Journées

ÉQUATIONS AUX DÉRIVÉES PARTIELLES

Biarritz, 6 juin–10 juin 2011

Michel Ledoux

Analytic and Geometric Logarithmic Sobolev Inequalities

Analytic and Geometric Logarithmic Sobolev Inequalities

Michel Ledoux

Abstract

We survey analytic and geometric proofs of classical logarithmic Sobolev inequalities for Gaussian and more general strictly log-concave probability measures. Developments of the last decade link the two approaches through heat kernel and Hamilton-Jacobi equations, inequalities in convex geometry and mass transportation.

Logarithmic Sobolev inequalities, going back to the works of L. Gross [18], P. Federbush [16], I. Stam [27] and others, are an essential tool in the analysis of the trend to equilibrium in the study of various analytic and probabilistic models for which they provide exponential decays in entropy. One specific feature with respect to classical Sobolev inequalities is their independence with respect to dimension, allowing for the investigation of infinite dimensional systems.

In this short exposition, we briefly survey a number of developments of the last decade at the interface between analysis, probability theory and geometry around this family of functional inequalities. We concentrate in particular on analytic heat kernel and geometric convexity proofs of logarithmic Sobolev inequalities, and analyze their links. Not every detail is make precise here, in particular the classes of functions used for the various inequalities are not always clearly described.

One basic form of the logarithmic Sobolev inequality is the one for the standard Gaussian probability measure $d\mu(x) = (2\pi)^{-n/2} e^{-|x|^2/2} dx$ on $\mathbb{R}^n$ stating that for every smooth positive function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\int_{\mathbb{R}^n} f d\mu = 1$,

$$\int_{\mathbb{R}^n} f \log f d\mu \leq \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} \, d\mu. \tag{1}$$

The constant is sharp and equality is achieved on exponential functions.

This logarithmic Sobolev inequality actually admits various formulations. We refer to [2, 1, 26, 28] for introductions to logarithmic Sobolev inequalities and for complete references. In particular, the transformations below may be performed similarly for any probability measure μ on $\mathbb{R}^n$ given the form (1) (up to constants). For example, f may be changed into f^2 so to yield that for any smooth function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\int_{\mathbb{R}^n} f^2 d\mu = 1$,

$$\int_{\mathbb{R}^n} f^2 \log f^2 d\mu \leq 2 \int_{\mathbb{R}^n} |\nabla f|^2 d\mu. \tag{2}$$

In this form, the logarithmic Sobolev inequality has much similarity with the classical Sobolev inequality. While when the gradient of f is in $\mathrm{L}^2(\mu)$, one cannot assert (and it is wrong in general) that the function itself is in $\mathrm{L}^p(\mu)$ for some $p > 2$, the scale degenerates to the Orlicz space $\mathrm{L}^2 \log \mathrm{L}(\mu)$ (which is critical). On the other hand, no constant depending on the dimension is reported in (2), a fundamental feature of this family of inequalities, allowing for an access to infinite dimensional analysis.

A further description of logarithmic Sobolev inequalities may be provided in a information theoretic terminology. For a given smooth positive function $f : \mathbb{R}^n \to \mathbb{R}$ with $\int_{\mathbb{R}^n} f d\mu = 1$, denote by $d\nu = f \, d\mu$ the probability measure with density f with respect to μ and set

$$H(\nu \,|\, \mu) = \int_{\mathbb{R}^n} f \log f \, d\mu$$

for the relative entropy of ν with respect to μ and

$$I(\nu \,|\, \mu) = \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} \, d\mu$$

for the relative Fisher information. Then the logarithmic Sobolev inequality (1) for the probability measure μ may be recast equivalently as

$$H(\nu \,|\, \mu) \leq \frac{1}{2} I(\nu \,|\, \mu) \tag{3}$$

for every probability measure ν absolutely continuous with respect to μ.

In particular, from a more PDE point of view, denoting by $V(x) = \frac{1}{2}|x|^2$ the quadratic potential underlying the Gaussian measure μ (but what follows may actually be formulated for general regular potentials V), let $L = \Delta - \nabla V \cdot \nabla$ be the heat operator with invariant measure $d\mu = \frac{e^{-V}}{Z} \, dx$. The dual picture with invariant measure the Lebesgue measure is described by the linear Fokker-Planck operator $\tilde{L} = \nabla \cdot [\rho \nabla (\log \rho + V)]$. Changing a smooth positive function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\int_{\mathbb{R}^n} f d\mu = 1$ into a (smooth positive) probability density $\rho = f \frac{e^{-V}}{Z} = f \rho_\infty$ with respect to the Lebesgue measure, the logarithmic Sobolev inequality (1) is then transformed into

$$\int_{\mathbb{R}^n} \rho \log(\rho/\rho_\infty) dx = H\big(\rho \,|\, \rho_\infty\big) \leq \frac{1}{2} I\big(\rho \,|\, \rho_\infty\big) = 2 \int_{\mathbb{R}^n} \big|\nabla\big(\sqrt{\rho/\rho_\infty}\big)\big|^2 \rho_\infty dx \tag{4}$$

for any such probability density ρ.

