Inégalité de Prékopa–Leindler sur la sphère

Dario CORDERO-ERAUSQUIN

Équipe d'analyse et de mathématiques appliquées, ESA 8050, Université de Marne-la-Vallée, 5, boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée cedex 2, France
Courriel : cordero@math.univ-mlv.fr

(Reçu le 17 septembre 1999, accepté le 27 septembre 1999)

Résumé. En utilisant le transport de McCann [6] on obtient une version sphérique de l'inégalité de Prékopa–Leindler. © 1999 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Prékopa–Leindler inequality on the sphere

Abstract. We derive from McCann's transport [6] a spherical version of the Prékopa–Leindler inequality. © 1999 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

L'inégalité de Prékopa–Leindler ([8], [4]) dans $\mathbb{R}^n$ s'énonce de la manière suivante : soient $f, g, h : \mathbb{R}^n \rightarrow \mathbb{R}_+$ trois fonctions positives intégrables et $\alpha \in [0, 1]$ tel que, pour tous $x, y \in \mathbb{R}^n$, on ait $h((1 - \alpha)x + \alpha y) \geq f^{1-\alpha}(x)g^\alpha(y)$, alors

$$\int_{\mathbb{R}^n} h(x) dx \geq \left(\int_{\mathbb{R}^n} f(x) dx \right)^{1-\alpha} \left(\int_{\mathbb{R}^n} g(x) dx \right)^\alpha.$$

Cette inégalité a de nombreuses applications en convexité et en probabilité. Appliquée à des fonctions indicatrices d'ensembles, elle donne l'inégalité de Brunn–Minkowski. Elle permet aussi d'obtenir certaines inégalités de concentration, de Poincaré et de log-Sobolev (voir par exemple [5] et [1]). Pour démontrer simplement cette inégalité, on utilise le transport de Brenier–McCann dans $\mathbb{R}^n$ (voir [7]). Récemment McCann [6] a étendu les résultats du transport optimal de mesures au cas des variétés. Nous allons voir que cela permet de trouver une version sphérique de l'inégalité de Prékopa–Leindler.

La sphère unité $\mathbb{S}^n$ de l'espace euclidien $(\mathbb{R}^{n+1}, \langle \cdot, \cdot \rangle, |\cdot|)$ sera munie de sa structure riemannienne induite. La distance géodésique sera notée d . On rappelle que si $x, y \in \mathbb{S}^n$, $d(x, y) = \arccos(\langle x, y \rangle) \in [0, \pi]$ est l'angle entre x et y . On notera σ la mesure riemannienne sur $\mathbb{S}^n$ (qui est aussi la mesure de surface usuelle sur $\mathbb{S}^n$). Pour $\alpha \in [0, 1]$ et $x, y \in \mathbb{S}^n$ non antipodaux ($y \neq -x$) on notera $m_\alpha(x, y)$ le barycentre de x et de y avec poids respectifs $(1 - \alpha)$ et α . En d'autres termes, $m_\alpha(x, y)$ est l'unique point z de la sphère qui vérifie :

$$d(x, z) = \alpha d(x, y) \quad \text{et} \quad d(z, y) = (1 - \alpha) d(x, y). \quad (1)$$

Note présentée par Gilles PISIER.

D. Cordero-Erausquin

Si x et y sont antipodaux, on écrira $v = m_\alpha(x, y)$ pour dire que v est l'un des z qui vérifie (1). Notons $S : [0, \pi] \rightarrow \mathbb{R}$ la fonction $S(\theta) = \frac{\sin \theta}{\theta}$. La fonction S est log-concave positive avec $S(0) = 1$ et $S(\pi) = 0$. Pour $\alpha \in [0, 1]$ et $\theta \in [0, \pi]$, on définit

$$L_\alpha(\theta) = \frac{S(\theta)}{S^\alpha(\alpha\theta) S^{1-\alpha}((1-\alpha)\theta)}.$$

Comme S est log-concave on a : $L_\alpha \leq 1$. L'inégalité de Prékopa–Leindler sur la sphère s'énonce de la manière suivante :

THÉORÈME 1. – Soient $\alpha \in [0, 1]$ et $f, g, h : \mathbb{S}^n \rightarrow \mathbb{R}_+$ trois fonctions intégrables. Supposons que pour tous $x, y \in \mathbb{S}^n$ et $z = m_\alpha(x, y)$,

$$h(z) \geq (L_\alpha(d(x, y)))^{n-1} f(x)^{1-\alpha} g(y)^\alpha. \quad (2)$$

Alors on a :

$$\int_{\mathbb{S}^n} h(u) d\sigma(u) \geq \left(\int_{\mathbb{S}^n} f(u) d\sigma(u) \right)^{1-\alpha} \left(\int_{\mathbb{S}^n} g(u) d\sigma(u) \right)^\alpha.$$

