Séminaire BOURBAKI Novembre 2003
56e année, 2003-2004, nO 928, p. 95 à 113

INÉGALITÉ DE BRUNN-MINKOWSKI-LUSTERNIK, ET AUTRES INÉGALITÉS GÉOMÉTRIQUES ET FONCTIONNELLES

par **Bernard MAUREY**

INTRODUCTION

La théorie des corps convexes a commencé à la fin du XIXe siècle avec l'inégalité de Brunn, généralisée ensuite en inégalité de Brunn-Minkowski-Lusternik qui s'applique à des ensembles non nécessairement convexes. Ce thème a depuis longtemps des contacts avec les problèmes isopérimétriques et avec des inégalités d'Analyse, telles que celles qui traduisent les plongements de Sobolev. Nous allons développer quelques aspects plus récents des inégalités géométriques, dont certains sont liés à la technique du transport de mesure, notamment le transport dit « de Brenier ». On ne trouvera pas ici l'approche d'un résultat faramineux, mais des pistes vers un ensemble convergent de techniques qui ont prouvé leur applicabilité.

L'essentiel de notre travail sera fait dans l'espace euclidien $\mathbb{R}^n$; on notera le produit scalaire par $x \cdot y = \sum_{i=1}^n x_i y_i$ et la norme par $|x| = (x \cdot x)^{1/2}$. On notera $\int_{\mathbb{R}^n} f(x)\,dx$ ou bien $\int_{\mathbb{R}^n} f$ l'intégrale d'une fonction f pour la mesure de Lebesgue sur $\mathbb{R}^n$. On notera $\mathbf{1}_A$ la fonction indicatrice d'un sous-ensemble A de $\mathbb{R}^n$, et $A + B$ la somme de Minkowski de deux sous-ensembles A, B de $\mathbb{R}^n$, égale à $\{a + b : a \in A, b \in B\}$.

1. BRUNN-MINKOWSKI ET INÉGALITÉS GÉOMÉTRIQUES

L'inégalité de Brunn-Minkowski *sans dimension* se formule ainsi : si A, B sont deux compacts non vides de $\mathbb{R}^n$, et si on note $|A|, |B|$ leurs volumes (pour la mesure de Lebesgue), on a

$$\Big|\frac{A+B}{2}\Big| \geqslant |A|^{1/2}|B|^{1/2}.$$

En réalité, l'inégalité $|(1-t)A + tB| \geqslant |A|^{1-t}|B|^t$ est valable pour tout $t \in [0,1]$, et elle se transforme facilement par des arguments d'homogénéité en la forme classique pour l'inégalité de Brunn-Minkowski dans $\mathbb{R}^n$,

$$|A+B|^{1/n} \geqslant |A|^{1/n} + |B|^{1/n}.$$

Si A est convexe et si B est un translaté d'un homothétique de A, l'inégalité devient une égalité. Appliquée en prenant pour B une boule euclidienne de rayon tendant vers 0, cette inégalité conduit à une démonstration de l'inégalité isopérimétrique dans $\mathbb{R}^n$. Démontrée d'abord par Brunn (1887) en dimension 2 ou 3 pour deux convexes, reprise par Minkowski au tout début du XX[e] siècle, l'inégalité a été étendue par Lusternik à des compacts quelconques de $\mathbb{R}^n$; curieusement, le nom de Lusternik est rarement associé de nos jours à cette extension considérable (voir [Lus], [HeM]). Hadwiger et Ohmann [HaO] en donnent une démonstration assez simple, en approchant A et B par des réunions finies de rectangles, et en réalisant une association bien choisie entre parties de A et B de même mesure ; la preuve finit par l'inégalité entre moyennes arithmétiques et géométriques. Ces deux ingrédients se retrouvent dans la plupart des autres preuves.

L'article récent de Richard Gardner [Gar] couvre presque tous les thèmes traités dans cet exposé, et bien d'autres qui ne seront pas abordés ici.

