Logarithmic Sobolev inequality for diffusion semigroups

Ivan Gentil

June 22, 2009

Abstract

Course given in Grenoble at the 2009 summer school : Optimal transportation, Theory and applications

1 Introduction

The goal of this course is to introduce inequalities as Poincaré or logarithmic Sobolev for diffusion semigroups. We will focus more on examples than on the general theory of diffusion semigroups.

A main tool to obtain those inequalities is the so called Bakry-Emery Γ_2-criterium. This criterium is well known to prove such inequalities and has been also used many times for other problems (see for example [BÉ85, ABC$^+$00, Bak06]).

In Section 4, we will explain an alternative method to get a logarithmic Sobolev inequality under the Γ_2-crierium. It is called the *Mass transportation method* and has been introduced recently (see [CE02, OV00, CENV04, Vil09]). By this way we will also get a another inequality called the *Talagrand inequality or $\mathcal{T}_2$ inequality.*

2 The Ornstein-Uhlenbeck semigroup and the Gaussian measure

A Markov semigroup on $\mathbb{R}^n$ (for $n > 0$) is associated to a Markov process, there are two famous example of diffusion semigroups. The first one is the heat semigroup which is associated to the Brownian motion on $\mathbb{R}^n$. In this course we will study the second one which is the Ornstein-Uhlenbeck semigroup. As we will see in the next section, the Ornstein-Uhlenbeck semigroup is associated to a linear stochastic differential equation driven by a Brownian motion.

In this note a smooth function f in $\mathbb{R}^n$ is a function such that all computation done as integration by parts are justified.

2.1 Definition and general properties

Definition 2.1 *Let define the family of operator $(\mathbf{P_t})_{t\geqslant 0}$: if $f \in \mathcal{C}_b(\mathbb{R}^n)$ then*

$$\mathbf{P_t} f(x) = \int f(e^{-t}x + \sqrt{1-e^{-2t}}y)d\gamma(y), \tag{1}$$

where

$$d\gamma(y) = \frac{e^{-|y|_2^2/2}}{(2\pi n)^{n/2}}dy$$

is the standard Gaussian distribution in $\mathbb{R}^n$ and $|\cdot|$ is the Euclidean norm on $\mathbb{R}^n$.

The family of operator $(\mathbf{P_t})_{t\geqslant 0}$ is called the Ornstein-Uhlenbeck semigroup.

Remark 2.2 *If $(X_t)_{t\geqslant 0}$ is the Markov process defined by the stochastic differential equation*

$$\begin{cases} dX_t = \sqrt{2}dB_t - X_t dt \\ X_0 = 0 \end{cases} \tag{2}$$

then the Itô formula gives that for all continuous and bounded functions f on $\mathbb{R}^n$,

$$\mathbf{P_t} f(x) = E_x(f(X_t)).$$

Since the stochastic differential equation is linear, one can give an explicit solution, equation (1) *is known as the Mehler Formula,*

$$X_t = e^{-t}X_0 + \int_0^t \sqrt{2}e^{s-t}dB_s.$$

Proposition 2.3 *The Ornstein-Uhlenbeck semigroup is a linear operator satisfying the following properties :*

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

(ii) *For all functions $f \in \mathcal{C}_b(\mathbb{R}^n)$, the map $t \mapsto \mathbf{P_t} f$ is continuous from $\mathbb{R}^+$ to $\mathcal{L}^2(d\gamma)$.*

(iii) *For all $s, t \geqslant 0$ one has $\mathbf{P_t} \circ \mathbf{P}_s = \mathbf{P}_{s+t}$.*

(iv) $\mathbf{P_t}1 = 1$ *and* $\mathbf{P_t} f \geqslant 0$ *if* $f \geqslant 0$.

(v) $\|\mathbf{P_t}\|_\infty \leq \|f\|_\infty$.

We say that the Ornstein-Uhlenbeck semigroup is a Markov semigroup on $(\mathcal{C}_b(\mathbb{R}^n), \|\cdot\|_\infty)$.

