au produit tensoriel est proposée ; une variante est amenée dans [13]. Cette technique de preuve très robuste permet d'étendre le théorème d'Otto-Villani dans le cadre général des espaces métriques. L'article [12] vient compléter le théorème d'Otto-Villani en montrant que, dans l'espace euclidien, l'inégalité $\mathbf{T}_2$ est en réalité équivalente à l'inégalité de Sobolev logarithmique restreinte à une sous classe de fonctions. Cette équivalence est ensuite établie en toute généralité dans les articles [13, 15]. Elle permet de démontrer la stabilité des inégalités de transport par perturbation bornée.

Par souci de clarté, seule l'inégalité $\mathbf{T}_2$ sera abordée dans ce chapitre, mais tous les résultats présentés ci-dessous sont valables pour des inégalités de transport plus générales.

5.1 Le théorème d'Otto-Villani sur un espace métrique

5.1.1 Définitions

Commençons par étendre la définition de l'inégalité de Sobolev logarithmique à un espace métrique $(\mathcal{X}, d)$ quelconque. Pour cela, on doit donner un sens à $|\nabla f|$. On adoptera la définition classique suivante : si $f : \mathcal{X} \to \mathbb{R}$, on posera

$$|\nabla^- f|(x) := \limsup_{y \to x} \frac{[f(y) - f(x)]_-}{d(y, x)}, \tag{5.1.1}$$

si x n'est pas un point isolé dans $\mathcal{X}$ (si x est isolé, on pose $|\nabla^- f|(x) = 0$). On définit de même $|\nabla^+ f|$.

Nous dirons que μ vérifie l'inégalité de Sobolev logarithmique $\mathbf{LSI}(C)$, si

$$\mathrm{Ent}_\mu(f^2) \leqslant C \int |\nabla^- f|^2 \, d\mu,$$

pour toute fonction $f : \mathcal{X} \to \mathbb{R}$ Lipschitz.

Enonçons à présent la généralisation du théorème 2.4.4 à laquelle est consacrée cette section.

Théorème 5.1.2. *Soit μ une mesure de probabilité sur un espace polonais $(\mathcal{X}, d)$; si μ vérifie l'inégalité de Sobolev logarithmique $\mathbf{LSI}(C)$, alors μ vérifie l'inégalité $\mathbf{T}_2(C)$.*

5.1.2 Historique

Dans un cadre riemannien, deux preuves ont été proposées pour ce théorème. La preuve d'Otto-Villani [OV00] utilise une interpolation $(\nu_t)_{t\geqslant 0}$ entre $\nu_0 = \nu$ et $\nu_\infty = \mu$. L'interpolation en question est définie par $\nu_t(dx) = h(t,x)\,\mu(dx)$, où h est solution de l'équation de Fokker-Planck :

$$\frac{\partial h}{\partial t} = \Delta h - \nabla V \cdot \nabla h, \qquad h(0,x) = \frac{d\nu}{d\mu}(x).$$

où par hypothèse $\mu(dx) = e^{-V(x)}\,dx$, pour une fonction V régulière. La preuve repose sur le calcul des dérivées des fonctionnelles $H(\,\cdot\,|\mu)$ et $W_2(\,\cdot\,,\mu)$ le long de cette interpolation.

Peu après [OV00], Bobkov, Gentil et Ledoux [BGL01] ont proposé une nouvelle preuve (dans $\mathbb{R}^k$), plus élémentaire, utilisant la forme duale de l'inégalité $\mathbf{T}_2$:

$$\mu \text{ vérifie } \mathbf{T}_2(a) \quad \Leftrightarrow \quad \int e^{\frac{1}{a}Qf}\,d\mu \leqslant e^{\frac{1}{a}\int f\,d\mu}, \quad \forall f \in \mathcal{C}_b(\mathbb{R}^k),$$

où $Qf(x) = \inf_{y\in\mathbb{R}^k}\{f(y) + \|y-x\|_2^2\}$. La preuve repose cette fois sur une dérivation le long du semi-groupe d'Hamilton-Jacobi :

$$Q_t f(x) = \inf_{y\in\mathbb{R}^k}\left\{f(y) + \frac{1}{2t}\|x-y\|_2^2\right\},$$

qui interpole naturellement entre f et Qf et suit l'équation

$$\frac{\partial}{\partial t} Q_t f = -\frac{1}{2}\|\nabla Q_t f\|_2^2. \tag{5.1.3}$$

Ces deux techniques de preuves ont été utilisées dans d'autres cadres (toujours riemanniens) par plusieurs auteurs. Dans [Wan04a], la technique d'Otto et Villani est utilisée pour comparer les inégalités de Beckner-Latała-Oleszkiewicz aux inégalités de transport. La technique de Bobkov, Gentil et Ledoux est utilisée dans [GGM05] pour généraliser le théorème d'Otto-Villani aux inégalités de Sobolev logarithmiques modifiées. Voir également [CG06], pour une utilisation des deux techniques.

Lott et Villani [LV07] ont été les premiers à montrer que l'implication $\mathbf{LSI} \Rightarrow \mathbf{T}_2$ restait valable dans des espaces métriques assez généraux. Leur preuve exploite la technique de [BGL01] et passe notamment par une extension de l'équation d'Hamilton-Jacobi (5.1.3) à un cadre métrique. Lott et Villani obtiennent l'implication $\mathbf{LSI} \Rightarrow \mathbf{T}_2$ sous des conditions supplémentaires sur l'espace et la mesure μ (condition de doublement, inégalité de Poincaré locale, existence de courbes géodésiques dans $\mathcal{X}$). Le théorème 5.1.2 a été obtenu dans [6] par une preuve faisant appel à la caractérisation de l'inégalité $\mathbf{T}_2$ par la concentration gaussienne adimensionnelle (Corollaire 4.2.6). Une variante de cette preuve est apparue dans [13] ; elle évite le passage par la concentration et ne fait appel qu'à la stabilité des inégalités **LSI** et $\mathbf{T}_2$ par passage au produit tensoriel. Enfin, dans un papier très récent [15], nous montrons que l'approche par les équations d'Hamilton-Jacobi peut être étendue au cadre métrique sans faire appel aux hypothèses structurelles utilisées par Lott et Villani dans [LV07]. Signalons que très récemment, la preuve originale d'Otto-Villani a été étendue au cas métrique par Gigli et Ledoux [GL11], en tirant partie des avancées réalisées par Ambrosio, Gigli et Savaré sur les flots gradients (chemins dans l'espace $\mathcal{P}(\mathcal{X})$ généralisant l'interpolation ν_t de [OV00] présentée plus haut) dans un cadre métrique [AGS11].

5.1.3 Preuve par la concentration

L'idée développée dans [6] pour montrer que **LSI** implique $\mathbf{T}_2$ tient en quelques lignes. Un argument bien connu, dû à Herbst (voir par exemple [Led01]), montre que l'inégalité de Sobolev logarithmique entraîne de la concentration gaussienne indépendante de la dimension. Comme cette dernière propriété est équivalente à l'inégalité $\mathbf{T}_2$, l'implication est démontrée. Pour une probabilité μ sur $\mathbb{R}^k$ absolument continue par rapport à la mesure de Lebesgue, le raisonnement précédent s'applique sans aucun problème (voir [6]). En revanche, dans le cadre général des espaces métriques, ce raisonnement nécessite quelques vérifications techniques. Même si l'esprit en est le même, ce qui suit est légèrement adapté de [6] et n'a pas fait l'objet d'une publication. Commençons par rappeler l'argument de Herbst.

Théorème 5.1.4. *Soit μ une probabilité vérifiant* **LSI**(C) *; pour tout $n \in \mathbb{N}^*$,*

$$\mu^n\left(f > \int f\,d\mu^n + r\right) \leqslant e^{-\frac{r^2}{C}}, \qquad \forall r \geqslant 0,$$

pour toute fonction $f : \mathcal{X}^n \to \mathbb{R}$ vérifiant

$$\sum_{i=1}^n |\nabla_i^{\pm} f|^2(x) \leqslant 1, \qquad \forall x \in \mathcal{X}^n, \tag{5.1.5}$$

où

$$|\nabla_i^{\pm} f|(x) = \limsup_{y \to x_i} \frac{[f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) - f(x)]_{\pm}}{d(y, x_i)}.$$

Le premier argument dans la preuve du théorème 5.1.4 est que l'inégalité **LSI** tensorise bien : si μ vérifie **LSI**(C), alors μ^n vérifie **LSI**(C) sur l'espace $\mathcal{X}^n$, avec $\sum_{i=1}^n |\nabla_i^- f|^2$ dans le membre de droite. Le second argument consiste à appliquer l'inégalité **LSI** en dimension n à une fonction f de la forme $f = e^{\lambda g/2}$, où g est une fonction de moyenne nulle et vérifiant $\sum_{i=1}^n |\nabla_i^- g|^2 \leqslant 1$ et $\lambda > 0$. Il est alors facile d'obtenir une inégalité différentielle vérifiée par $Z(\lambda) = \log \int e^{\lambda g}\,d\mu^n$ qui entraîne $Z(\lambda) \leqslant \lambda^2/C$.

La concentration s'exprime au travers des inégalités de déviations de fonctions 1-Lipschitz pour la distance d_2. Il faut donc vérifier qu'une fonction de ce type vérifie bien la condition (5.1.5). C'est le but du lemme suivant.

Lemme 5.1.6. *Soit* $f : \mathcal{X}^n \to \mathbb{R}$*, une fonction* 1*-Lipschitz pour la distance* d_2*. Pour tout* $K \subset \mathcal{X}$ *compact de* $\mathcal{X}$*, et* $\varepsilon > 0$*, la fonction* $f_{K,\varepsilon}$ *définie par*

$$f_{K,\varepsilon}(x) = \inf_{z \in K} \left\{ f(z) + \sqrt{\varepsilon^2 + d_2^2(x,z)} \right\}, \qquad \forall x \in \mathcal{X}^n,$$

vérifie la condition $\sum_{i=1}^n |\nabla^+ f_{K,\varepsilon}|^2 \leqslant 1$ *et l'encadrement*

$$f(x) \leqslant f_{K,\varepsilon}(x) \leqslant f(x) + \varepsilon + 2d_2(x,K), \qquad \forall x \in \mathcal{X}^n.$$

Admettons le lemme un instant, et complétons la preuve du théorème 5.1.2.