One specific feature of logarithmic Sobolev inequalities is their independence with respect to the dimension of the underlying state space $\mathbb{R}^n$. However, dimension may actually be dug out for example by developing (4) as

$$\int_{\mathbb{R}^n} \rho \log \rho \, dx + \int_{\mathbb{R}^n} V \rho \, dx + \log Z \leq \frac{1}{2} \int_{\mathbb{R}^n} \left| \nabla (\log \rho + V) \right|^2 \rho \, dx. \qquad (5)$$

Dimension enters here into the picture through the normalization Z, equal to $Z_n = (2\pi)^{n/2}$ for the Gaussian measure. For this example, this dimension effect may also be visualized after the action of dilations in (5) and optimization which then produce the so-called Euclidean logarithmic Sobolev inequality (with respect to the Lebesgue measure thus)

$$\int_{\mathbb{R}^n} \rho \log \rho \, dx \leq \frac{n}{2} \log \left(\frac{1}{2n\pi e} \int_{\mathbb{R}^n} \frac{|\nabla \rho|^2}{\rho} \, dx \right) \qquad (6)$$

for every smooth positive probability density ρ. The Euclidean logarithmic Sobolev inequality (6) is equivalent to the Gaussian logarithmic Sobolev inequality (1), and is sharp on Gaussian functions. Moreover, it is formally equivalent, up to constant, to the classical L^2 Sobolev inequality.

For both the heat and Fokker-Planck descriptions, logarithmic Sobolev inequalities describe equivalently the trend to equilibrium at an exponential rate. For example, if ρ_t is the probability density solution of the evolution

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left[\rho \nabla (\log \rho + V) \right]$$

with initial condition ρ_0, then under the logarithmic Sobolev inequality (4),

$$H\left(\rho_t \mid \rho_\infty \right) \leq e^{-2t} H\left(\rho_0 \mid \rho_\infty \right)$$

for every $t \geq 0$, where we recall that $H(\rho \mid \rho_\infty) = \int_{\mathbb{R}^n} \rho \log(\rho / \rho_\infty) dx$.

In this short expository paper, we thus present two basic approaches to the logarithmic Sobolev inequality (1) (for Gaussian and more general strictly log-concave measures). The first one will be analytic, through heat kernel and semigroup arguments that will actually reveal a number of deeper gradient bounds. The second one is geometric (convexity) in nature, relying on the Brunn-Minkowski inequality. This geometric approach is actually deeply linked with aspects of mass transportation. The two approaches may be related by the concept of hypercontractivity, and vanishing viscosity may be used to directly connect them. In the last section, we briefly outline how the convexity arguments may be used towards (classical) Sobolev inequalities (with their sharp constants). Most of the results presented here are taken from the works [8, 10, 6, 9] (cf. [5, 28] for general references).

1. Analytic heat kernel proof

There are at least fifteen different proofs of the logarithmic Sobolev inequality (1). The following analytic proof, going back to the work of D. Bakry and M. Émery [4], is perhaps the simplest one. Start from the standard heat semigroup on $\mathbb{R}^n$ acting on a suitable function $f : \mathbb{R}^n \to \mathbb{R}$ as

$$P_t f(x) = \int_{\mathbb{R}^n} f(y) \, e^{-|x-y|^2/4t} \, \frac{dy}{(4\pi t)^{n/2}}, \quad t > 0, \ x \in \mathbb{R}^n,$$

with generator the Laplace operator Δ. Note that when $t = \frac{1}{2}$, P_t defines a Gaussian probability measure centered at x (and at $x = 0$ is exactly the standard Gaussian μ).

Fix now $f : \mathbb{R}^n \to \mathbb{R}$ a smooth positive function and $t > 0$. At any point (omitted below),

$$P_t(f \log f) - P_t f \log P_t f = \int_0^t \frac{d}{ds} P_s \Big(P_{t-s} f \log P_{t-s} f \Big) ds.$$

Since $\frac{d}{ds} P_s = \Delta P_s = P_s \Delta$,

$$\frac{d}{ds} P_s \Big(P_{t-s} f \log P_{t-s} f \Big) = P_s \Big(\Delta \big(P_{s-t} f \log P_{t-s} f \big) - \Delta P_{t-s} f \log P_{t-s} f - \Delta P_{t-s} f \Big)$$

$$= P_s \Big(\frac{|\nabla P_{t-s} f|^2}{P_{t-s} f} \Big).$$

Hence

$$P_t(f \log f) - P_t f \log P_t f = \int_0^t P_s \Big(\frac{|\nabla P_{t-s} f|^2}{P_{t-s} f} \Big) ds. \tag{7}$$

Now, gradient and semigroup obviously commute $\nabla P_u = P_u(\nabla)$ so that, by the Cauchy-Schwarz inequality for the Gaussian kernel P_u, for every $u \geq 0$,

$$|\nabla P_u f|^2 \leq \Big[P_u \big(|\nabla f| \big) \Big]^2 \leq P_u \Big(\frac{|\nabla f|^2}{f} \Big) P_u f. \tag{8}$$

This inequality applied for $u = t - s$ thus shows that

$$\frac{|\nabla P_{t-s} f|^2}{P_{t-s} f} \leq P_{t-s} \Big(\frac{|\nabla f|^2}{f} \Big).$$

Inserting in (7), by the semigroup property,

$$P_t(f \log f) - P_t f \log P_t f \leq \int_0^t P_s \Big(P_{t-s} \big(\frac{|\nabla f|^2}{f} \big) \Big) ds = t \, P_t \Big(\frac{|\nabla f|^2}{f} \Big). \tag{9}$$

As announced, at $t = \frac{1}{2}$, this is exactly the logarithmic Sobolev inequality (1) for the standard Gaussian probability measure μ. Note that the only inequality sign in this proof comes up from the Cauchy-Schwarz inequality allowing thus for an easy description of extremal functions (exponentials).