Proof

$\triangleleft$ We will give only some indications of the proof. First it is easy to prove items (i), (ii), (iv) and (v).

For the item (iii), you just have to compute the Ornstein-Uhlenbeck as follow : $\mathbf{P_t} f(x) = E(f(e^{-t}x + \sqrt{1-e^{-2t}}Y))$ where Y is a random variable with a Gaussian distribution. Then compute $\mathbf{P_t}(\mathbf{P}_s f)$ to obtain $\mathbf{P}_{t+s} f$. In fact, since the solution of the stochastic differential equation (2) is a Markov process then (iii) is a natural property of the Ornstein-Uhlenbeck semigroup. $\triangleright$

Proposition 2.4 *For all functions $f \in \mathcal{C}^2(\mathbb{R}^n)$ bounded with bounded derivatives then one has*

$$\forall x \in \mathbb{R}^n,\ \forall t \geqslant 0,\ \frac{\partial}{\partial t}\mathbf{P_t} f(x) = \mathbf{L}(\mathbf{P_t} f)(x) = \mathbf{P_t}(\mathbf{L} f)(x),$$

where for all smooth functions f, $\mathbf{L} f = \Delta f - x \cdot \nabla f$.

The linear operator $\mathbf{L}$ is known as the infinitesimal generator of the Ornstein-Uhlenbeck semigroup.

Proof

$\triangleleft$ Let us give a sketch of the proof. Let f be a smooth function, then

$$\frac{\partial}{\partial t}\mathbf{P_t} f(x) = \int \left(-e^{-t}x + \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}y\right) \cdot \nabla f\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right) d\gamma(y).$$

By definition of the Ornstein-Uhlenbeck semigroup one gets

$$-xe^{-t} \cdot \int \nabla f\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right) d\gamma(y) = -x \cdot \nabla \mathbf{P_t} f(x)$$

whereas the second term, after an integration by parts gives

$$\frac{e^{-2t}}{\sqrt{1-e^{-2t}}} \int y \cdot \nabla f\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right) d\gamma(y) = \Delta \mathbf{P_t} f(x),$$

which finishes the proof.

Using the same computation one can prove the commutation property between $\mathbf{P_t}$ and the generator $\mathbf{L}$. $\triangleright$

More generally, if $\mathbf{L}$ is an infinitesimal generator associated to a linear semigroup $(\mathbf{P_t})_{t\geqslant 0}$ (not necessary a Markov semigroup) then the commutation $\mathbf{LP_t} = \mathbf{P_t L}$ holds.

Proposition 2.5 (Some properties of the O-U semigroup) *The Ornstein-Uhlenbeck semigroup is γ-ergodic, that means for all $f \in \mathcal{C}_b(\mathbb{R}^n)$,*

$$\forall x \in \mathbb{R}^n, \ \lim_{t\to\infty} \mathbf{P_t} f(x) = \int f d\gamma, \tag{3}$$

in $L^2(d\gamma)$.

The probability measure γ is then the unique invariant measure, for all smooth functions $f \in \mathcal{C}_b(\mathbb{R}^n)$:

$$\int \mathbf{P_t} f d\gamma = \int f d\gamma, \tag{4}$$

or equivalently for all smooth functions f,

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

In fact we have the fundamental identity,

$$\int g\mathbf{L} f d\gamma = \int f\mathbf{L} g d\gamma = -\int \nabla f \cdot \nabla g d\gamma, \tag{5}$$

for all smooth functions on $\mathbb{R}^n$. We say that the Gaussian distribution is reversible with respect to the Ornstein-Uhlenbeck semigroup, $\mathbf{L}$ is symmetric in $L^2(d\gamma)$.

Proof

$\triangleleft$ Let us give the proof of (5):

$$\begin{aligned} \textstyle\int f\mathbf{L} g d\gamma &= \int f\Delta g d\gamma - \int (fx\cdot\nabla g) d\gamma \\ &= -\int \nabla\cdot(f\gamma)\cdot\nabla g dx - \int fx\cdot\nabla g d\gamma \\ &= -\int \nabla f\cdot\nabla g d\gamma, \end{aligned}$$

where $\nabla \cdot f$ stands for the divergence of f.