Preuve du théorème 5.1.2. Soient f une fonction 1-Lipschitz pour la distance d_2 sur $\mathcal{X}^n$, $\varepsilon > 0$ et K un compact de $\mathcal{X}^n$. Le théorème 5.1.4 appliqué à la fonction $f_{K,\varepsilon}$ du lemme 5.1.6 entraîne que $\mu^n(f_{K,\varepsilon} > \int f_{K,\varepsilon}\, d\mu^n + r) \leqslant e^{-r^2/C}$, pour tout $r \geqslant 0$. En utilisant l'encadrement fourni par le lemme 5.1.6, on obtient

$$\mu^n \left(f > \int f \, d\mu^n + r + \varepsilon + 2 \int d_2(x,K)\, \mu^n(dx) \right) \leqslant e^{-r^2/C}, \qquad \forall r \geqslant 0. \tag{5.1.7}$$

Fixons un point $a \in \mathcal{X}^n$, et considérons la mesure de probabilité $\nu(dx) = \frac{1}{Z} d_2(a,x)\, \mu^n(dx)$, où Z est une constante de renormalisation[1]. Comme l'espace $\mathcal{X}^n$ est polonais, la probabilité ν est tendue. Il existe donc une suite croissante de compacts K_p telle que $a \in K_p$ et telle que $\nu(K_p^c) \to 0$, lorsque $p \to \infty$. Cela entraîne que

$$\int d_2(x,K_p)\, \mu^n(dx) \leqslant \int d_2(x,a) \mathbf{1}_{K_p^c}(x)\, \mu^n(dx) = Z\nu(K_p^c) \to 0,$$

lorsque $p \to +\infty$. En appliquant l'inégalité (5.1.7), avec la suite K_p et une suite ε_p décroissant vers 0, on obtient par le théorème de convergence monotone

$$\mu^n \left(f > \int f \, d\mu^n + r \right) \leqslant e^{-r^2/C}, \qquad \forall r \geqslant 0.$$

Il n'est pas difficile de déduire de cela une inégalité de déviation par rapport à la médiane de f :

$$\mu^n \left(f > m(f) + r \right) \leqslant e^{-(r-r_o)^2/C}, \qquad \forall r \geqslant r_o := \sqrt{C \log(2)}.$$

On conclut grâce au corollaire 4.2.6 que μ vérifie $\mathbf{T}_2(C)$. $\square$

Donnons maintenant la preuve du lemme 5.1.6.

Preuve du lemme 5.1.6. L'encadrement est laissé au lecteur. Montrons que $f_{K,\varepsilon}$ vérifie la condition $\sum_{i=1}^n |\nabla_i^+ f_{K,\varepsilon}|^2 \leqslant 1$. Fixons $x \in \mathcal{X}^n$, et pour tout $y \in \mathcal{X}$ et $i \in \{1, 2, \ldots, n\}$, posons $\bar{x}^i y =$

1. L'argument de Herbst pour $n = 1$ entraîne que les fonctions Lipschitz sont intégrables. En particulier $Z \leqslant \sum_{i=1}^n \int d(a_i, x)\, \mu(dx) < +\infty$.

$(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)$. Comme K est compact, il existe $a \in K$, tel que $f_{K,\varepsilon}(x) = f(a) + \sqrt{\varepsilon^2 + d_2^2(x,a)}$. On a aussi, $f_{K,\varepsilon}(\bar{x}^i y) \leqslant f(a) + \sqrt{\varepsilon^2 + d_2^2(\bar{x}^i y, a)}$. Par conséquent,

$$\begin{aligned} f(\bar{x}^i y) - f(x) &\leqslant \sqrt{\varepsilon^2 + d_2^2(\bar{x}^i y, a)} - \sqrt{\varepsilon^2 + d_2^2(x,a)} \\ &\leqslant d(y, x_i) \frac{d(y, a_i) + d(x_i, a_i)}{\sqrt{\varepsilon^2 + d_2^2(\bar{x}^i y, a)} + \sqrt{\varepsilon^2 + d_2^2(x,a)}} \end{aligned}$$

On en déduit que $|\nabla_i^+ f_{K,\varepsilon}|(x) \leqslant \frac{d(x_i,a_i)}{\sqrt{\varepsilon^2 + d_2^2(x,a)}}$, puis que $\sum_{i=1}^n |\nabla_i^+ f_{K,\varepsilon}|^2(x) \leqslant 1$. □

Remarque 5.1.8. *La preuve donnée dans [6], dans le cas métrique, est légèrement différente. L'argument de Herbst n'est appliqué qu'à la fonction* F_n *définie sur* $\mathcal{X}^n$ *par* $F_n(x) = W_2(L_n^x, \mu)$ *(qui est* $1/\sqrt{n}$*-Lipschitz pour la distance* d_2 *et vérifie* $\sum_{i=1}^n |\nabla_i^+ F_n|^2(x) \leqslant 1/n$ *- voir la proposition 4.4 de [6]). Cela conduit à la conclusion*

$$\limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{P}(\mathcal{T}_2(L_n, \mu) > r) \leqslant -r/C, \qquad \forall r \geqslant 0,$$

ce qui, d'après la proposition 4.2.3, implique que μ *vérifie* $\mathbf{T}_2(C)$.

L'article [13] apporte une petite variante. En appliquant **LSI** en dimension n à la fonction $f = e^{\lambda Qf/2}$, où $Qf(x) = \inf_{y \in \mathcal{X}^n}\{f(y) + d_2^2(x,y)\}$ et $f : \mathcal{X}^n \to \mathbb{R}^+$ est continue bornée positive, on obtient par un argument de type Herbst que

$$\int e^{\lambda Qf}\, d\mu^n \leqslant e^{\beta(\lambda) \int f\, d\mu^n}, \qquad \forall \lambda \in [0, 1/C[,$$

où $\beta(\lambda) = \frac{\lambda}{1 - C\lambda}$, $0 \leqslant \lambda < 1/C$. Comme $\beta(\lambda) > \lambda$, on ne peut pas appliquer le critère de Bobkov et Götze (théorème 2.2.1). On conclut alors à l'aide de la version suivante du critère de Bobkov-Götze, démontrée dans [13].

Lemme 5.1.9. *S'il existe* $a, b, c > 0$ *telles que pour tout* $n \in \mathbb{N}^*$, *et pour toute fonction* $f : \mathcal{X}^n \to \mathbb{R}^+$, *continue bornée positive, on a*

$$\int e^{cQf}\, d\mu^n \leqslant a e^{b \int f\, d\mu^n},$$

alors μ *vérifie* $\mathbf{T}_2(1/c)$.

5.1.4 Preuve par les équations d'Hamilton-Jacobi

Dans l'article [15], nous montrons que la méthode de [BGL01] peut être adaptée à un cadre métrique sans hypothèse de structure particulière. A cette fin, nous démontrons le théorème suivant qui généralise la formule de Hopf-Lax-Oleinik classique pour les équations d'Hamilton-Jacobi dans $\mathbb{R}^k$ [Eva10].

On supposera dans toute cette section que les boules fermées de $\mathcal{X}$ sont compactes. Cette hypothèse pourrait être enlevée, au prix de quelques complications techniques.

Théorème 5.1.10. *Pour toute fonction* $f : \mathcal{X} \to \mathbb{R}$ *semi-continue inférieurement et minorée, posons* $Q_t f(x) = \inf_{y \in \mathcal{X}} \left\{ f(y) + \frac{1}{2t} d^2(x,y) \right\}$*, pour tout* $x \in \mathcal{X}$*, et tout* $t > 0$.
On a toujours

$$\frac{d}{dt_+} Q_t f(x) \leqslant -\frac{|\nabla^- Q_t f|^2(x)}{2}, \qquad \forall t > 0,\ \forall x \in \mathcal{X}.$$

Si, en plus, l'espace $(\mathcal{X}, d)$ *est géodésique, alors il y a égalité dans la formule précédente.*

Rappelons que $(\mathcal{X}, d)$ est géodésique, si pour tout couple $(x, y) \in \mathcal{X}^2$, il existe au moins une courbe $(\gamma_t)_{t\in[0,1]}$ telle que $\gamma_0 = x$, $\gamma_1 = y$ et $d(\gamma_s, \gamma_t) = |t-s| d(x,y)$, $s,t \in [0,1]$. Une courbe de ce type porte le nom de géodésique à vitesse constante.

Faisons quelques commentaires sur le théorème 5.1.10.

– On peut considérer des Q_t plus généraux : $Q_t f(x) = \inf_{y\in\mathcal{X}}\{f(y) + t\theta(d(x,y)/t)\}$, avec $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ convexe, strictement croissante et de classe $\mathcal{C}^1$ telle que $\theta(0) = 0$. Voir le théorème 1.6 de [15].
– Le théorème 5.1.10 a été démontré indépendamment par Ambrosio, Gigli et Savaré dans [AGS11], par une preuve analogue. Il joue un rôle significatif dans leur étude des flots gradients dans des espaces métriques. Lott et Villani [LV07] démontrent également que $Q_t f$ vérifie l'équation d'Hamilton-Jacobi, mais sous une forme un peu plus faible et en imposant des hypothèses supplémentaires sur l'espace métrique mesuré (X, d, μ). Plus précisément, ils montrent que si l'espace est géodésique compact et si la mesure μ vérifie une inégalité de Poincaré locale, ainsi qu'une propriété de doublement, alors l'égalité
$$\frac{d}{dt_+} Q_t f(x) = -\frac{|\nabla^- Q_t f|^2(x)}{2}$$
a lieu pour tout $t > 0$ et pour tout x en dehors d'un ensemble N_t de mesure nulle pour μ. Le rôle de ces hypothèses est de permettre l'utilisation des résultats de différentiabilité de Cheeger [Che99].

La preuve que nous proposons du théorème 5.1.10 ne repose pas sur les outils développés dans [Che99] et reste à un niveau élémentaire. Le point clé est la proposition suivante, établie dans [15], qui donne une expression pour la dérivée à droite de $Q_t f$.