Remarquons que si x et y sont antipodaux, $L_\alpha(d(x, y)) = 0$. Par conséquent, la condition (2) ne doit être vérifiée que pour des points non antipodaux. D'autre part, comme $L_\alpha \leq 1$, la condition (2) est plus faible que la condition de Prékopa–Leindler dans $\mathbb{R}^n$.

Dans la suite, chaque fois que nous aurons du calcul différentiel à faire sur une fonction définie sur la sphère, nous considérerons le prolongement 0-homogène à $\mathbb{R}^{n+1} \setminus \{0\}$ de cette fonction et nous calculerons les dérivées dans $\mathbb{R}^{n+1}$. Évidemment, le gradient sphérique coïncide sur la sphère avec le gradient du prolongement 0-homogène.

Rappelons le résultat de McCann [6]. Si μ et ν sont deux mesures sur $\mathbb{S}^n$ de même masse totale avec $\mu \ll \sigma$, alors il existe une fonction $\varphi : \mathbb{S}^n \rightarrow \mathbb{R}$ c -concave telle que l'application $t(x) = \exp_x(-\nabla\varphi(x))$ transporte μ sur ν . Si on note $c(x, y) = \frac{1}{2} d^2(x, y)$, on dit que φ est c -concave si $\varphi^{cc} = \varphi$, où $\varphi^c(x) = \inf_{y \in \mathbb{S}^n} (c(x, y) - \varphi(y))$. Signalons que l'application t précédente réalise le transport optimal (pour le coût quadratique) de μ sur ν . Nous allons utiliser le résultat de McCann dans le cas où les mesures μ et ν sont à densités continues strictement positives. Pour ne pas entrer dans des détails trop techniques nous supposons qu'alors la fonction φ est $\mathcal{C}^2$ (ce qui implique en particulier que $|\nabla\varphi| \in [0, \pi[$). On peut se passer de cette hypothèse en procédant comme dans [7]. En effet, en suivant [2], on remarque que, lue dans une carte, la fonction φ est semi-concave (ce qui implique qu'elle admet des dérivées secondes presque partout au sens d'Aleksandrov). Cette régularité est suffisante pour faire le changement de variable dont nous avons besoin (voir [7]).

Démonstration du théorème 1. – On se ramène au cas où les fonctions f et g sont continues et strictement positives. Notons $\sigma(f) = \int_{\mathbb{S}^n} f d\sigma$ et $\sigma(g) = \int_{\mathbb{S}^n} g d\sigma$, et introduisons les mesures suivantes sur $\mathbb{S}^n$: $d\mu(x) = \frac{\sigma(g)}{\sigma(f)} f(x) d\sigma(x)$ et $d\nu(x) = g(x) d\sigma(x)$. Soit donc φ (que l'on a supposé $\mathcal{C}^2$) la fonction c -concave telle que $t(x) = \exp_x(-\nabla\varphi(x))$ transporte μ sur ν . On notera $\phi = -\varphi$ de sorte que $t(x) = \exp_x(\nabla\phi(x))$. On a donc pour tout $x \in \mathbb{S}^n$:

$$g(t(x)) J(x) = \frac{\sigma(g)}{\sigma(f)} f(x), \quad (3)$$

où $J(x)$ est le déterminant jacobien de t en x . Si on note $\theta(x) = d(x, t(x)) = |\nabla\phi(x)| \in [0, \pi[$ et $u(x) = \frac{\nabla\phi(x)}{\theta(x)}$, on a : $t(x) = \cos(\theta)x + \sin(\theta)u$ (où l'on a omis d'écrire toutes les dépendances en x). Introduisons, pour $\alpha \in [0, 1]$, $t_\alpha : \mathbb{S}^n \rightarrow \mathbb{S}^n$ défini par :

$$t_\alpha(x) = \exp_x(\alpha \nabla\phi(x)) = \cos(\alpha\theta(x))x + \sin(\alpha\theta(x))u(x), \quad \forall x \in \mathbb{S}^n. \quad (4)$$