In fact (4) is clear due to the fact if a semigroup is ergodic for some probability measure then the measure is always invariant. $\triangleright$

As we have seen in the proof of the proposition 2.4 the Ornstein-Uhlenbeck semigroup satisfies the equality for all f and x:

$$\forall t \geqslant 0, \ \nabla \mathbf{P_t} f(x) = e^{-t}\mathbf{P_t}\nabla f(x), \tag{6}$$

where $\mathbf{P_t}\nabla f = (\mathbf{P_t}\partial_i f)_{1\leq i\leq n}$ and for all norms $\|\cdot\|$ in $\mathbb{R}^n$, one gets easily

$$\forall t \geqslant 0, \ \|\nabla \mathbf{P_t} f(x)\| \leq e^{-t}\mathbf{P_t}\|\nabla f\|(x), \tag{7}$$

those equations are known as the commutation property of the gradient and the semigroup. In the next part we will use inequality (7) applied to the Euclidean norm.

2.1.1 The Poincaré and logarithmic Sobolev inequalities

Theorem 2.6 *The following Poincaré inequality for the Gaussian measure holds, for all smooth functions f on $\mathbb{R}^n$,*

$$\mathbf{Var}_\gamma(f) := \int f^2 d\gamma - \left(\int f d\gamma\right)^2 \leq \int |\nabla f|^2 d\gamma. \tag{8}$$

The term $\mathbf{Var}_\gamma(f)$ is called the variance of f under the probability measure γ. Moreover, the inequality is optimal and extremal functions are given by smooth functions satisfying $\nabla f = C$ for some constant $C \in \mathbb{R}^n$.

Proof

$\triangleleft$ Let f be a smooth function on $\mathbb{R}^n$ then $\mathbf{P}_0 f = f$ and by the ergodicity property gives $\mathbf{P}_\infty f = \int f d\gamma$ (see (3)). The Ornstein-Uhlenbeck semigroup gives a nice interpolation between f and $\int f d\gamma$.

$$\begin{aligned}
\mathbf{Var}_\gamma(f) &= -\int_0^{+\infty} \frac{d}{dt} \int (\mathbf{P_t} f)^2 d\gamma dt \\
&= -2 \int_0^{+\infty} \int \mathbf{L} \mathbf{P_t} f \mathbf{P_t} f d\gamma dt \\
&= 2 \int_0^{+\infty} \int |\nabla \mathbf{P_t} f|^2 d\gamma dt \\
&\leq 2 \int_0^{+\infty} \int e^{-2t} (\mathbf{P_t} |\nabla f|)^2 d\gamma dt \\
&\leq 2 \int_0^{+\infty} \int e^{-2t} \mathbf{P_t} \left(|\nabla f|^2\right) d\gamma dt \\
&= 2 \int_0^{+\infty} \int e^{-2t} |\nabla f|^2 d\gamma dt \\
&= \int |\nabla f|^2 d\gamma,
\end{aligned}$$

where we use equality (7), Cauchy-Schwarz inequality and the invariance property of the standard Gaussian distribution (4). $\triangleright$

Theorem 2.7 *The following logarithmic Sobolev inequality for the Gaussian measure holds, for all smooth and non-negative functions f on $\mathbb{R}^n$,*

$$\mathbf{Ent}_\gamma(f) := \int f \log \frac{f}{\int f d\gamma} d\gamma \leq \frac{1}{2} \int \frac{|\nabla f|^2}{f} d\gamma. \tag{9}$$

The term $\mathbf{Ent}_\gamma(f)$ is known as the entropy of f under the measure γ. Moreover, the inequality (9) is optimal and extremal functions are given by $\nabla f = Cf$ for some constant $C \in \mathbb{R}^n$.