Proposition 5.1.11. *Si* $f : \mathcal{X} \to \mathbb{R}$ *est une fonction semi-continue inférieurement et minorée, alors on a toujours*

$$\frac{d}{dt_+} Q_t f(x) = -\max_{\bar{y}\in m(t,x)} \frac{d^2(x,\bar{y})}{2t^2},$$

où $m(t,x)$ *est l'ensemble, non-vide, des points* $\bar{y}$ *réalisant le minimum dans la définition de* $Q_t f$ *:*

$$\bar{y} \in m(t,x) \quad \Leftrightarrow \quad Q_t f(x) = f(\bar{y}) + \frac{1}{2t} d^2(x, \bar{y}).$$

La preuve est complétée en comparant $\max_{\bar{y}\in m(t,x)} d^2(x,\bar{y})/2t^2$ à $|\nabla^- Q_t f|^2(x)/2$.

5.2 Les inégalités de log-Sobolev restreintes

Les articles [11, 13, 15] abordent également la question d'une réciproque au théorème d'Otto-Villani. Nous montrons que l'inégalité $\mathbf{T}_2$ est équivalente à l'inégalité **LSI** restreinte à une sous classe de fonctions.

Le théorème suivant regroupe les différentes formes équivalentes de l'inégalité $\mathbf{T}_2$. Nous aurons besoin de la définition suivante. Soit $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ une fonction de coût ; une fonction $f : \mathcal{X} \to \mathbb{R} \cup \{\pm\infty\}$ est dite c-convexe si elle peut s'écrire sous la forme

$$f(x) = \sup_{y \in \mathcal{X}} \{g(y) - c(x, y)\}, \qquad \forall x \in \mathcal{X},$$

pour une certaine fonction $g : \mathcal{X} \to \mathbb{R} \cup \{\pm\infty\}$.

Théorème 5.2.1. *Soit* μ *une mesure de probabilité sur un espace* $(\mathcal{X}, d)$ *géodésique dont les boules fermées sont compactes.*
Les propositions suivantes sont équivalentes.

1. *Il existe* $C > 0$ *telle que* μ *vérifie* $\mathbf{T}_2(C)$.
2. *Il existe* $D > 0$ *telle que* μ *vérifie l'inégalité suivante : pour toute fonction* $f : \mathcal{X} \to \mathbb{R}$ *continue et bornée et pour tout* $\lambda \in]0, 1/D[$, *on a*

$$\mathrm{Ent}_\mu(e^f) \leqslant \frac{1}{1 - \lambda D} \int (f - Q_{1/\lambda} f) e^f \, d\mu \tag{5.2.2}$$

3. *Il existe* $E > 0$ *telle que* μ *vérifie l'inégalité suivante : pour toute fonction* $f : \mathcal{X} \to \mathbb{R}$ *qui est* $K\frac{d^2}{2}$*-convexe, avec* $0 < K < 1/E$,

$$\mathrm{Ent}_\mu(e^f) \leqslant \frac{2E}{(1 - KE)^2} \int |\nabla^+ f|^2 e^f \, d\mu \tag{5.2.3}$$

De plus, les constantes optimales dans les inégalités ci-dessus sont reliées par les relations

$$E_{opt} \leqslant D_{opt} \leqslant C_{opt} \leqslant \exp(2) E_{opt}.$$

Les inégalités (5.2.2) et (5.2.3) ont été introduites dans [11]. Nous avons appelé l'inégalité (5.2.2) *inégalité* (τ)*-log-Sobolev* en référence à la propriété (τ) de Maurey [Mau91]. L'inégalité (5.2.3) est appelée *inégalité log-Sobolev restreinte*. L'idée de restreindre l'inégalité de Sobolev logarithmique pour étudier $\mathbf{T}_2$ vient de Cattiaux et Guillin [CG06] (qui considéraient une autre classe de fonctions et n'obtenaient pas d'équivalence).

Le théorème précédent peut être étendu à une large famille d'inégalités de transport, notamment les inégalités $\mathcal{T}_p \leqslant CH$. Détaillons les apports successifs des articles [11, 13, 15] :

– L'article [11] prouve l'équivalence entre les trois propositions ci-dessus dans un cadre euclidien : $\mathcal{X} = \mathbb{R}^k$ muni de sa norme euclidienne standard $\| \cdot \|_2$. Remarquons au passage que dans ce cas une fonction $f : \mathbb{R}^k \to \mathbb{R}$ est $Kd^2/2$-convexe si et seulement si la fonction $x \mapsto f(x) + K\|x\|_2^2/2$ est convexe sur $\mathbb{R}^k$, ce qui se traduit, par la condition

$$f((1-t)x+ty) \leqslant (1-t)f(x)+tf(y)+\frac{K}{2}t(1-t)\|x-y\|_2^2, \qquad \forall x, y \in \mathbb{R}^k, \ \forall t \in [0, 1]. \tag{5.2.4}$$

  Dans le cas où f est de classe $\mathcal{C}^2$, ceci est encore équivalent à la condition

$$\mathrm{Hess}\, f \geqslant -K.$$

  C'est sous cette forme qu'est définie l'inégalité log-Sobolev restreinte dans [11].

– L'article [13] prouve l'équivalence entre $\mathbf{T}_2$ et l'inégalité (τ)-log-Sobolev (5.2.2) dans un cadre métrique. Notons que l'équivalence entre ces inégalités est vraie même si l'espace n'est pas géodésique.
– L'article [15] complète le travail précédent en montrant l'équivalence avec l'inégalité log-Sobolev restreinte. L'une des difficultés que nous avons rencontrées était d'identifier la bonne classe de fonctions pour définir l'inégalité restreinte. Dans [11], nous avions tenté de définir la classe de fonctions en généralisant la condition (5.2.4) au cas d'un espace géodésique. Ce point de vue semble moins bien adapté, en particulier en ce qui concerne l'extension à des fonctions de coûts $\theta(t) = t^p$, avec $p > 2$.

Commentons les principaux points de la preuve du théorème 5.2.1. Le passage de $\mathbf{T}_2$ à l'inégalité (τ)-log-Sobolev se fait grâce à un argument très simple que nous allons présenter rapidement. Soit $f \in \mathcal{C}_b(\mathcal{X})$; considérons la probabilité $\nu_f(dx) = \frac{e^{f(x)}}{Z_f}\,\mu(dx)$, où Z_f est une constante de renormalisation. En appliquant l'inégalité de Jensen, puis la formule de dualité de Kantorovich et enfin l'inégalité $\mathbf{T}_2$, on voit que

$$\begin{aligned} H(\nu_f|\mu) &= \int f\, d\nu_f - \log\left(\int e^f\, d\mu\right) \\ &\leqslant \int f\, d\nu_f - \int f\, d\mu \\ &\leqslant \int f - Q_{1/\lambda} f\, d\nu_f + \sup_{g \in \mathcal{C}_b(\mathcal{X})} \left\{\int Q_{1/\lambda} g\, d\nu_f - \int g\, d\mu\right\} \\ &= \int f - Q_{1/\lambda} f\, d\nu_f + \lambda C H(\nu_f|\mu). \end{aligned}$$

Par conséquent, si $\lambda < 1/C$, on a $H(\nu_f|\mu) \leqslant \frac{1}{1-\lambda C}\int f - Q_{1/\lambda} f\, d\nu_f$, ce qui revient à l'inégalité τ-log-Sobolev.

Le passage de l'inégalité log-Sobolev restreinte à l'inégalité $\mathbf{T}_2$ peut se faire en utilisant les mêmes techniques que dans le cas classique (utilisation d'Hamilton-Jacobi ou de la concentration). On peut, par exemple, se convaincre que l'utilisation de la technique de [BGL01] est possible, en observant que la classe des fonctions $K\frac{d^2}{2}$-convexe dans $\mathbb{R}^k$ est stable sous l'action des opérateurs d'inf-convolution Q_t : on peut montrer que si f est $\frac{1}{s}\frac{d^2}{2}$-convexe dans $\mathbb{R}^k$, alors $Q_t f$ est $\frac{1}{s-t}\frac{d^2}{2}$-convexe, pour tout $0 < t < s$.

Voyons à présent comment le théorème précédent peut être utilisé pour démontrer une propriété de stabilité de l'inégalité $\mathbf{T}_2$ par perturbation bornée. Ce résultat a été établi pour la première fois dans [11].

Théorème 5.2.5. *Si μ est une mesure de probabilité sur un espace métrique $(\mathcal{X}, d)$ vérifiant $\mathbf{T}_2(C)$ $C > 0$, alors pour toute fonction $\varphi : \mathcal{X} \to \mathbb{R}$ bornée, la mesure de probabilité $\tilde{\mu}(dx) = \frac{e^{\varphi}}{Z}\,\mu(dx)$ vérifie l'inégalité $\mathbf{T}_2$ avec la constante $\tilde{C} = e^2 e^{\mathrm{Osc}(\varphi)} C$, où $\mathrm{Osc}(\varphi) = \sup \varphi - \inf \varphi$.*

Ce type de stabilité par perturbation bornée est très classique pour les inégalités de Poincaré ou de Sobolev logarithmiques (et leurs variantes). Pour ces inégalités, la stabilité par perturbation bornée est obtenu par un argument très simple de Holley et Stroock [HS87] (voir aussi [ABC+00]) qu'on va rappeler ci-dessous dans le cas de l'inégalité de Poincaré : supposons que

μ vérifie l'inégalité de Poincaré suivante

$$\mathrm{Var}_\mu(f) \leqslant C \int |\nabla f|^2 \, d\mu, \qquad \forall f$$

Remarquons que

$$\mathrm{Var}_{\tilde{\mu}}(f) = \inf_{a \in \mathbb{R}} \int (f-a)^2 \, d\tilde{\mu} \leqslant \int \left(f - \int f \, d\mu \right)^2 \, d\tilde{\mu} \leqslant \frac{e^{\sup \varphi}}{Z} \mathrm{Var}_\mu(f)$$

Par ailleurs,

$$\int |\nabla f|^2 \, d\mu \leqslant Z e^{-\inf \varphi} \int |\nabla f|^2 \, d\tilde{\mu}.$$

On en déduit immédiatement que $\tilde{\mu}$ vérifie l'inégalité de Poincaré.

Cet argument ne fonctionne plus pour l'inégalité $\mathbf{T}_2$ et la question de la stabilité est restée ouverte plusieurs années. Maintenant qu'on sait que $\mathbf{T}_2$ est équivalente à **LSI** en restriction à la classe des fonctions d^2-convexes, il suffit d'appliquer l'argument de Holley-Stroock à l'inégalité log-Sobolev restreinte pour obtenir la stabilité de $\mathbf{T}_2$ (au facteur multiplicatif e^2 près).