At this stage, it should be pointed out that actually exactly the same proof produces a reverse logarithmic Sobolev inequality. Namely, starting again from (7) and (8), use now the latter with $u = s$ and $P_{t-s}f$ instead of f to get

$$P_t(f \log f) - P_t f \log P_t f \geq \int_0^t \frac{|\nabla P_s(P_{t-s}f)|^2}{P_s(P_{t-s}f)} = t \frac{|\nabla P_t f|^2}{P_t f}. \qquad (10)$$

As will be developed below, the latter actually entails useful gradient bounds.

The preceding heat kernel proof may be developed similarly (cf. [3, 20, 5]) for the semigroups $(P_t)_{t \geq 0}$ with generators $L = \Delta - \nabla V \cdot \nabla$ under a suitable uniform lower bound on the Hessian of the potential V. The semigroup $(P_t)_{t \geq 0}$ is described as the solution $u = u(x, t) = P_t f(x)$ of the initial value problem

$$\frac{\partial u}{\partial t} - Lu = 0 \quad \text{in} \quad \mathbb{R}^n \times (0, \infty)$$

$$u = f \quad \text{on} \quad \mathbb{R}^n \times \{t = 0\}$$

Now whenever, as symmetric matrices, $V'' \geq c \in \mathbb{R}$, it may be shown that

$$|\nabla P_u f| \leq e^{-cu} P_u(|\nabla f|), \quad u \geq 0,$$

and the previous arguments may then be performed identically for both the logarithmic Sobolev inequality (9) and its reverse form (10). The result holds more generally for heat kernels on Riemannian manifolds (or weighted Riemannian manifolds) with a non-negative lower bound on the Ricci curvature. In particular, when $c > 0$, we may let $t \to \infty$ to reach the analogue of (1) for the invariant measure $d\mu = e^{-V} dx$ as the inequality

$$\int_{\mathbb{R}^n} f \log f \, d\mu \leq \frac{1}{2c} \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} \, d\mu$$

for every smooth positive function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\int_{\mathbb{R}^n} f d\mu = 1$. As in the Gaussian case, this inequality is independent of the dimension n and as a result actually shares a basic stability by product (allowing, for product measures, to deduce the mutidimensional inequality for the one-dimensional one). Note also that the reverse logarithmic Sobolev (10) in this context yield simple and robust (dimension free) gradient bounds. For example, if $c = 0$, whenever $0 \leq f \leq 1$, $|\nabla P_t f| \leq t^{-1/2}$ for every $t > 0$.

While independent of the dimension, the Gaussian logarithmic Sobolev inequality still reflects dimension through its Euclidean version (6). This dimension effect may actually also be perceived on the previous semigroup proof along the following lines. Namely the above proof shows that, under the commutation property $\nabla P_u = P_u(\nabla)$, the function

$$\phi(s) = P_s\left(\frac{|\nabla P_{t-s}f|^2}{P_{t-s}f}\right), \quad s \leq t,$$

is non-decreasing. Of course, an alternate way to reach this conclusion is to look for its derivative. Now ϕ may also be written as $\phi(s) = P_s(P_{t-s}f|\nabla \log P_{t-s}f|^2)$ and it is then easily seen that

$$\phi'(s) = 2\, P_s\Big(P_{t-s}f\,\Gamma_2\left(\log P_{t-s}f\right)\Big)$$

where Γ_2 is the so-called Bakry-Émery operator

$$\Gamma_2(h) = \frac{1}{2}\,\Delta\big(|\nabla h|^2\big) - \nabla h \cdot \nabla(\Delta h).$$

On $\mathbb{R}^n$, or more generally on a Riemannian manifold (X, g) equipped with the Laplace-Beltrami operator Δ, Bochner's formula indicates that

$$\Gamma_2(h) = \big\|\mathrm{Hess}(h)\big\|_2^2 + \mathrm{Ric}_g(\nabla h, \nabla h)$$

where Ric_g is the Ricci curvature tensor (which is 0 on the flat space $\mathbb{R}^n$). Then, on $\mathbb{R}^n$ or a Riemannian manifold with non-negative Ricci curvature, $\Gamma_2(h) \geq 0$ and therefore $\phi' \geq 0$ which is the announced claim. The argument may be similarly pushed to weighted Riemannian manifolds X equipped with $d\mu(x) = e^{-V} dx$ where dx is the Riemannian measure and V a smooth potential for which

$$\Gamma_2(h) = \big\|\mathrm{Hess}(h)\big\|_2^2 + \big[\mathrm{Ric}_g + \nabla\nabla V\big](\nabla h, \nabla h). \tag{11}$$