1.1. Prékopa-Leindler

On attribue généralement à Prékopa le résultat suivant : si $\varphi(x,t)$ est convexe sur $\mathbb{R}^n \times \mathbb{R}$, la fonction Φ définie sur $\mathbb{R}$ par

$$\mathrm{e}^{-\Phi(t)} = \int_{\mathbb{R}^n} \mathrm{e}^{-\varphi(x,t)}\, dx$$

est convexe (on admettra les valeurs infinies). L'inégalité de Prékopa-Leindler a un rapport plus direct avec Brunn-Minkowski-Lusternik : on se donne $0 < \theta < 1$ et trois fonctions réelles s.c.i. f_0, f_θ, f_1 sur $\mathbb{R}^n$ qui vérifient pour tous $x_0, x_1 \in \mathbb{R}^n$ l'inégalité

$$(H_{PL}) \qquad f_\theta\big((1-\theta)x_0 + \theta x_1\big) \leqslant (1-\theta) f_0(x_0) + \theta f_1(x_1)$$

et on obtient l'énoncé qui suit.

THÉORÈME 1.1 (Prékopa, Leindler). — *Si les fonctions f_0, f_θ, f_1 satisfont (H_{PL}), on a*

$$\int_{\mathbb{R}^n} \mathrm{e}^{-f_\theta(x)}\, dx \geqslant \Big(\int_{\mathbb{R}^n} \mathrm{e}^{-f_0(x)}\, dx\Big)^{1-\theta} \Big(\int_{\mathbb{R}^n} \mathrm{e}^{-f_1(x)}\, dx\Big)^{\theta}.$$

En attribuant ce résultat aux seuls Prékopa [Pre] et Leindler [Lei], nous faisons un choix de simplicité qui n'est sans doute pas historiquement tout à fait correct ; en effet, la paternité de ce type de résultat a été revendiquée par deux groupes différents (voir Das Gupta [DaG] pour une autre vision de l'histoire). Le théorème 1.1 redonne immédiatement Brunn-Minkowski-Lusternik en prenant les fonctions f_0, f_θ, f_1 égales à 0 sur les ensembles A_0, $A_\theta = (1-\theta)A_0 + \theta A_1$, A_1, et égales à $+\infty$ en dehors, de sorte que pour $j = 0, \theta, 1$ on ait l'égalité $\mathbf{1}_{A_j} = \mathrm{e}^{-f_j}$.

Pour éviter ces valeurs infinies, il est souvent plus agréable d'écrire l'hypothèse sous la forme suivante : trois fonctions réelles positives s.c.s. g_0, g_θ, g_1 sur $\mathbb{R}^n$ vérifient pour tous $x_0, x_1 \in \mathbb{R}^n$ l'inégalité

$$(H_{PL2}) \qquad g_\theta\big((1-\theta)x_0 + \theta x_1\big) \geqslant g_0(x_0)^{1-\theta}\, g_1(x_1)^\theta$$

et on déduit que $\int g_\theta \geqslant (\int g_0)^{1-\theta}(\int g_1)^\theta$. Cette inégalité est adaptée à l'étude des mesures log-concaves. Une mesure μ sur $\mathbb{R}^n$ à densité log-concave s'écrit $d\mu(x) = \mathrm{e}^{-\varphi(x)}\, dx$, avec φ convexe sur $\mathbb{R}^n$; on voit facilement que le triplet $g_j = \mathbf{1}_{A_j}\, \mathrm{e}^{-\varphi}$, $j = 0, \theta, 1$ avec $A_\theta = (1-\theta)A_0 + \theta A_1$ vérifie l'hypothèse (H_{PL2}), ce qui conduit à l'inégalité

$$(1) \qquad \mu\big((1-\theta)A_0 + \theta A_1\big) \geqslant \mu(A_0)^{1-\theta}\mu(A_1)^\theta.$$

Appliquée à une mesure log-concave symétrique (c'est-à-dire invariante par $x \to -x$), cette inégalité donne le résultat de T.W. Anderson (en fait, le résultat d'Anderson [And] est un peu plus général que le théorème 1.2, et il a été montré directement à partir de Brunn-Minkowski) : si $\theta = 1/2$, si $A_0 = C + v$ et $A_1 = C - v$, où C est un convexe symétrique et v un vecteur quelconque, on trouve que $(A_0 + A_1)/2 = C$, pendant que $\mu(A_0) = \mu(A_1)$ par la symétrie de μ ; ainsi, on obtient

$$\mu(C) = \mu\big((A_0 + A_1)/2\big) \geqslant \mu(A_0)^{1/2}\mu(A_1)^{1/2} = \mu(C + v),$$

ce qui signifie ceci : parmi les translatés d'un convexe symétrique, le convexe centré en 0 a la plus grande mesure.