Chapitre 6

Inégalités fonctionnelles en courbure minorée

Ce chapitre présente les résultats obtenus dans [11].

6.1 Inégalités fonctionnelles sous hypothèses de courbure.

Soit (M, g) une variété riemannienne connexe complète. Revenons sur le critère de Bakry-Emery [BE85] brièvement présenté dans le chapitre 2. Si $\mu(dx) = e^{-V(x)}\,dx$ est une mesure de probabilité définie par un potentiel $V : M \to \mathbb{R}$ de classe $\mathcal{C}^2$ vérifiant la condition $CD(K, \infty)$, avec $K > 0$,

$$\mathrm{Ric} + \mathrm{Hess}\, V \geqslant K, \tag{6.1.1}$$

alors la probabilité μ vérifie l'inégalité de Sobolev logarithmique **LSI**$(2/K)$. Cette conclusion peut être renforcée ; comme l'ont montré Bakry et Ledoux [BL96], l'hypothèse $CD(K, \infty)$, $K > 0$, permet de montrer que μ vérifie une inégalité isopérimétrique gaussienne. Rappelons que le profil isopérimétrique de μ est défini par

$$\mathrm{Is}_\mu(t) = \inf\{\mu^+(A);\ \mu(A) = t\}, \qquad \forall t \in [0, 1],$$

où pour tout ensemble $A \subset M$, la mesure de bord, également appelée contenu de Minkowski, de A est définie par

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

D'après [BL96], sous l'hypothèse $CD(K, \infty)$, $K > 0$, μ vérifie l'inégalité suivante

$$\sqrt{C/2}\, \mathrm{Is}_\mu(t) \geqslant \Phi' \circ \Phi^{-1}(t), \qquad \forall t \in (0, 1), \tag{6.1.2}$$

avec $C = \frac{2}{K}$, où $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-u^2/2}\,du$, $x \in \mathbb{R}$ est la fonction de répartition gaussienne. Si $\mu = \gamma^k$ est la gaussienne standard sur $M = \mathbb{R}^k$, il y a égalité dans (6.1.2), ce qui exprime le fait que les demi-espaces sont solutions du problème isopérimétrique gaussien : à un volume fixé, ce sont les ensembles ayant une mesure de surface minimale. Ce dernier résultat a été obtenu indépendamment par Borell [Bor75] et Sudakov-Tsirelson [SC74].

L'inégalité isopérimétrique gaussienne est plus forte que l'inégalité log-Sobolev ([Led00]) ; on a donc la chaîne d'implication suivante

$$(6.1.2) \Rightarrow \mathbf{LSI}(C) \Rightarrow \mathbf{T}_2(C) \Rightarrow \text{Concentration gaussienne } \mu(A_r) \geqslant 1 - e^{-(r-r_o)^2/C}.$$

Si μ vérifie $CD(K, \infty)$ avec $K > 0$, toutes ces inégalités ont donc lieu avec la constante $C = 2/K$. Qu'en est il lorsque $K \leqslant 0$?

Dans une série de papiers, E. Milman démontre qu'en courbure minorée on peut remonter de la concentration à l'isopérimétrie [Mil09c, Mil09b, Mil09a, Mil10]. Ses résultats sont valables pour tout type de concentration. Dans le cas gaussien, son théorème s'énonce ainsi.

Théorème 6.1.3. *Soit $\mu(dx) = e^{-V(x)}\,dx$ une probabilité sur M vérifiant la condition $CD(K, \infty)$, avec $K \leqslant 0$. S'il existe $M, a > 0$ tels que μ vérifie la propriété de concentration gaussienne suivante*

$$\mu(A_r) \geqslant 1 - Me^{-ar^2}, \qquad \forall r \geqslant 0, \qquad \forall A \subset M,\ \mu(A) \geqslant 1/2,$$

avec $|K|/(2a) < 1$, alors il existe une constante $C(a, M, K)$ ne dépendant que de a, M et K telle que μ vérifie l'inégalité isopérimétrique gaussienne (6.1.2) *avec la constante C.*

Ce résultat généralise notamment un résultat de Wang [Wan97] montrant que sous l'hypothèse $CD(K, \infty)$, avec $K < 0$, une condition d'intégrabilité exponentielle d'ordre 2 du type $I(a) := \int e^{ad^2(x_o,x)}\,\mu(dx) < +\infty$, avec $a > |K|/2$ entraîne $\mathbf{LSI}(C)$, avec une constante C dépendant de $I(a)$. Comme nous l'avons observé dans le chapitre 3, $I(a)$ dépend en général de la dimension de la variété, tandis que le résultat de Milman fournit une dépendance universelle minimale : la constante C ne dépend que des paramètres a, K et M et pas du tout de la variété considérée.

Le but de l'article [11] est de généraliser en partie le théorème précédent à un cadre métrique. La notion de courbure sous laquelle nous allons travailler est celle introduite par Lott-Villani [LV09] et Sturm [Stu06a, Stu06b] que nous allons maintenant brièvement présenter.

6.2 Courbure minorée au sens de Lott-Sturm-Villani

Rappelons qu'un espace métrique $(\mathcal{X}, d)$ est géodésique si pour tous $x_0, x_1 \in \mathcal{X}$, il existe une courbe $(x_t)_{t\in[0,1]}$ telle que $d(x_t, x_s) = |t - s| d(x_0, x_1)$ pour tout $s, t \in [0, 1]$. Un chemin $(x_t)_{t\in[0,1]}$ de ce type est appelé géodésique à vitesse constante. Si $(\mathcal{X}, d)$ est un espace géodésique, l'espace $\mathcal{P}_2(\mathcal{X})$ des probabilités ayant un moment d'ordre deux fini, muni de la distance W_2, est également géodésique. Les chemins géodésiques à vitesse constante dans $(\mathcal{P}_2(\mathcal{X}), W_2)$ s'obtiennent grâce à la proposition suivante (voir par exemple [Vil09]). On notera $\mathcal{G}$, l'ensemble des géodésiques à vitesse constante dans $\mathcal{X}$, et pour tout $t \in [0, 1]$, $e_t : \mathcal{G} \to \mathcal{X} : \gamma \mapsto \gamma_t$.

Proposition 6.2.1. *Soient $(\mathcal{X}, d)$ un espace géodésique et $\nu_0, \nu_1 \in \mathcal{P}_2(\mathcal{X})$. Un chemin $(\nu_s)_{s\in[0,1]}$ joignant $\nu_0, \nu_1 \in \mathcal{P}_2(\mathcal{X})$ est une géodésique à vitesse constante pour W_2 si et seulement s'il existe une mesure de probabilité Π sur $\mathcal{G}$ telle que $\Pi_\# e_t = \nu_t$, pour tout $t \in [0, 1]$ et $\pi = \Pi_\#(e_0, e_1)$ est un couplage optimal entre ν_0 et ν_1. Il existe toujours au moins une géodésique entre ν_0 et ν_1.*

Dans le cas où $\mathcal{X} = M$ est une variété riemanienne connexe et complète, si ν_0, ν_1 sont à support compact, et ν_0 est absolument continue par rapport à la mesure de Lebesgue sur M, il existe une unique application de transport T entre ν_0 et ν_1. Elle se présente sous la forme

$$T(x) = \exp_x(\nabla\varphi(x)), \qquad \forall x \in M,$$

où $\varphi : M \to \mathbb{R}$ est une fonction $\frac{d^2}{2}$-convexe. De plus, pour ν_0 presque tout $x \in M$, le chemin $T_t(x) = \exp_x(t\nabla\varphi(x))$ $t \in [0,1]$ est une géodésique (minimisante) à vitesse constante. De plus $\nu_t := T_{t\#}\nu_0$ est l'unique géodésique à vitesse constante pour W_2 (voir [Vil09]).

Le point de départ des travaux [LV09, Stu06a, Stu06b] est le résultat suivant.

Théorème 6.2.2. *Soit* $\mu(dx) = e^{-V(x)}\,dx$ *une probabilité sur* M*, variété riemannienne connexe et complète. Les propositions suivantes sont équivalentes.*

1. *La probabilité* μ *vérifie la condition* $\mathrm{CD}(K,\infty)$, $K \in \mathbb{R}$.
2. *Pour toute* $\nu_0, \nu_1 \in \mathcal{P}_2(M)$ *dont le support est inclus dans celui de* μ*, il existe une géodésique* ν_t *joignant* ν_0 *à* ν_1 *telle que*

$$H(\nu_t|\mu) \leqslant (1-t)H(\nu_0|\mu) + tH(\nu_1|\mu) - \frac{Kt(1-t)}{2}W_2^2(\nu_0,\nu_1), \qquad \forall t \in [0,1].$$

Ce théorème a une histoire assez longue. Il résulte des travaux réalisés par McCann (dans sa thèse de doctorat), Otto-Villani [OV00], Cordero-McCann-Scmuckenschläger [CEMS01, CEMS06], Sturm-Von Renesse [vRS05], Sturm [Stu06a, Stu06b], Lott-Villani [LV09]. Nous renvoyons au chapitre 17 de [Vil09] pour plus de détails.

Le théorème précédent ouvre la voie à une définition de la condition $\mathrm{CD}(K,\infty)$ sur un espace géodésique.

Définition 6.2.3. *On dira qu'une mesure de probabilité* μ *sur un espace géodésique* $(\mathcal{X}, d)$ *vérifie la condition* $\mathrm{CD}(K,\infty)$, $K \in \mathbb{R}$*, si pour toute* $\nu_0, \nu_1 \in \mathcal{P}_2(\mathcal{X})$ *dont le support est inclus dans celui de* μ*, il existe une géodésique* $(\nu_t)_{t\in[0,1]}$ *dans* $(\mathcal{P}_2(\mathcal{X}), W_2)$ *joignant* ν_0 *à* ν_1 *telle que*

$$H(\nu_t|\mu) \leqslant (1-t)H(\nu_0|\mu) + tH(\nu_1|\mu) - \frac{Kt(1-t)}{2}W_2^2(\nu_0,\nu_1), \qquad \forall t \in [0,1].$$

Cette définition est due à Lott-Villani [LV09] et Sturm [Stu06a]. Il est également possible de définir un critère de courbure dimension $\mathrm{CD}(K,N)$, avec $N < +\infty$, mais nous n'en parlerons pas ici. Notons que la définition de [LV09] est en réalité un peu plus restrictive, car elle impose que la condition de convexité soit vérifiée par toute une famille de fonctionnelles (incluant l'entropie relative).