Now actually, as is clear for the definition of the Γ_2 operator, under $\mathrm{Ric}_g \geq 0$ we actually have (by a trace inequality) that

$$\Gamma_2(h) \geq \big\|\mathrm{Hess}(h)\big\|_2^2 \geq \frac{1}{n}\,(\Delta h)^2.$$

Therefore

$$\phi'(s) = 2\, P_s\Big(P_{t-s}f\,\Gamma_2\left(\log P_{t-s}f\right)\Big) \geq \frac{2}{n}\, P_s\Big(P_{t-s}f\,[\Delta \log P_{t-s}f]^2\Big).$$

Integrating the latter (which requires some work) then shows that

$$P_t(f\log f) - P_tf\log P_tf \leq t\,\Delta P_tf + \frac{n}{2}\,P_tf\log\left(1 - \frac{2t}{n}\,\frac{P_t(f\Delta\log f)}{P_tf}\right). \tag{12}$$

This inequality is not immediately appreciable. Note that since $f\Delta\log f = \Delta f - \frac{|\nabla f|^2}{f}$, it may be written equivalently as

$$P_t(f\log f) - P_tf\log P_tf \leq t\,\Delta P_tf + \frac{n}{2}\,P_tf\log\left(1 - \frac{2t}{n}\,\frac{\Delta P_tf}{P_tf} + \frac{2t}{n}\,\frac{1}{P_tf}\,P_t\Big(\frac{|\nabla f|^2}{f}\Big)\right).$$

Using that $\log(1 + u) \leq u$, or better letting $n \to \infty$, we then recover the standard logarithmic Sobolev inequality (9) for P_t. Inequality (12) holds for heat kernels on weighted Riemannian manifolds with non-negative curvature $\Gamma_2 \geq 0$ in the sense of (11).

For $(P_t)_{t\geq 0}$ the standard heat semigroup on $\mathbb{R}^n$, it is of interest as an illustration to take $t = \frac{1}{2}$ to reach the following inequality for the Gaussian measure μ

$$\int_{\mathbb{R}^n} f \log f \, d\mu \leq \frac{1}{2} \int_{\mathbb{R}^n} \Delta f \, d\mu + \frac{n}{2} \log\left(1 - \frac{1}{n}\int_{\mathbb{R}^n} f \Delta \log f \, d\mu\right)$$

for every smooth positive function f such that $\int_{\mathbb{R}^n} f \, d\mu = 1$. This dimensional logarithmic Sobolev inequality is thus stronger than the dimensional logarithmic Sobolev inequality (1). However, if f is changed back into a probability density $\rho = f\mu$ with respect to the Lebesgue measure, we actually end back exactly with the Euclidean logarithmic Sobolev inequality (6). This self-improving property is actually a by-product of the action of dilations in $\mathbb{R}^n$.

Another feature of this investigation is that, as for the standard reverse logarithmic Sobolev inequality (10), there is a reverse form of the dimensional logarithmic Sobolev inequality (12), namely

$$P_t(f \log f) - P_t f \log P_t f \geq t \, \Delta P_t f - \frac{n}{2} \, P_t f \log\left(1 + \frac{2t}{n} \Delta \log P_t f\right).$$

This inequality actually implicitly contains the fact that

$$1 + \frac{2t}{n} \Delta \log P_t f > 0$$

which may be translated equivalently as

$$\frac{|\nabla P_t f|^2}{(P_t f)^2} - \frac{\Delta P_t f}{P_t f} \leq \frac{n}{2t}.$$

This inequality is actually the famous Li-Yau parabolic inequality in Riemannian manifolds with non-negative Ricci curvature [23] which has been proved as a main tool in the investigation of Harnack type inequalities and heat kernel bounds (cf. [13]). It is classically established using the maximum principle while it is imbedded here in a family of logarithmic Sobolev heat kernel inequalities, and holds similarly for weighted manifolds.

2. Geometric convexity proof

The geometric proof will put us to start with in a somewhat different world. The very first starting point is the classical Brunn-Minkowski-Lusternik inequality in Euclidean space which indicates that for bounded (compact) subsets A, B of $\mathbb{R}^n$,

$$\mathrm{vol}_n(A + B)^{1/n} \geq \mathrm{vol}_n(A)^{1/n} + \mathrm{vol}_n(B)^{1/n},$$

$A + B = \{x + y; x \in A, y \in B\}$ being the Minkowski sum of A and B. This inequality is for example typically used to prove the standard isoperimetric inequality in Euclidean space by choosing B a ball with small radius ε tending then to 0. See [17] for a general introduction to classical Brunn-Minkowski inequalities in Euclidean geometry.

Towards our goal, the next step is the functional form of the Brunn-Minkowski-Lusternik inequality known as the Prékopa-Leindler theorem. This theorem indicates that whenever $\theta \in [0,1]$ and u, v, w are non-negative measurable functions on $\mathbb{R}^n$ such that

$$w\big(\theta x + (1-\theta)y\big) \geq u(x)^\theta v(y)^{1-\theta}, \quad x, y \in \mathbb{R}^n, \tag{13}$$

then

$$\int_{\mathbb{R}^n} w \, dx \geq \left(\int_{\mathbb{R}^n} u \, dx \right)^\theta \left(\int_{\mathbb{R}^n} v \, dx \right)^{1-\theta}. \tag{14}$$

Choosing for u and v respectively the characteristic functions of A and B yields the (equivalent, by homogeneity, and dimension free) multiplicative form

$$\mathrm{vol}_n\big(\theta A + (1-\theta)B\big) \geq \mathrm{vol}_n(A)^\theta \mathrm{vol}_n(B)^{1-\theta}$$

of the Brunn-Minkowski-Lusternik inequality.