On notera J_α le déterminant jacobien de t_α . On vérifie facilement que t_α est injective. D'autre part, $t_\alpha(x) = m_\alpha(x, t(x))$, $t_0 = \text{id}_{\mathbb{S}^n}$, $t_1 = t$ (et donc $J_0 = 1$, $J_1 = J$). On a donc, en faisant le changement de variables $z = t_\alpha(x)$ et en utilisant (2) puis (3) :

$$\begin{aligned} \int_{\mathbb{S}^n} h(z) d\sigma(z) &= \int_{\mathbb{S}^n} h(t_\alpha(x)) J_\alpha(x) d\sigma(x) \\ &\geq \int_{\mathbb{S}^n} f(x)^{1-\alpha} g(t(x))^\alpha (L_\alpha(\theta(x)))^{n-1} J_\alpha(x) d\sigma(x) \\ &\geq \left(\frac{\sigma(g)}{\sigma(f)} \right)^\alpha \int_{\mathbb{S}^n} f(x) (L_\alpha(\theta(x)))^{n-1} \frac{J_\alpha(x)}{J(x)^\alpha} d\sigma(x). \end{aligned}$$

Il suffit donc de montrer que, pour tout $x \in \mathbb{S}^n$, on a : $(L_\alpha(\theta(x)))^{n-1} J_\alpha(x) \geq J(x)^\alpha$. Pour cela, calculons J_α . Si $x \in \mathbb{S}^n$ on notera $x^\perp$ l'orthogonal de x et P_x la projection orthogonale sur $x^\perp$. Les déterminants et les volumes en dimension n sont normalisés par la structure euclidienne de $\mathbb{R}^{n+1}$. Le déterminant jacobien de t_α en x est $J_\alpha(x) = \det(P_{t_\alpha(x)} dt_\alpha(x)|_{x^\perp})$, où $dt_\alpha|_{x^\perp}$ est la restriction de la dérivée de t_α à $x^\perp$. Soit $\mathcal{B}$ une base orthogonale de $x^\perp$ commençant par le vecteur $u(x)$ et soit $\mathcal{B}' = \{x\} \cup \mathcal{B}$. Par homogénéité on a : $\nabla\phi(x) + (\text{Hess}\phi)_x(x) = 0$, où $(\text{Hess}\phi)_x$ est la Hessienne de ϕ au point x . Dans la base $\mathcal{B}'$ la matrice de $(\text{Hess}\phi)_x$ est donc :

$$\begin{pmatrix} 0 & -\theta(x) & 0 \\ -\theta(x) & a & {}^t b \\ 0 & b & D \end{pmatrix}$$

avec $a \in \mathbb{R}$, $b \in \mathbb{R}^{n-1}$ et D une matrice symétrique $(n-1) \times (n-1)$ (que l'on peut supposer diagonale). En utilisant (4) on a, pour $x \in \mathbb{S}^n$ et $v \in x^\perp$:

$$\begin{aligned} dt_\alpha(x)(v) &= -\alpha \sin(\alpha\theta) \langle \nabla\theta, v \rangle x + \cos(\alpha\theta)v \\ &\quad + \left(\alpha \cos(\alpha\theta) - \frac{\sin(\alpha\theta)}{\theta} \right) \langle \nabla\theta, v \rangle u + \frac{\sin(\alpha\theta)}{\theta} (\text{Hess}\phi)_x(v). \end{aligned}$$

Remarquons que pour $v \in x^\perp$: $\langle \nabla\theta(x), v \rangle = \langle (\text{Hess}\phi)_x(u), v \rangle$. Pour calculer J_α on introduit la rotation R dans le plan $x, u(x)$ qui transforme $t_\alpha(x)$ en x (c'est-à-dire la translation parallèle de $t(x)$ vers x). Évidemment, $J_\alpha(x) = \det(RP_{t_\alpha(x)} dt_\alpha(x)|_{x^\perp})$. On a : $RP_{t_\alpha(x)}(x) = -\sin(\alpha\theta(x))u$ et $RP_{t_\alpha(x)}(u) = \cos(\alpha\theta(x))u$. La matrice de $RP_{t_\alpha(x)} dt_\alpha(x)|_{x^\perp}$ dans la base $\mathcal{B}$ est donc :

$$\begin{pmatrix} 1 + \alpha a & \alpha {}^t b \\ \frac{\sin(\alpha\theta)}{\theta} b & \cos(\alpha\theta)I + \frac{\sin(\alpha\theta)}{\theta} D \end{pmatrix}.$$