La définition précédente a immédiatement des conséquences en terme d'inégalités fonctionnelles. Par exemple, lorsque $K > 0$, il est clair que la condition $\mathrm{CD}(K,\infty)$ entraîne

$$\mathcal{T}_2(\nu_0,\nu_1) \leqslant \frac{2}{K}\left[\frac{1}{t}H(\nu_0|\mu) + \frac{1}{1-t}H(\nu_1|\mu)\right], \qquad \forall \nu_0,\nu_1,$$

qui entraîne (et est en fait équivalente à) l'inégalité $\mathbf{T}_2(2/K)$.

De plus, on peut déduire de la condition $\mathrm{CD}(K,\infty)$ une inégalité fonctionnelle appelée HWI, car reliant l'entropie relative H, l'information de Fisher I, et la distance de Wasserstein W_2.

Théorème 6.2.4. *Soit* μ *une mesure de probabilité sur un espace géodésique* $(\mathcal{X}, d)$ *vérifiant la condition* $CD(K, \infty)$, $K \in \mathbb{R}$. *Pour toute probabilité* $\nu \in \mathcal{P}_2(\mathcal{X})$ *absolument continue par rapport à* μ, *avec* $f = \frac{d\nu}{d\mu} > 0$ *et Lipschitz, on a*

$$H(\nu|\mu) \leqslant \sqrt{I_\mu(\nu)} W_2(\nu, \mu) - \frac{K}{2} W_2^2(\nu, \mu),$$

avec $I_\mu(\nu) = \int \frac{|\nabla^- f|^2}{f} \, d\mu$.

Rappelons que la définition de $|\nabla^- f|$ est donnée par (5.1.1).

L'inégalité HWI permet de généraliser le critère de Bakry-Emery à un cadre métrique. En effet, si $K > 0$, en utilisant l'inégalité $ab - \frac{K}{2} b^2 \leqslant \frac{a^2}{2K}$, on déduit de l'inégalité HWI que μ vérifie **LSI**$(2/K)$.

6.3 De la concentration gaussienne à l'inégalité de Sobolev logarithmique

La résultat principal de l'article [11] est le suivant.

Théorème 6.3.1. *Soit* μ *une mesure de probabilité sur un espace géodésique* $(\mathcal{X}, d)$ *vérifiant la condition* $CD(K, \infty)$, *avec* $K \leqslant 0$. *Si* μ *vérifie l'inégalité de concentration gaussienne suivante*

$$\mu(A_r) \geqslant 1 - Me^{-ar^2}, \qquad \forall r \geqslant 0, \qquad \forall A \subset M, \ \mu(A) \geqslant 1/2, \tag{6.3.2}$$

avec

$$|K|/(2a) < \frac{\log(2)}{(2\sqrt{M} + \sqrt{\log(2)})^2},$$

alors μ *vérifie* **LSI**(C) *pour une constante* C *ne dépendant que de* K, a *et* M. *En particulier, si* $K = 0$, *on a*

$$\mathrm{Ent}_\mu(f^2) \leqslant \frac{\kappa M}{a} \int |\nabla^- f|^2 \, d\mu,$$

pour toute fonction f *Lipschitz, où* κ *est une constante numérique.*

Ce théorème généralise donc une partie du théorème 6.1.3 de Milman. On retrouve bien l'esprit du résultat de Milman : en courbure minorée, une concentration suffisamment forte permet de remonter dans la hiérarchie des inégalités fonctionnelles, avec un contrôle universel des constantes. En revanche, la condition portant sur le rapport $|K|/(2a)$ n'est pas la bonne.

La technique utilisée pour démontrer le théorème 6.3.1 est inspirée d'un argument développé par Otto et Villani dans [OV00] pour établir le résultat suivant.

Proposition 6.3.3. *Si* μ *vérifie* $CD(K, \infty)$, *avec* $K \leqslant 0$ *et l'inégalité* $\mathbf{T}_2(C)$ *avec une constante* C *telle que* $C|K|/2 < 1$, *alors* μ *vérifie* **LSI** *avec une constante dépendant uniquement de* K, C.

En effet, en injectant l'inégalité $W_2(\nu, \mu) \leqslant \sqrt{CH(\nu|\mu)}$ dans l'inégalité HWI, on obtient

$$H(\nu|\mu) \leqslant \sqrt{I_\mu(\nu)} \sqrt{CH(\nu|\mu)} + \frac{|K|C}{2} H(\nu|\mu).$$

Si $|K|C/2 < 1$, on obtient immédiatement **LSI** avec la constante $C/(1-|K|C/2)^2$.

Nous avons cherché à adapter cet argument en faisant appel à l'inégalité $\mathbf{T}_2$ non tendue (développée à cet effet), présentée dans le chapitre 5. Rappelons que le théorème 4.3.2 montre que si μ vérifie la propriété de concentration (6.3.6), alors elle vérifie l'inégalité $\mathbf{T}_2$ non tendue suivante :

$$\mathcal{T}_2(\nu,\mu) \leqslant \frac{u}{a} H(\nu|\mu) + \frac{4Mu}{a(u-1)}, \qquad \forall \nu \in \mathcal{P}(\mathcal{X}),\ \forall u > 1.$$

En injectant l'inégalité $\mathbf{T}_2$ non tendue dans l'inégalité HWI, nous obtenons une famille d'inégalités Sobolev logarithmiques non tendues :

$$\mathrm{Ent}_\mu(f^2) \leqslant \beta(r) \int |\nabla^- f|^2 \, d\mu + r \int f^2 \, d\mu, \qquad \forall f,$$

pour tout $r > r_o(a, K, M)$ et avec $\beta(r)$ dépendant uniquement de a, K, M et r.

Pour tendre une inégalité de ce type, on utilise généralement l'inégalité de Poincaré (qui permet d'éliminer le terme parasite). Dans notre contexte, nous ne savons pas que μ vérifie l'inégalité de Poincaré (même si ce sera bien le cas rétrospectivement puisque **LSI** entraîne Poincaré). Pour palier à cette absence d'information sur le trou spectral, nous faisons appel à un résultat assez étonnant de Wang [Wan04b] mettant en évidence un phénomène d'auto-tension lorsque le paramètre r est petit.

Proposition 6.3.4. *Si μ vérifie l'inégalité*

$$\mathrm{Ent}_\mu(f^2) \leqslant b_1 \int |\nabla^- f|^2 \, d\mu + b_2 \int f^2 \, d\mu, \qquad \forall f,$$

avec $b_2 < \log(2)$, alors μ vérifie **LSI** *avec une constante dépendant uniquement de b_1 et b_2.*

La condition sur le rapport $|K|/(2a)$ apparaissant dans le théorème 6.3.1 est là justement pour garantir que $r_o(a, K, M) < \log(2)$.

L'approche précédente peut être généralisée à d'autres inégalités fonctionnelles. Nous obtenons en particulier le résultat suivant concernant l'inégalité de Poincaré en courbure nulle.

Théorème 6.3.5. *Soit μ une mesure de probabilité sur un espace géodésique $(\mathcal{X}, d)$ vérifiant la condition $CD(0,\infty)$ (de Lott-Villani [LV09]). Si μ vérifie l'inégalité de concentration exponentielle suivante*

$$\mu(A_r) \geqslant 1 - Me^{-ar}, \qquad \forall r \geqslant 0, \qquad \forall A \subset M,\ \mu(A) \geqslant 1/2, \tag{6.3.6}$$

alors μ vérifie l'inégalité de Poincaré suivante

$$\mathrm{Var}_\mu(f) \leqslant \frac{\kappa(M)}{a^2} \int |\nabla^- f|^2 \, d\mu,$$

pour toute fonction f Lipschitz, où $\kappa(M)$ est une constante dépendant uniquement de M.

Signalons que dans un cadre riemannien, Milman [Mil09c, Mil10] a obtenu un résultat spectaculaire montrant que sous l'hypothèse $CD(0,\infty)$ (qui revient à la log-concavité de μ lorsque $\mathcal{X} = \mathbb{R}^k$) n'importe quelle propriété de concentration entraîne l'inégalité de Poincaré.

Chapitre 7

Inégalités fonctionnelles pour des mesures à queues lourdes

Dans ce dernier chapitre, nous présentons les articles [7] et [10] qui aborde la question assez vaste de déterminer des inégalités fonctionnelles adaptées aux mesures de probabilité ayant des queues de distributions décroissant lentement. Dans $\mathbb{R}^k$, les exemples typiques de telles distributions sont les mesures de type Cauchy

$$\mu(dx) = Z^{-1} \frac{1}{(1 + \|x\|_2^2)^{(k+\alpha)/2}}\, dx, \qquad \alpha > 0,$$

ou les mesures sous-exponentielles

$$\mu(dx) = Z^{-1} e^{-\|x\|_2^p}\, dx, \qquad p \in]0,1[.$$

Pour de telles mesures de probabilité, aucune des inégalités fonctionnelles rencontrées jusque là dans ce manuscrit (Log-Sobolev, transport, super Poincaré, ...) n'est vérifiée. En effet, l'inégalité de Poincaré, qui est la plus faible des inégalités classiques, entraîne de la concentration exponentielle, ce qui exclut les probabilités ci-dessus. Dans [7], nous affaiblissons la définition de diverses inégalités fonctionnelles, en considérant des versions pondérées de celles-ci, afin d'englober les probabilités à queues lourdes.

La motivation principale des articles [7, 10] est de chercher à estimer le profil isopérimétrique (pour la norme euclidienne) en dehors du cadre classique des probabilités fortement concentrées [BL96, Bob99, BCR07, Mil10].

7.1 Inégalités à poids

Rappelons que si $\mu \in \mathcal{P}(\mathbb{R}^k)$, le profil isopérimétrique de μ est défini par

$$\mathrm{Is}_\mu(t) = \inf\{\mu^+(A); A \subset \mathbb{R}^k, \mu(A) = t\}, \qquad t \in [0,1],$$

où $\mu^+(A) = \liminf_{r\to 0} \mu(A_r \setminus A)/r$, avec $A_r = \{x \in \mathbb{R}^k, \exists y \in A, \|x-y\|_2 \leqslant r\}$.