Modern proof of the Prékopa-Leindler theorem involve mass transportation methods (cf. e.g. [7, 28]). Dimension one is achieved by a suitable parametrization, and dimension n may then be proved by induction. Direct multidimensional mass transportation may also be developed on the basis of the Knothe map, or the Brenier-Rüschendorf transport by the gradient of a convex function. At any rate, the various proofs all boil down at some point to the arithmetic-geometric mean inequality. Mass transportation methods have been significantly developed recently towards notions of Ricci curvature lower bounds in metric measure spaces as well as functional and transportation cost inequalities. We refer to [29] for a comprehensive account on these achievements, and to [21] for a modest introduction in the spirit of this exposition.

To make use of the Prékopa-Leindler theorem in our context, it is necessary to first rewrite it with respect to the standard probability Gaussian measure or more generally a probability measure of the type $d\mu = e^{-V} dx$ for some smooth potential V on $\mathbb{R}^n$. Then, the hypothesis (13) is turned into

$$w\big(\theta x + (1-\theta)y\big) \geq u(x)^\theta v(y)^{1-\theta} \, e^{Z_\theta(x,y)}, \quad x, y \in \mathbb{R}^n$$

where

$$Z_\theta(x, y) = V\big(\theta x + (1-\theta)y\big) - \theta V(x) - (1-\theta)V(y),$$

while the conclusion (14) takes the form

$$\int_{\mathbb{R}^n} w \, d\mu \geq \left(\int_{\mathbb{R}^n} u \, d\mu \right)^\theta \left(\int_{\mathbb{R}^n} v \, d\mu \right)^{1-\theta}.$$

The intervention of the quantity $Z_\theta(x, y)$ actually reflects curvature aspects as is clear on the example of the quadratic potential. More generally, if $V - c\frac{|x|^2}{2}$ is

convex for some $c > 0$ (in other words whenever V is smooth, $V'' \geq c > 0$ as symmetric matrices), for every $x, y \in \mathbb{R}^n$,

$$Z_\theta(x, y) = V\big(\theta x + (1 - \theta)y\big) - V(x) - (1 - \theta)V(y) \leq \frac{c\,\theta(1 - \theta)}{2}\,|x - y|^2. \quad (15)$$

Let now $f : \mathbb{R}^n \to \mathbb{R}$ be a bounded measurable and let $\theta \in [0, 1]$. Choose then

$$w(z) = e^{f(z)}$$
$$v(y) = 1$$
$$u(x) = e^{g(x)}$$

where the function g has to be chosen in order to satisfy the hypothesis in the Prékopa-Leindler theorem, that is such that for every $x, y \in \mathbb{R}^n$,

$$w\big(\theta x + (1 - \theta)y\big) \geq u(x)^\theta v(y)^{1-\theta}\, e^{Z_\theta(x,y)}.$$

In other words,

$$f\big(\theta x + (1 - \theta)y\big) \geq \theta g(x) + Z_\theta(x, y).$$

Now, by (15), it is enough for this purpose that

$$f\big(\theta x + (1 - \theta)y\big) \geq \theta g(x) + \frac{c\,\theta(1 - \theta)}{2}\,|x - y|^2$$

for all $x, y \in \mathbb{R}^n$ so that the optimal choice for g is

$$g(x) = \frac{1}{\theta}\, Q_{(1-\theta)/c\theta} f(x), \quad x \in \mathbb{R}^n,$$

where

$$Q_t f(x) = \inf_{y \in \mathbb{R}^n} \left\{ f(y) + \frac{1}{2t}\,|x - y|^2 \right\}, \quad t > 0,\ x \in \mathbb{R}^n$$

is the infimum-convolution of f with the quadratic cost. The conclusion of the Prékopa-Leindler theorem for $d\mu = e^{-V}dx$, $V'' \geq c > 0$, is then that

$$\int_{\mathbb{R}^n} e^f d\mu \geq \left(\int_{\mathbb{R}^n} e^{\frac{1}{\theta}Q_{(1-\theta)/c\theta}f} d\mu \right)^\theta.$$

Setting $\frac{1}{\theta} = 1 + ct$ yields that

$$\int_{\mathbb{R}^n} e^f d\mu \geq \left(\int_{\mathbb{R}^n} e^{(1+ct)Q_t f} d\mu \right)^{1/(1+ct)} \quad (16)$$

for every $t > 0$.

Now $Q_t f$, $t > 0$, is classically known (cf. [15, 28]) as the Hopf-Lax representation of solutions of the basic Hamilton-Jacobi equation and thus

$$\partial_t Q_t f_{|t=0} = -\frac{1}{2}\,|\nabla f|^2.$$

Differentiating therefore (16) at $t = 0$ yields

$$\int_{\mathbb{R}^n} f\, e^f d\mu - \int_{\mathbb{R}^n} e^f d\mu \log \int_{\mathbb{R}^n} e^f d\mu \leq \frac{1}{2c} \int_{\mathbb{R}^n} e^f |\nabla f|^2 d\mu$$

which, after the change of f into $\log f$, is nothing else than the standard logarithmic Sobolev inequality for μ. Conversely, this logarithmic Sobolev inequality implies back the family of inequalities (16).