Mettons en facteur $\frac{\sin(\alpha\theta)}{\alpha\theta}$ dans les $(n-1)$ dernières lignes. On obtient

$$J_\alpha = \det(RP_{t_\alpha(x)} dt_\alpha(x)|_{x^\perp}) = \left(\frac{\sin(\alpha\theta)}{\alpha\theta} \right)^{n-1} \det H_\alpha = (S(\alpha\theta))^{n-1} \det H_\alpha, \quad (5)$$

D. Cordero-Erausquin

où H_α est la matrice symétrique suivante :

$$H_\alpha = \begin{pmatrix} 1 + \alpha a & \alpha^t b \\ \alpha b & T(\alpha\theta)I + \alpha D \end{pmatrix}$$

et $T : [0, \pi[\rightarrow \mathbb{R}$ est donné par $T(x) = \frac{x}{\tan x}$.

LEMME 1. – La matrice H_1 est positive.

Démonstration. – Cela provient de la construction même de t . Soit $x \in \mathbb{S}^n$ fixé. Alors $y = t(x)$ est le point pour lequel $\varphi(x) + \varphi^c(y) = \frac{1}{2} d^2(x, y)$. Donc la fonction $K(v) = \phi(v) + \frac{1}{2} (d^2(v, y) - d^2(x, y))$ atteint un minimum en $v = x$. Si on calcule la matrice de $P_x (\text{Hess } K)_{x|x^\perp}$ dans la base $\mathcal{B}$, on obtient exactement H_1 . $\square$

On poursuit en remarquant que $T(\alpha\theta) - \alpha T(\theta) = (1 - \alpha) \frac{S((1 - \alpha)\theta)}{S(\theta)S(\alpha\theta)}$. Si on note

$$I_{\alpha,\theta} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{S((1 - \alpha)\theta)}{S(\theta)S(\alpha\theta)} I \end{pmatrix},$$

on a donc $H_\alpha = (1 - \alpha)I_{\alpha,\theta} + \alpha H_1$. On utilise alors la log-concavité du déterminant (pour les matrices positives) et on obtient $\det H_\alpha \geq (\det I_{\alpha,\theta})^{1-\alpha} (\det H_1)^\alpha$. Et donc, en utilisant (5),

$$J_\alpha \geq (S(\alpha\theta))^{n-1} (\det I_{\alpha,\theta})^{1-\alpha} (\det H_1)^\alpha = \left(\frac{S(\alpha\theta)}{S(\theta)} \right)^{n-1} (\det I_{\alpha,\theta})^{1-\alpha} J_1^\alpha = \left(\frac{1}{L_\alpha(\theta)} \right)^{n-1} J_1^\alpha.$$

Ce qui montre que $(L_\alpha(\theta))^{n-1} J_\alpha \geq J_1^\alpha$ et achève la démonstration du théorème. $\square$

Remarques. – 1) La même méthode s'applique au cas de l'espace hyperbolique $\mathbb{H}^n$. La fonction S correspondante est alors $S(t) = \frac{\sinh t}{t}$ définie pour tout $t \geq 0$. Cette fonction est log-convexe et le L_α correspondant est plus grand que 1.

2) On peut améliorer l'inégalité en utilisant la $\frac{1}{n}$ -concavité du déterminant au lieu de sa log-concavité. On obtient alors une inégalité de type Henstock–MacBeath [3].

Remerciements. Nous tenons à remercier chaleureusement Bernard Maurey pour ses indications et ses commentaires, et Robert J. McCann pour nous avoir transmis ses résultats.

Références bibliographiques

- [1] Bobkov S.G., Ledoux M., From Brunn–Minkowski to Brascamp–Lieb and to logarithmic Sobolev inequalities, Preprint, 1999.
- [2] Gangbo W., McCann R.J., The geometry of optimal transportation, Acta Math. 177 (1996) 113–161.
- [3] Henstock R., MacBeath A.M., On the measure of sum-sets. I. The theorems of Brunn, Minkowski, and Lusternik, Proc. London Math. Soc. 3 (1953) 182–194.
- [4] Leindler L., On a certain converse of Hölder's inequality, Acta Sci. Math. 33 (1972) 217–233.
- [5] Maurey B., Some deviation inequalities, Geom. and Funct. Anal. 1 (1991) 188–197.
- [6] McCann R.J., Polar factorization of maps on Riemannian manifolds, Preprint, 1999.
- [7] McCann R.J., A convexity principle for interacting gases, Adv. Math. 128 (1997) 153–179.
- [8] Prekopa A., On logarithmically concave measures and functions, Acta Sci. Math. 34 (1973) 335–343.