On dit que μ vérifie une inégalité isopérimétrique si, pour une certaine fonction F, on a

$$\mu^+(A) \geqslant F(\mu(A)), \qquad \forall A \subset \mathbb{R}^k,$$

ce qui se traduit de manière équivalente par l'inégalité $\mathrm{Is}_\mu(t) \geqslant F(t)$ $t \in [0,1]$.

Une inégalité isopérimétrique peut se traduire par une inégalité de Cheeger faible, comme le montre le résultat suivant de Bobkov [Bob07].

Proposition 7.1.1. *Soit $\mu \in \mathcal{P}(\mathbb{R}^k)$; les propositions suivantes sont équivalentes.*

1. *La probabilité μ vérifie l'inégalité isopérimétrique*

$$\mu^+(A) \geqslant F(\mu(A)), \qquad \forall A \subset \mathbb{R}^k,$$

avec $F : [0,1] \to \mathbb{R}^+$ symétrique par rapport à $1/2$.

2. *La probabilité μ vérifie l'inégalité de Cheeger faible : pour une certaine fonction $\beta :]0,1/2[\to \mathbb{R}^+$,*

$$\int |f - m(f)|\, d\mu \leqslant \beta(s) \int \|\nabla f\|_2 \, d\mu + s\mathrm{Osc}\,(f), \qquad \forall s \in]0,1/2[,$$

pour toute $f : \mathbb{R}^k \to \mathbb{R}$ Lipschitz, où $m(f)$ désigne la médiane de f.

De plus, les fonctions F et β sont reliées par les relations suivantes :

$$\beta(s) = \sup_{s \leqslant t \leqslant 1/2} \frac{t-s}{F(t)},\ s \in]0,1/2[\qquad\qquad F(t) = \sup_{0 \leqslant s \leqslant t} \frac{t-s}{\beta(s)},\ t \leqslant 1/2.$$

On retrouve en particulier le fait qu' une probabilité vérifie l'inégalité de Cheeger classique

$$\mu^+(A) \geqslant C \min(\mu(A); 1 - \mu(A)), \qquad \forall A \subset \mathbb{R}^k, \tag{7.1.2}$$

si et seulement si elle vérifie l'inégalité de Poincaré L_1

$$\int |f - m(f)|\, d\mu \leqslant \frac{1}{C} \int \|\nabla f\|_2 \, d\mu,$$

pour toute fonction $f : \mathbb{R}^k \to \mathbb{R}$ régulière.

Dans [7], nous développons, pour les mesures à queues lourdes, différents types d'inégalités de Poincaré ou de Cheeger à poids, plus faibles que leurs homologues sans poids. Ces variantes pondérées des inégalités classiques apparaissent également dans [BL09]. Pour éviter d'alourdir la présentation, nous ne mentionnerons dans la suite que l'inégalité de Cheeger avec poids, qui est en lien directe avec l'isopérimétrie.

Définition 7.1.3. *Nous dirons que $\mu \in \mathcal{P}(\mathbb{R}^k)$ vérifie l'inégalité de Cheeger pondérée à gauche par $\omega : \mathbb{R}^k \to \mathbb{R}^+$, avec la constante $C > 0$, si*

$$\inf_{a \in \mathbb{R}} \int |f - a| \omega \, d\mu \leqslant C \int \|\nabla f\|_2 \, d\mu, \tag{7.1.4}$$

pour toute fonction $f : \mathbb{R}^k \to \mathbb{R}$ régulière.

L'intérêt de ces inégalités de Cheeger pondérées à gauche est qu'elles entraînent des inégalités de Cheeger faibles, et conduisent donc, d'après la proposition 7.1.1, à une estimation du profil isopérimétrique de μ.

Lemme 7.1.5. *Si* μ *vérifie l'inégalité de Cheeger pondérée à gauche* (7.1.4), *alors elle vérifie l'inégalité de Cheeger faible suivante :*

$$\int |f - m(f)|\, d\mu \leqslant \frac{C}{H^{-1}(s)} \int \|\nabla f\|_2\, d\mu + s\mathrm{Osc}\,(f), \qquad \forall s \in]0, 1/2[,$$

où $H(u) = \mu(\omega \leqslant u)$ *et* $H^{-1}(s) = \inf\{u, H(u) \geqslant s\}$.

L'argument étant très simple, nous le rappelons ci-dessous. Supposons donc que μ vérifie (7.1.4). Prenons $f : \mathbb{R}^k \to \mathbb{R}$ une fonction bornée et régulière et considérons une médiane a de f pour la probabilité $\tilde{\mu} = \frac{1}{Z}\omega\, \mu$. On a alors $\inf_c \int |f - c|\omega d\mu = \int |f - a|\omega\, d\mu$ et $|f - a| \leqslant \mathrm{Osc}\,(f)$. On en déduit, pour tout $u > 0$,

$$\begin{aligned} \int |f - a|\, d\mu &\leqslant \frac{1}{u} \int_{\omega > u} |f - a|\omega\, d\mu + \int_{\omega \leqslant u} |f - a|\, d\mu \\ &\leqslant \frac{1}{u} \int |f - a|\omega\, d\mu + H(u)\mathrm{Osc}\,(f). \end{aligned}$$

D'où le résultat en appliquant l'inégalité de Cheeger pondérée.

Pour obtenir une inégalité de Cheeger à poids, nous mettons en oeuvre une méthode reposant sur les fonctions de Lyapunov (voir [BBCG08], [BCG08], [CGWW09] pour d'autres utilisations des fonctions de Lyapunov). Le résultat principal de [7] est le théorème suivant.

Théorème 7.1.6. *Soit* $\mu(dx) = e^{-V(x)}\, dx$ *une mesure de probabilité associée à un potentiel* $V : \mathbb{R}^k \to \mathbb{R}$ *de classe* $\mathcal{C}^1$. *Pour toute fonction* $f : \mathbb{R}^k \to \mathbb{R}$ *de classe* $\mathcal{C}^2$, *on note* $Lf = \Delta f - \nabla f \cdot \nabla V$.
On suppose qu'il existe une fonction $W : \mathbb{R}^k \to [1, +\infty[$ *de classe* $\mathcal{C}^2$, *une fonction* $\varphi : \mathbb{R}^+ \to \mathbb{R}$ *de classe* $\mathcal{C}^1$ *croissante et un ensemble* $K \subset \mathbb{R}^k$ *compact tels que pour un certain* $b \in \mathbb{R}$, *on ait*

$$LW \leqslant -\varphi(W) + b\mathbf{1}_K. \tag{7.1.7}$$

On suppose également que μ *vérifie une inégalité de Cheeger locale sur un ensemble* U *contenant* K, *c'est-à-dire qu'il existe une constante* $C_U > 0$ *telle que*

$$\int_U |f|\, d\mu \leqslant C_U \int_E \|\nabla f\|_2\, d\mu, \tag{7.1.8}$$

pour toute $f : \mathbb{R}^k \to \mathbb{R}$ *régulière et de médiane nulle sous la probabilité* $\mu_U(dx) = \frac{1}{\mu(U)}\mathbf{1}_U(x)\, dx$.
Sous ces hypothèses, μ *vérifie l'inégalité de Cheeger pondérée à droite suivante :*

$$\int |f - m(f)|\, d\mu \leqslant \max\left(\frac{bC_U}{\varphi(1)}\right) \int \left(1 + \frac{\|\nabla W\|_2}{\varphi(W)}\right) \|\nabla f\|_2\, d\mu,$$

pour toute fonction $f : \mathbb{R}^k \to \mathbb{R}$ *régulière.*
Si en plus $\mathrm{Hess}\, W \geqslant 0$ *en dehors de* K, *alors il existe une constante* $C > 0$, *telle que* μ *vérifie l'inégalité de Cheeger pondérée à gauche suivante :*

$$\inf_c \int |f - c| \frac{\varphi(W)}{\sqrt{1 + \|\nabla W\|_2^2}}\, d\mu \leqslant C \int \|\nabla f\|_2\, d\mu,$$

pour toute fonction $f : \mathbb{R}^k \to \mathbb{R}$ *régulière.*

Une fonction W vérifiant (7.1.7) est appelée *fonction de Lyapunov* pour l'opérateur L.

Pour illustrer la méthode de preuve, nous allons démontrer l'inégalité de Cheeger pondérée à droite. L'inégalité pondérée à gauche est un peu plus délicate à obtenir. Par hypothèse, nous avons

$$1 \leqslant -\frac{LW}{\varphi(W)} + \frac{b\mathbf{1}_K}{\varphi(W)} \leqslant -\frac{LW}{\varphi(W)} + \frac{b\mathbf{1}_K}{\varphi(1)}.$$

Par conséquent, en intégrant, puis en utilisant la formule d'intégration par partie classique

$$\int g(-Lh)\, d\mu = \int \nabla g \cdot \nabla h\, d\mu,$$

on obtient

$$\begin{aligned}
\int |f-a|\, d\mu &\leqslant -\int \frac{|f-a|}{\varphi(W)} LW\, d\mu + \frac{b}{\varphi(1)} \int_K |f-a|\, d\mu \\
&= \int \nabla \left(\frac{|f-a|}{\varphi(W)} \right) \cdot \nabla W\, d\mu + \frac{b}{\varphi(1)} \int_K |f-a|\, d\mu \\
&= \int \frac{\nabla |f-a| \cdot \nabla W}{\varphi(W)}\, d\mu - \int |f-a| \frac{\varphi'(W)}{\varphi^2(W)} \|\nabla W\|_2^2\, d\mu + \frac{b}{\varphi(1)} \int_K |f-a|\, d\mu \\
&\leqslant \int \|\nabla f\|_2 \frac{\|\nabla W\|_2}{\varphi(W)}\, d\mu + \frac{b}{\varphi(1)} \int_U |f-a|\, d\mu.
\end{aligned}$$

En prenant a égal à la médiane de f sous la probabilité μ_U de l'énoncé, l'inégalité de Cheeger locale permet de majorer le dernier terme par $\frac{bC_U}{\varphi(1)} \int \|\nabla f\|_2\, d\mu$, ce qui achève la preuve.

Le théorème précédent, nous permet d'estimer le profil isopérimétrique des distributions de types Cauchy ou sous-exponentiel.

Théorème 7.1.9.