Théorème 1.2 (Anderson). — *Si μ est une mesure log-concave symétrique sur $\mathbb{R}^n$, si C est un convexe symétrique et v un vecteur quelconque, on a*

$$\mu(C + v) \leqslant \mu(C).$$

Un cas particulier important est celui de la mesure gaussienne. Désignons par γ_n la mesure gaussienne standard sur $\mathbb{R}^n$, dont la densité est $(2\pi)^{-n/2}\exp(-|x|^2/2)$. Elle est bien sûr log-concave, symétrique. Le principe précédent s'applique donc à γ_n, et il joue un rôle important dans certaines questions de statistique.

1.2. Isopérimétrie gaussienne

Le problème isopérimétrique gaussien a été résolu par Christer Borell [Bo1], Vladimir Sudakov et Boris Tsirelson [SuT]. Plus tard, Antoine Ehrhard [Ehr] a donné une autre démonstration et apporté d'autres informations, que nous discuterons plus loin. Si A est un sous-ensemble fermé de $\mathbb{R}^n$, on désigne par A_ε son épaississement de taille $\varepsilon > 0$ pour la distance euclidienne,

$$A_\varepsilon = \{x \in \mathbb{R}^n : \mathrm{dist}(x, A) \leqslant \varepsilon\}.$$

Théorème 1.3. — *Parmi les ensembles fermés A de mesure gaussienne $\gamma_n(A) = a$ fixée, les demi-espaces affines minimisent l'accroissement de mesure $\gamma_n(A_\varepsilon) - \gamma_n(A)$, ou ce qui revient au même, minimisent la mesure $\gamma_n(A_\varepsilon)$.*

Par exemple, lorsque $\gamma_n(A) = a = 1/2$, un demi-espace affine B de même mesure que A est donné par $B = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 \leqslant 0\}$, et on peut énoncer en passant aux complémentaires

$$\gamma_n(A_\varepsilon^c) \leqslant \gamma_n(B_\varepsilon^c) = \gamma_n\bigl(\{x_1 > \varepsilon\}\bigr) = \int_\varepsilon^{+\infty} d\gamma_1(t).$$

Si f est une fonction 1-lipschitzienne sur $\mathbb{R}^n$ (vérifiant $|f(x)-f(y)| \leqslant |x-y|$ pour tous x, y) possédant une *valeur médiane* m, c'est-à-dire une valeur telle que $\gamma_n(f \leqslant m) = 1/2$, on voit que f reste inférieure ou égale à $m + \varepsilon$ sur l'épaissi A_ε de $A = \{f \leqslant m\}$, donc

$$\gamma_n\bigl(\{f > m + \varepsilon\}\bigr) \leqslant \gamma_n(A_\varepsilon^c) \leqslant \int_\varepsilon^{+\infty} d\gamma_1(t)\,;$$

on a la même majoration pour $\gamma_n\bigl(\{f < m - \varepsilon\}\bigr)$, ce qui donne une bonne borne pour la probabilité $\gamma_n\bigl(\{|f - m| > \varepsilon\}\bigr)$ que f dévie de plus de ε de sa valeur médiane m,

$$\gamma_n\bigl(\{|f - m| > \varepsilon\}\bigr) \leqslant \gamma_1\bigl(\{t \in \mathbb{R} : |t| > \varepsilon\}\bigr) \leqslant \mathrm{e}^{-\varepsilon^2/2}\,.$$

Cette propriété de concentration a eu de nombreuses applications en *théorie asymptotique*, la théorie des espaces normés de dimension finie tendant vers l'infini, voir par exemple le fameux théorème de Dvoretzky dans le livre de Gilles Pisier [Pis]. Cependant, il est suffisant pour ce type d'applications d'avoir un résultat moins précis, de la forme

$$\gamma_n\bigl(\{|f - m_1| > \varepsilon\}\bigr) \leqslant 2\ \mathrm{e}^{-\varepsilon^2/4}$$

par exemple, où m_1 peut désigner la moyenne de f au lieu de sa médiane, et ce type d'inégalité peut être obtenu de bien des manières (intégrale stochastique, méthodes d'espace gaussien, voir [Led]), mais aussi à partir de Prékopa-Leindler comme ci-dessous.