Denoting by $\|\cdot\|_p$ the norm in $L^p(\mu)$, $1 \leq p \leq \infty$ (even extended to $p > 0$), the family of inequalities (16) is rewritten as

$$\left\| e^{Q_t f} \right\|_{1+ct} \leq \left\| e^f \right\|_1, \quad t > 0.$$

Actually, for any $a > 0$, the family of inequalities

$$\left\| e^{Q_t f} \right\|_{a+ct} \leq \left\| e^f \right\|_a, \quad t > 0, \tag{17}$$

is still equivalent to the logarithmic Sobolev inequality for μ. An interesting feature happens as $a \to 0$. The logarithmic Sobolev inequality for μ still implies (17) for $a = 0$ but not conversely, and for $a = 0$ these inequalities may be interpreted as the dual form of the quadratic transportation cost inequality

$$W_2(\nu, \mu)^2 \leq \frac{1}{c} H(\nu \mid \mu) \tag{18}$$

between Wasserstein distance W_2 and entropy H for any probability measure ν (absolutely continuous with respect to μ). The implication from the logarithmic Sobolev inequality for μ to the quadratic transportation cost inequality (18) is the famous Otto-Villani theorem [24] (at the starting point of many developments around mass transportation, PDE, geometry of metric measure spaces - cf. [28, 29]) and has been revisited this way in [10].

The property (16) is actually a version for solutions of Hamilton-Jacobi equations of the celebrated hypercontractivity property of E. Nelson for the Ornstein-Uhlenbeck semigroup. Actually, and this is one basic contribution of the seminal paper [18] by L. Gross, the logarithmic Sobolev inequality

$$\int_{\mathbb{R}^n} f \log f \, d\mu \leq \frac{C}{2} \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} \, d\mu \tag{19}$$

for the probability measure $d\mu = e^{-V} dx$ invariant for the operator $L = \Delta - \nabla V \cdot \nabla$ is equivalent to the hypercontractivity property of the associated heat semigroup $P_t = e^{tL}$, $t > 0$, in the sense that whenever $1 < p < q < \infty$ and

$$e^{2t/C} \geq \frac{q-1}{p-1},$$

then

$$\|P_t f\|_q \leq \|f\|_p.$$

The (clever) proof consists in showing that

$$\frac{d}{dt} \|P_t f\|_{q(t)} \leq 0$$

where $q(t) = 1 + e^{2t/C}(p-1)$, $t \geq 0$, if and only if the logarithmic Sobolev inequality (19) holds. This is thus exactly the same picture as the one used for (16). Actually,

one may transit smoothly from the heat equation to the Hamilton-Jacobi equation by means of the vanishing viscosity method which amounts to perturb the latter by a small noise. Consider namely, for $\varepsilon > 0$,

$$\frac{\partial v^\varepsilon}{\partial t} + \frac{1}{2}|\nabla v^\varepsilon|^2 - Lv^\varepsilon = 0 \quad \text{in} \quad \mathbb{R}^n \times (0,\infty)$$
$$v^\varepsilon = f \quad \text{on} \quad \mathbb{R}^n \times \{t = 0\}.$$

Now $u^\varepsilon = e^{-v^\varepsilon/2\varepsilon}$ solves

$$\frac{\partial u^\varepsilon}{\partial t} = \varepsilon\, Lu^\varepsilon$$

and hence is represented as $u^\varepsilon = P_{\varepsilon t}\left(e^{-f/2\varepsilon}\right)$. Thus

$$v^\varepsilon = -2\varepsilon \log P_{\varepsilon t}\left(e^{-f/2\varepsilon}\right)$$

and standard Laplace-Varadhan asymptotics show that

$$\lim_{\varepsilon \to 0} v^\varepsilon = -\lim_{\varepsilon \to 0} 2\varepsilon \log P_{\varepsilon t}\left(e^{-f/2\varepsilon}\right) = Q_t f.$$

With some technical effort, heat hypercontractivity may then be transferred to Hamilton-Jacobi hypercontractivity, providing thus a link between the analytic the geometric approaches (see [10] for details).

3. Classical Sobolev inequalities

One may wonder whether the preceding approaches have anything to say on the classical Sobolev inequalities in $\mathbb{R}^n$,

$$\|f\|_q \leq C_n(p)\, \|\nabla f\|_p$$

for every smooth compactly supported function $f : \mathbb{R}^n \to \mathbb{R}$. Here the norms are with respect to the Lebesgue measure and $1 \leq p < n$, $\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$. In the following brief discussion, we only consider $p = 2$ ($n \geq 3$) but the same strategy applies for any values of $1 < p < n$.