1. *Soit* $\mu(dx) = V(x)^{-(n+\alpha)}\, dx$, *avec* $V : \mathbb{R}^k \to]0, +\infty[$ *convexe et* $\alpha > 0$. *Il existe une constante* $C > 0$ *telle que*
$$\mu^+(A) \geqslant C \min(\mu(A); 1-\mu(A))^{1+\frac{1}{\alpha}}, \qquad \forall A \subset \mathbb{R}^k.$$
2. *Soit* $\mu(dx) = \frac{1}{Z} e^{-V(x)^p}\, dx$, *avec* $V : \mathbb{R}^k \to]0, +\infty[$ *convexe et* $p \in]0, 1[$. *Il existe une constante* $C > 0$ *telle que*
$$\mu^+(A) \geqslant C \min(\mu(A); 1-\mu(A)) \log \left(\frac{1}{\min(\mu(A); 1-\mu(A))} \right)^{1-\frac{1}{p}}, \qquad \forall A \subset \mathbb{R}^k.$$

Ce résultat est surtout qualitatif, car dans l'approche par les fonctions de Lyapunov, les constantes sont assez mal contrôlées. Le premier point avait été obtenu par Bobkov dans [Bob07], avec un contrôle fin de la constante C. Formellement, en prenant $p = 1$ dans le deuxième résultat, on retrouve le fait bien connu qu'une mesure log-concave vérifie l'inégalité de Cheeger [KLS95, Bob99], conclusion déjà obtenue par les fonctions de Lyapunov dans [BBCG08].

7.2 Tensorisation de l'isopérimétrie

Un autre point abordé dans [7] et [10] est l'étude en fonction de n du profil isopérimétrique de μ^n, pour une mesure $\mu \in \mathcal{P}(\mathbb{R})$ à queues lourdes. Nous présentons ici le point de vue développé dans [10], qui donne une preuve plus satisfaisante de résultats obtenus dans [7].

Soit μ une mesure de probabilité sur $\mathbb{R}$; pour tout $n \in \mathbb{N}^*$, le profil isopérimétrique de μ^n sur $\mathbb{R}^n$ pour la métrique euclidienne standard est noté Is_{μ^n}. On notera $\mathrm{Is}_{\mu^\infty} = \inf_{n \geqslant 1} \mathrm{Is}_{\mu^n}$.

L'inégalité $\mathrm{Is}_\mu \geqslant \mathrm{Is}_{\mu^n} \geqslant \mathrm{Is}_{\mu^\infty}$ est toujours vraie. L'égalité $\mathrm{Is}_\mu = \mathrm{Is}_{\mu^\infty}$ caractérise les mesures gaussiennes [BH96]. Par ailleurs, une inégalité

$$\mathrm{Is}_\mu \geqslant \mathrm{Is}_{\mu^\infty} \geqslant c\mathrm{Is}_\mu, \tag{7.2.1}$$

pour un certain $c > 0$ impose à μ d'être comprise entre la mesure exponentielle et la mesure gaussienne (voir par exemple [BCR07]). L'inégalité (7.2.1) est par exemple vérifiée par des mesures de la forme $\mu(dx) = Z^{-1} e^{-V(x)}\, dx$ pour une fonction $V : \mathbb{R} \to \mathbb{R}^+$ convexe et telle que $\sqrt{V}$ est concave (voir [BCR06, BCR07]). Pour des mesures à queues lourdes, il faut donc s'attendre à ce que $\mathrm{Is}_{\mu^n} \downarrow 0$, lorsque $n \to \infty$.

Le résultat principal de [10] est une propriété de tensorisation des inégalités isopérimétriques et de Cheeger faibles introduites plus haut.

Théorème 7.2.2. *Soit $\mu \in \mathcal{P}(\mathbb{R}^k)$ une probabilité vérifiant l'inégalité isopérimétrique*

$$\mu^+(A) \geqslant J(\mu(A)), \qquad \forall A \subset \mathbb{R}^k, \tag{7.2.3}$$

pour une certaine fonction $J : [0,1] \to \mathbb{R}^+$ symétrique par rapport à $1/2$ et telle que $J(s)/s$ soit croissante sur $]0,1/2]$. Sous cette hypothèse, pour tout $n \in \mathbb{N}^*$, la probabilité μ^n vérifie l'inégalité isopérimétrique suivante

$$\mu^{n\,+}(A) \geqslant \frac{n}{2} J\left(\frac{1}{4\sqrt{6n}} \min(\mu^n(A); 1-\mu^n(A))\right), \qquad \forall A \subset \left(\mathbb{R}^k\right)^n.$$

Pour l'inégalité de Cheeger ($J(t) = C\min(t; 1-t)$), ce résultat a été démontré pour la première fois par Bobkov et Houdré dans [BH97a] (avec des constantes un peu meilleures). Dans ce cas, leur conclusion est que si μ vérifie l'inégalité de Cheeger (7.1.2) avec une constante C, alors pour tout n la probabilité μ^n vérifie l'inégalité de Cheeger avec la constante indépendante de la dimension $C/(2\sqrt{6})$. La preuve de [BH97a] a été récemment simplifiée par Bobkov dans [Bob09]. C'est cette preuve que nous avons adaptée pour établir le théorème 7.2.2.

Lorsque $k = 1$, et $\mu(dx) = e^{-V(x)}\, dx \in \mathcal{P}(\mathbb{R})$, où V est une fonction paire, la meilleure fonction J dans l'inégalité (7.2.3) est donnée par (voir par exemple [BH97b])

$$J_\mu(t) = \min(I_\mu(t); 2I_\mu(\min(t; 1-t)/2)), \quad t \in (0,1),$$

où $I_\mu(t) = F'_\mu \circ F_\mu^{-1}(t)$, $t \in (0,1)$ avec $F_\mu(x) = \mu]-\infty, x]$, $x \in \mathbb{R}$. Si V est concave, on peut montrer que $J_\mu(s)/s$ est croissante sur $]0,1/2[$ et au contraire, si V est convexe, alors $J_\mu(s)/s$ est décroissante. Le théorème 7.2.2 est donc bien adapté aux mesures à queues lourdes.

Appliqué par exemple aux mesures de Cauchy et aux mesures sous-exponentielles, le théorème 7.2.2 entraîne le résultat suivant (voir [7]).

Proposition 7.2.4.

1. *Soit $\mu(dx) = \frac{\alpha}{2(1+|x|)^{1+\alpha}}\, dx$, avec $\alpha > 0$. Il existe une constante $c_\alpha > 0$ ne dépendant que de α telle que*

$$I_{\mu^n}(t) \geqslant c_\alpha \frac{t^{1+\frac{1}{\alpha}}}{n^{\frac{1}{\alpha}}}, \qquad \forall t \in]0,1/2[$$

2. *Soit* $\mu(dx) = \frac{1}{Z} e^{-|x|^p} dx$, *avec* $p \in]0, 1[$. *Il existe une constante* $c_p > 0$ *ne dépendant que de* p *telle que*

$$I_{\mu^n}(t) \geqslant c_p t \left(\log \left(\frac{n}{t} \right) \right)^{1-\frac{1}{p}}, \qquad \forall t \in]0, 1/2[$$

Dans les deux cas, on peut montrer que la dépendance en n est du bon ordre [7].

Bibliographie

[ABC+00] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.

[AFHS97] M. Avellaneda, C. Friedman, R. Holmes, and D. Samperi. Calibrating volatility surfaces via relative-entropy minimization. *Applied Mathematical Finance*, 4(1) :37–64, 1997.

[AGS11] L. Ambrosio, N. Gigli, and G. Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *ArXiv :1106.2090v3*, June 2011.

[BBCG08] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Commun. Probab.*, 13 :60–66, 2008.

[BCG08] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes : Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3) :727–759, 2008.

[BCR06] F. Barthe, P. Cattiaux, and C. Roberto. Interpolated inequalities between exponential and Gaussian, Orlicz hypercontractivity and isoperimetry. *Rev. Mat. Iberoam.*, 22(3) :993–1067, 2006.

[BCR07] F. Barthe, P. Cattiaux, and C. Roberto. Isoperimetry between exponential and Gaussian. *Electron. J. Probab.*, 12 :no. 44, 1212–1237 (electronic), 2007.

[BE85] D. Bakry and M. Emery. Diffusions hypercontractives. In *Séminaire de probabilités*, volume 1123 of *Lecture Notes in Mathematics*, pages 177–206. Springer-Verlag, 1985.

[Bec89] W. Beckner. A generalized Poincaré inequality for Gaussian measures. *Proc. Amer. Math. Soc.*, 105 :397–400, 1989.

[BG99] S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1) :1–28, 1999.

[BGL01] S. G. Bobkov, I. Gentil, and M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pures Appl. (9)*, 80(7) :669–696, 2001.

[BGV07] F. Bolley, A. Guillin, and C. Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields*, 137(3-4) :541–593, 2007.

[BH96] S. G. Bobkov and C. Houdré. Characterization of Gaussian measures in terms of the isoperimetric property of half-spaces. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 228(Veroyatn. i Stat. 1) :31–38, 356, 1996.

[BH97a] S. Bobkov and C. Houdré. Isoperimetric constants for product probability measures. *Ann. Probab.*, 25(1) :184–205, 1997.

[BH97b] S. G. Bobkov and C. Houdré. Some connections between isoperimetric and Sobolev-type inequalities. *Mem. Amer. Math. Soc.*, 129(616) :viii+111, 1997.

[BH00] S. G. Bobkov and C. Houdré. Weak dimension-free concentration of measure. *Bernoulli*, 6(4) :621–632, 2000.

[BK00] V. V. Buldygin and Yu.V. Kozachenko. *Metric characterization of random variables and random processes*. American Mathematical Society, 2000.

[BL96] D. Bakry and M. Ledoux. Lévy-Gromov's isoperimetric inequality for an infinite-dimensional diffusion generator. *Invent. Math.*, 123(2) :259–281, 1996.

[BL97] S. G. Bobkov and M. Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107(3) :383–400, 1997.

[BL09] S. G. Bobkov and M. Ledoux. Weighted Poincaré-type inequalities for Cauchy and other convex measures. *Ann. Probab.*, 37(2) :403–427, 2009.

[Bob94] S. G. Bobkov. Isoperimetric inequalities for distributions of exponential type. *Ann. Probab.*, 22(2) :978–994, 1994.

[Bob99] S. G. Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Probab.*, 27(4) :1903–1921, 1999.

[Bob07] S. G. Bobkov. Large deviations and isoperimetry over convex probability measures. *Electr. J. Prob.*, 12 :1072–1100, 2007.