Proposition 1.4. — *Pour toute fonction f réelle 1-lipschitzienne sur $\mathbb{R}^n$ et pour tout nombre réel t, on a*

$$\int \mathrm{e}^{tf(x)-tf(y)}\ d\gamma_n(x) d\gamma_n(y) \leqslant \mathrm{e}^{t^2}\,.$$

Sous cette forme le résultat est optimal, puisque l'inégalité ci-dessus est une égalité pour les fonctions linéaires. À partir de cette estimation de transformée de Laplace, on obtient les inégalités de concentration par Markov, comme il est habituel.

Démonstration. — Posons

$$g_t(x) = \min\{tf(x+h) + |h|^2/4 : h \in \mathbb{R}^n\}.$$

Le triplet de fonctions $\varphi_0(x) = -g_t(x) + |x|^2/2$, $\varphi_{1/2}(z) = |z|^2/2$, $\varphi_1(y) = tf(y) + |y|^2/2$ vérifie l'hypothèse de Prékopa-Leindler (H_{PL}), car pour tous $x, y \in \mathbb{R}^n$ on a

$$\varphi_0(x) + \varphi_1(y) \geqslant -tf(y) - |x-y|^2/4 + |x|^2/2 + tf(y) + |y|^2/2 = 2\,\varphi_{1/2}\Big(\frac{x+y}{2}\Big).$$

On en déduit que

$$\Big(\int \mathrm{e}^{g_t(x)}\ d\gamma_n(x)\Big)\Big(\int \mathrm{e}^{-tf(y)}\ d\gamma_n(y)\Big) \leqslant \Big(\int d\gamma_n(z)\Big)^2 = 1.$$

Comme $tf(x+h) \geqslant tf(x) - |th|$ par la condition de Lipschitz, et que $tf(x) - |th| + |h|^2/4 \geqslant tf(x) - t^2$ pour tout h, on a $tf(x) \leqslant g_t(x) + t^2$, donc

$$\int \mathrm{e}^{tf(x)-tf(y)}\ d\gamma_n(x)d\gamma_n(y) = \Big(\int \mathrm{e}^{tf(x)}\ d\gamma_n(x)\Big)\Big(\int \mathrm{e}^{-tf(y)}\ d\gamma_n(y)\Big) \leqslant \mathrm{e}^{t^2}. \qquad \square$$

Avec une preuve apparentée à la précédente mais nettement plus subtile, Sergueï Bobkov et Michel Ledoux ont obtenu une démonstration d'une inégalité Log-Sobolev pour γ_n ([BoL]).

1.3. Une inégalité de Brascamp-Lieb

Une inégalité de Brascamp et Lieb [BrL] a trouvé des applications nombreuses en convexité ; on va l'énoncer sous une forme particulière due à Keith Ball [Ba2]. On se place dans $\mathbb{R}^n$ ou dans un espace euclidien E de dimension n. Désignons par $v \otimes v$ l'opérateur de rang un $x \to (x \cdot v)\,v$, et supposons que

$$(J) \qquad \mathrm{Id}_{\mathbb{R}^n} = \sum_{j=1}^N c_j\, v_j \otimes v_j,$$

avec des vecteurs v_j de norme un et des scalaires $c_j > 0$, qui satisfont $\sum_{j=1}^N c_j = n$ par un calcul de trace immédiat. Il en résulte que pour tous $x, y \in \mathbb{R}^n$

$$x \cdot y = \sum_{j=1}^N c_j\,(x \cdot v_j)\,(y \cdot v_j), \quad |x|^2 = \sum_{j=1}^N c_j\,(x \cdot v_j)^2.$$

Cette décomposition de l'identité est fortement liée à une notion géométrique importante, *l'ellipsoïde de John*, et à ses propriétés remarquables ; l'ellipsoïde de John pour un corps convexe symétrique C est l'ellipsoïde maximal contenu dans C ; Fritz John a montré que lorsque cet ellipsoïde est égal à la boule euclidienne unité B, l'identité de $\mathbb{R}^n$ admet la décomposition (J) avec des v_j choisis parmi les points de contact de B et du bord de C.

Théorème 1.5 (Brascamp-Lieb, version Ball). — *Si les vecteurs (v_j) et les nombres (c_j) vérifient la relation (J), on a pour toutes les fonctions positives (f_j) intégrables sur $\mathbb{R}$*

$$\int_{\mathbb{R}^n} \Big(\prod_{j=1}^N f_j(x \cdot v_j)^{c_j}\Big)\,dx \leqslant \prod_{j=1}^N \Big(\int_{\mathbb{R}} f_j(t)\,dt\Big)^{c_j}.$$