As presented in [28], mass transportation methods have been used earlier by D. Cordero-Erausquin, B. Nazaret, C. Villani [12] (based on [11] dealing with transportation cost and logarithmic Sobolev inequalities) to reach these Sobolev inequalities with their sharp constants (and with a description of the extremal functions). Semigroup tools are not so well suited to this task, although fast diffusion may be used at some point (cf. [14]). This last section is concerned with the relevance of the geometric Brunn-Minkowski approach and is taken from [9]. The picture here is that the Brunn-Minkowski inequality is a way to reach the standard isoperimetric inequality in Euclidean space which in turn is esquivalent to the (sharp) L^1-Sobolev inequality. In the scale of Sobolev inequalities, this L^1-Sobolev inequality implies the L^2-Sobolev inequality, however the optimal constant is lost in this implication.

The point of the approach below is to derive the sharp L^2-Sobolev inequality (actually any p) from the Brunn-Minkowski inequality. For simplicity, we only deal here with $p = 2$ (see [9] for the general case, including further Gagliardo-Nirenberg inequalities).

The first observation is that the Prékopa-Leindler theorem may be used in this context with the choice of, for a given smooth $f > 0$ on $\mathbb{R}^n$ and $\theta \in [0, 1]$,

$$u'x) = f(\theta x)^{-n}$$

$$v(y) = v_\sigma\left(\sqrt{\theta}\, y\right)^{-n}$$

$$w(z) = \left[(1-\theta)\sigma + \theta Q_{1-\theta}f(z)\right]^{-n}$$

where $v_\sigma(x) = \sigma + \frac{|x|^2}{2}$, $\sigma > 0$, $x \in \mathbb{R}^n$, which are related to the extremals functions of the Sobolev inequality. Unfortunately, the application of the Prékopa-Leindler theorem is missing the Sobolev inequality by a dimension defect (essentially n instead of $n - 1$).

In order to overcome this difficulty, it is necessary to consider a sharpened version of the Prékopa-Leindler theorem based on the following elementary but fundamental lemma (in dimension one).

Lemma 1. *Let $\theta \in [0, 1$ and u, v, w be non-negative measurable functions on $\mathbb{R}$. Assume that for all $x, y \in \mathbb{R}$,*

$$w\left(\theta x + (1-\theta)y\right) \geq \min\left(u(x), v(y)\right)$$

and that $\sup_{x \in \mathbb{R}} u(x) = \sup_{x \in \mathbb{R}} v(x) = 1$. Then

$$\int_{\mathbb{R}} w\, dx \geq \theta \int_{\mathbb{R}} u\, dx + (1-\theta) \int_{\mathbb{R}} v\, dx.$$

This lemma has a long history throughout the 20th century (see [19]). It is actually purely equivalent to the Brunn-Minkowski inequality in dimension one and a proof may be found in [9].

A version of the preceding fundamental lemma in dimension n may be obtained by induction on the number of coordinates which yields that whenever $\theta \in [0, 1$ and u, v, w are non-negative measurable functions on $\mathbb{R}^n$ such that for all $x, y \in \mathbb{R}^n$,

$$w\left(\theta x + (1-\theta)y\right)^{-1/(n-1)} \leq \theta u(x)^{-1/(n-1)} + (1-\theta)v(y)^{-1/(n-1)}$$

and for some $i = 1, \ldots, n,, m_i(u) = m_i(v) < \infty$, then

$$\int_{\mathbb{R}^n} w\, dx \geq \theta \int_{\mathbb{R}^n} u\, dx + (1-\theta) \int_{\mathbb{R}^n} v\, dx.$$

Here $m_i(f)$ denotes a constraint in L^∞ of the form

$$m_i(f) = \sup_{x_i \in \mathbb{R}} \int_{\mathbb{R}^{n-1}} f(x)dx_1 \cdots dx_{i-1}dx_{i+1}\cdots dx_n.$$

This version then produces the correct form towards Sobolev inequalities. Namely, given again $f > 0$ smooth and $v_\sigma(x) = \sigma + \frac{|x|^2}{2}$ as above, set now

$$u'x) = f(\theta x)^{1-n}$$

$$v(y) = v_\sigma\left(\sqrt{\theta}\, y\right)^{1-n}$$

$$w(z) = \left[(1-\theta)\sigma + \theta Q_{1-\theta}f(z)\right]^{1-n}$$

where $\sigma = \kappa\theta$ ($\kappa = \kappa(n,f) > 0$) has been chosen so that $m_1(u) = m_1(v)$. The dimensional parameter is here well-set and the fundamental lemma in dimension n then yields that

$$\int_{\mathbb{R}^n} \left[(1-\theta)\sigma + \theta Q_{1-\theta}f(z)\right]^{1-n} dz \geq \theta \int_{\mathbb{R}^n} f(\theta x)^{1-n} dx + (1-\theta) \int_{\mathbb{R}^n} v_\sigma\left(\sqrt{\theta}\, y\right)^{1-n} dy.$$

On the basis of this inequality, the argument develops as in Section 2. Letting $t = 1 - \theta \in\,]0,1[$

$$\int_{\mathbb{R}^n} \left(\kappa t + Q_t f\right)^{1-n} dx \geq \int_{\mathbb{R}^n} f^{1-n} dx + t\, \kappa^{(2-n)/2} \int_{\mathbb{R}^n} v_1^{1-n} dx,$$

the derivative at $t = 0$ yields that

$$(1-n) \int_{\mathbb{R}^n} f^{-n}\left(\kappa - \frac{1}{2}|\nabla f|^2\right) dx \geq \kappa^{(2-n)/2} \int_{\mathbb{R}^n} v_1^{1-n} dx.$$

After the change $g = f^{-2/(n-2)}$,

$$\frac{2}{(n-2)^2} \int_{\mathbb{R}^n} |\nabla g|^2 dx \geq \kappa \int_{\mathbb{R}^n} g^{2n/(n-2)} dx + \frac{1}{(n-1)\kappa^{(n-2)/2}} \int_{\mathbb{R}^n} v_1^{1-n} dx.$$

Taking the infimum over $\kappa > 0$ produces

$$\int_{\mathbb{R}^n} |\nabla g|^2 dx \geq C_n^{-2} \|g\|_{2n/(n-2)}^2$$

where C_n can be checked as the optimal constant in the Sobolev inequality.