[Bob09] S. G. Bobkov. On the isoperimetric constants for product measures. *J. Math. Sci. (N. Y.)*, 159(1) :47–53, 2009. Problems in mathematical analysis. No. 40.

[Bor75] C. Borell. The Brunn-Minkowski inequality in Gauss space. *Invent. Math.*, 30(2) :207–216, 1975.

[BR03] F. Barthe and C. Roberto. Sobolev inequalities for probability measures on the real line. *Studia Math.*, 159(3) :481–497, 2003. Dedicated to Professor Aleksander Pełczyński on the occasion of his 70th birthday (Polish).

[BR08] F. Barthe and C. Roberto. Modified logarithmic Sobolev inequalities on $\mathbb{R}$. *Potential Anal.*, 29(2) :167–193, 2008.

[BV05] F. Bolley and C. Villani. Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities. *Ann. Fac. Sci. Toulouse Math. (6)*, 14(3) :331–352, 2005.

[BZ05] S. Bobkov and B. Zegarlinski. Entropy bounds and isoperimetry. *Mem. Amer. Math. Soc.*, 176(829) :x+69, 2005.

[CEMS01] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.*, 146(2) :219–257, 2001.

[CEMS06] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger. Prékopa-Leindler type inequalities on Riemannian manifolds, Jacobi fields, and optimal transport. *Ann. Fac. Sci. Toulouse Math. (6)*, 15(4) :613–635, 2006.

[CG06] P. Cattiaux and A. Guillin. On quadratic transportation cost inequalities. *J. Math. Pures Appl. (9)*, 86(4) :341–361, 2006.

[CGW10] P. Cattiaux, A. Guillin, and L. M. Wu. A note on Talagrand's transportation inequality and logarithmic Sobolev inequality. *Probab. Theory Related Fields*, 148(1-2) :285–304, 2010.

[CGWW09] P. Cattiaux, A. Guillin, F-Y. Wang, and L. Wu. Lyapunov conditions for super Poincaré inequalities. *J. Funct. Anal.*, 256(6) :1821–1841, 2009.

[Che99] J. Cheeger. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9(3) :428–517, 1999.

[CL94] P. Cattiaux and C. Léonard. Minimization of the Kullback information of diffusion processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 30(1) :83–132, 1994.

[CL95] P. Cattiaux and C. Léonard. Large deviations and Nelson processes. *Forum Math.*, 7(1) :95–115, 1995.

[Cra38] H. Cramér. Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Scientifiques et Industrielles*, 736 :5–23, 1938.

[Csi67] I. Csiszár. Information-type measures of difference of probability distributions and indirect observations. *Stud. Sci. Math. Hung.*, 2 :299–318, 1967.

[Csi84] I. Csiszár. Sanov property, generalized I-projection and a conditional limit theorem. *Ann. Probab.*, 12(3) :768–793, 1984.

[CSS76] S. Cambanis, G. Simons, and W. Stout. Inequalities for $\mathtt{Ek}(X, Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 36(4) :285–294, 1976.

[Dal56] G. Dall'Aglio. Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Ann. Scuoloa Norm. Sup. Pisa (3)*, 10 :35–74, 1956.

[DGW04] H. Djellout, A. Guillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.*, 32(3B) :2702–2732, 2004.

[DZ98] A. Dembo and O. Zeitouni. *Large deviations techniques and applications. Second edition*. Applications of Mathematics 38. Springer Verlag, 1998.

[Eva10] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[Fré60] M. Fréchet. Sur les tableaux dont les marges et des bornes sont données. *Rev. Inst. Internat. Statist.*, 28 :10–32, 1960.

[GG97] F. Gamboa and E. Gassiat. Bayesian methods and maximum entropy for ill-posed inverse problems. *Ann. Statist.*, 25(1) :328–350, 1997.

[GGM05] I. Gentil, A. Guillin, and L. Miclo. Modified logarithmic Sobolev inequalities and transportation inequalities. *Probab. Theory Related Fields*, 133(3) :409–436, 2005.

[GL11] N. Gigli and M. Ledoux. From log sobolev to talagrand : a quick proof. Submitted, 2011.

[GM83] M. Gromov and V. D. Milman. A topological application of the isoperimetric inequality. *Amer. J. Math.*, 105(4) :843–854, 1983.

[Gro75] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4) :1061–1083, 1975.

[Hoe40] W. Hoeffding. Maßstabinvariante korrelationstheorie. *Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin*, 5 :181–233, 1940.

[Hoe63] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58 :13–30, 1963.

[HS87] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Statist. Phys.*, 46(5-6) :1159–1194, 1987.

[KLS95] R. Kannan, L. Lovász, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.*, 13(3-4) :541–559, 1995.

[KO85] Yu. V. Kozachenko and E. I. Ostrovskiĭ. Banach spaces of random variables of sub-Gaussian type. *Teor. Veroyatnost. i Mat. Statist.*, (32) :42–53, 134, 1985.

[Kul67] S. Kullback. A lower bound for discrimination information in terms of variation. *IEEE Trans. Inform. Theory*, 4 :126–127, 1967.

[KZ95] S. R. Kulkarni and O. Zeitouni. A general classification rule for probability measures. *Ann. Statist.*, 23(4) :1393–1407, 1995.

[Led96] M. Ledoux. On Talagrand's deviation inequalities for product measures. *ESAIM Probab. Statist.*, 1 :63–87, 1996.

[Led00] M. Ledoux. The geometry of Markov diffusion generators. *Ann. Fac. Sci. Toulouse Math. (6)*, 9(2) :305–366, 2000. Probability theory.

[Led01] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[LO00] R. Latała and K. Oleszkiewicz. Between Sobolev and Poincaré. In *Geometric aspects of functional analysis*, volume 1745 of *Lecture Notes in Math.*, pages 147–168. Springer, Berlin, 2000.

[LV07] J. Lott and C. Villani. Hamilton-Jacobi semigroup on length spaces and applications. *J. Math. Pures Appl. (9)*, 88(3) :219–229, 2007.

[LV09] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3) :903–991, 2009.

[LW08] R. Latała and J. O. Wojtaszczyk. On the infimum convolution inequality. *Studia Math.*, 189(2) :147–187, 2008.

[Mar86] K. Marton. A simple proof of the blowing-up lemma. *IEEE Trans. Inform. Theory*, 32(3) :445–446, 1986.

[Mar96a] K. Marton. Bounding $\overline{\mathrm{d}}$-distance by informational divergence : a method to prove measure concentration. *Ann. Probab.*, 24(2) :857–866, 1996.

[Mar96b] K. Marton. A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.*, 6(3) :556–571, 1996.

[Mar98] K. Marton. Measure concentration for a class of random processes. *Probab. Theory Related Fields*, 110(3) :427–439, 1998.

[Mar04] K. Marton. Measure concentration for Euclidean distance in the case of dependent random variables. *Ann. Probab.*, 32(3B) :2526–2544, 2004.

[Mas07] P. Massart. *Concentration inequalities and model selection*, volume 1896 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007. Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6–23, 2003, With a foreword by Jean Picard.

[Mau91] B. Maurey. Some deviation inequalities. *Geom. Funct. Anal.*, 1(2) :188–197, 1991.

[Mil09a] E. Milman. Concentration and isoperimetry are equivalent assuming curvature lower bound. *C. R. Math. Acad. Sci. Paris*, 347(1-2) :73–76, 2009.

[Mil09b] E. Milman. On the role of convexity in functional and isoperimetric inequalities. *Proc. Lond. Math. Soc. (3)*, 99(1) :32–66, 2009.

[Mil09c] E. Milman. On the role of convexity in isoperimetry, spectral gap and concentration. *Invent. Math.*, 177(1) :1–43, 2009.

[Mil10] E. Milman. Isoperimetric and concentration inequalities : equivalence under curvature lower bound. *Duke Math. J.*, 154(2) :207–239, 2010.

[Muc72] B. Muckenhoupt. Hardy's inequality with weights. *Studia Math.*, 44 :31–38, 1972. Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, I.

[Naj02] J. Najim. A Cramér type theorem for weighted random variables. *Electron. J. Probab.*, 7 :no. 4, 32 pp. (electronic), 2002.

[OV00] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2) :361–400, 2000.

[Pin64] M.S. Pinsker. *Information and information stability of random variables and processes*. Holden-Day, San Francisco, 1964.

[RR91] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces*, volume 146 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 1991.

[RR98] S. T. Rachev and L. Rüschendorf. *Mass transportation problems. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, 1998. Theory.

[San61] I. N. Sanov. On the probability of large deviations of random variables. In *Select. Transl. Math. Statist. and Probability, Vol. 1*, pages 213–244. Inst. Math. Statist. and Amer. Math. Soc., Providence, R.I., 1961.

[SC74] V. N. Sudakov and B. S. Cirel'son. Extremal properties of half-spaces for spherically invariant measures. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 41 :14–24, 165, 1974. Problems in the theory of probability distributions, II.

[Stu06a] K.T. Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1) :65–131, 2006.

[Stu06b] K.T. Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1) :133–177, 2006.

[SZ91] D. W. Stroock and O. Zeitouni. Microcanonical distributions, Gibbs states, and the equivalence of ensembles. In *Random walks, Brownian motion, and interacting particle systems*, volume 28 of *Progr. Probab.*, pages 399–424. Birkhäuser Boston, Boston, MA, 1991.

[Tal94] M. Talagrand. The supremum of some canonical processes. *Amer. J. Math.*, 116(2) :283–325, 1994.

[Tal95] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'I.H.E.S.*, 81 :73–203, 1995.

[Tal96] M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3) :587–600, 1996.

[Vil03] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Vil09] C. Villani. *Optimal transport : Old and New*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009.

[vRS05] M-K. von Renesse and K-T. Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7) :923–940, 2005.

[Wan97] F-Y. Wang. Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. *Probab. Theory Related Fields*, 109(3) :417–424, 1997.

[Wan00] F-Y. Wang. Functional inequalities for empty essential spectrum. *J. Funct. Anal.*, 170(1) :219–245, 2000.

[Wan04a] F-Y. Wang. Probability distance inequalities on Riemannian manifolds and path spaces. *J. Funct. Anal.*, 206(1) :167–190, 2004.

[Wan04b] F-Y. Wang. Spectral gap for hyperbounded operators. *Proc. Amer. Math. Soc.*, 132(9) :2629–2638, 2004.