References

[1] Ané, C. et al. Sur les inégalités de Sobolev logarithmiques. Panoramas et Synthèses, vol. **10**. Soc. Math. de France (2000).

[2] Bakry, D. L'hypercontractivité et son utilisation en théorie des semigroupes. Ecole d'Eté de Probabilités de St-Flour. Springer Lecture Notes in Math. **1581**, 1-114 (1994).

[3] Bakry, D. Functional inequalities for Markov semigroups. Probability Measures on Groups: Recent Directions and Trends. Proceedings of the CIMPA-TIFR School (2002). Tata Institute of Fundamental Research, New Delhi, 91-147 (2006).

[4] Bakry, D. and Émery, M. Diffusions hypercontractives. Séminaire de Probabilités, XIX. Springer Lecture Notes in Math. **1123**, 177-206 (1985).

[5] Bakry, D., Gentil, I. and Ledoux, M. Forthcoming monograph (2012).

[6] Bakry, D. and Ledoux, M. A logarithmic Sobolev form of the Li-Yau parabolic inequality. Revista Mat. Iberoamericana **22**, 683-702 (2006).

[7] Barthe, F. Autour de l'inégalité de Brunn-Minkowski. Ann. Fac. Sci. Toulouse Math. **12**, 127-178 (2003).

[8] Bobkov, S. and Ledoux, M. From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. Geom. Funct. Anal. **10**, 1028-1052 (2000).

[9] Bobkov, S. and Ledoux, M. From Brunn-Minkowski to sharp Sobolev inequalities. Annali di Matematica Pura ed Applicata **187**, 369-384 (2008).

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

[11] Cordero-Erausquin, D. Some applications of mass transport to Gaussian type inequalities (2000). Arch. Rational Mech. Anal. **161**, 257-269 (2002).

[12] Cordero-Erausquin, D., Nazaret, B. and Villani, C. A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities. Adv. Math. **182**, 307-332 (2004).

[13] Davies, E. B. Heat kernel and spectral theory. Cambridge Univ. Press (1989).

[14] Demange, J. Porous media equation and Sobolev inequalities under negative curvature. Bull. Sci. Math. **129**, 804-830 (2005).

[15] Evans, L. C. Partial differential equations. Graduate Studies in Math. **19**. Amer. Math. Soc. (1997).

[16] Federbush, P. A partially alternate derivation of a result of Nelson. J. Math. Phys. **10**, 50-52 (1969).

[17] Gardner, R. J. The Brunn-Minkowski inequality. Bull. Amer. Math. Soc. **39**, 355-405 (2002).

[18] Gross, L. Logarithmic Sobolev inequalities. Amer. J. Math. **97**, 1061-1083 (1975).

[19] Das Gupta, S. Brunn-Minkowski inequality and its aftermath. J. Multivariate Anal. **10**, 296-318 (1980).

[20] Ledoux, M. The geometry of Markov diffusion generators. Ann. Fac. Sci. Toulouse **IX**, 305-366 (2000).

[21] Ledoux, M. Géométrie des espaces métriques mesurés : les travaux de Lott, Villani, Sturm. Séminaire Bourbaki, Astérisque **326**, 257-280 (2009).

[22] Leindler, L. On a certain converse of Hölder's inequality II. Acta Sci. Math. Szeged **33**, 217-223 (1972).

[23] Li, P. and Yau, S.-T. On the parabolic kernel of the Schrödinger operator. Acta Math. **156**, 153-201 (1986).

[24] Otto, F. and Villani, C. Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality. J. Funct. Anal. **173**, 361-400 (2000).

[25] Prékopa, A. On logarithmic concave measures and functions. Acta Sci. Math. Szeged **34**, 335-343 (1973).

[26] Royer, G. An initiation to logarithmic Sobolev inequalities. Translated from the 1999 French original. SMF/AMS Texts and Monographs **14**. Amer. Math. Soc. / Soc. Math. de France (2007).

[27] Stam, A. Some inequalities satisfied by the quantities of information of Fisher and Shannon. Inform. Control **2**, 101-112 (1959).

[28] Villani, C. Topics in optimal transportation. Graduate Studies in Mathematics **58**. Amer. Math. Soc. (2003).

[29] Villani, C. Optimal transport, old and new. Grundlehren der Mathematischen Wissenschaften, **338**. Springer (2009).

Institut de Mathématiques de Toulouse, Université de Toulouse, F-31062 Toulouse, France, and Institut Universitaire de France
`ledoux@math.univ-toulouse.fr`