Proof

$\triangleleft$ Let us mimic the proof of the Poincaré inequality, let f be a smooth and non-negative function on $\mathbb{R}^n$.

$$\begin{aligned}
\mathbf{Ent}_\gamma(f) &= -\int_0^{+\infty} \frac{d}{dt} \int \mathbf{P_t} f \log \mathbf{P_t} f d\gamma dt \\
&= -\int_0^{+\infty} \int \mathbf{L} \mathbf{P_t} f \log \mathbf{P_t} f d\gamma dt \\
&= \int_0^{+\infty} \int \nabla \mathbf{P_t} f \cdot \nabla \log \mathbf{P_t} f d\gamma dt \\
&= \int_0^{+\infty} \int \frac{|\nabla \mathbf{P_t} f|^2}{\mathbf{P_t} f} d\gamma dt, \\
&\leq \int_0^{+\infty} \int e^{-2t} \frac{(\mathbf{P_t} |\nabla f|)^2}{\mathbf{P_t} f} d\gamma dt
\end{aligned}$$

where we have used the same argument as for Poincaré inequality. Now Cauchy-Schwarz inequality implies

$$\frac{(\mathbf{P_t}|\nabla f|)^2}{\mathbf{P_t}f} \leq \mathbf{P_t}\left(\frac{|\nabla f|^2}{f}\right),$$

or the convexity of the

$$(x,y) \mapsto x^2/y$$

for $x, y > 0$, then one gets

$$\mathbf{Ent}_\gamma(f) \leq \int_0^{+\infty} \int e^{-2t}\mathbf{P_t}\left(\frac{|\nabla f|^2}{f}\right) d\gamma dt = \frac{1}{2}\int \frac{|\nabla f|^2}{f} d\gamma.$$

▷

The logarithmic Sobolev inequality is often noted for the f^2 instead of f, which gives for all smooth functions f,

$$\mathbf{Ent}_\gamma(f^2) \leq 2\int |\nabla f|^2 d\gamma.$$

At the light of the Theorems 2.6 and 2.7, we say that the standard Gaussian satisfies a Poincaré and a logarithmic Sobolev inequality.

More generally a logarithmic Sobolev inequality always implies a Poincaré inequality by using a Taylor expansion (see Chapter 1 of [ABC$^+$00]).

In proposition 2.5, we proved that the Ornstein-Uhlenbeck semigroup is ergodic with respect to the Gaussian distribution. In fact one of the main application of the Poincaré and the logarithmic Sobolev inequalities is to give an estimate of the speed of convergence in two different spaces.

Theorem 2.8 *The Poincaré inequality* (8) *is equivalent to the following inequality*

$$\mathbf{Var}_\gamma(\mathbf{P_t}f) \leq e^{-2t}\mathbf{Var}_\gamma(f)\,, \tag{10}$$

for all smooth functions f.

And in the same way, the logarithmic Sobolev inequality (9) *is equivalent to*

$$\mathbf{Ent}_\gamma(\mathbf{P_t}f) \leq e^{-2t}\mathbf{Ent}_\gamma(f)\,, \tag{11}$$

for all non-negative and smooth functions f.

Proof

◁ For the first assertion, an elementary computation gives that

$$\frac{d}{dt}\mathbf{Var}_\gamma(\mathbf{P_t}f) = -2\int |\nabla \mathbf{P_t}f|^2 d\gamma,$$

then the Poincaré inequality and Grönwall lemma gives (25). Conversely, the derivation at time $t = 0$ of (25) implies the Poincaré inequality.

For the second assertion, we use the same method and the derivation of the entropy,

$$\frac{d}{dt}\mathbf{Ent}_\gamma(\mathbf{P_t}f) = -\int \frac{|\nabla \mathbf{P_t}f|^2}{\mathbf{P_t}f} d\gamma. \tag{12}$$

▷

One of the main difference between the two inequalities is that the initial condition is in $L^2(d\gamma)$ for the Poincaré inequality whereas the initial condition is in $L\log L(d\gamma)$ for the logarithmic Sobolev inequality.

3 Poincaré and Logarithmic Sobolev inequalities under curvature criterium

The main idea of this section is to obtain criteria for a probability measure μ such that the two inequalities (8) and (9) hold for the measure μ. We will study a particular case of the curvature-dimension criterium (or Γ_2-criterium) introduced by D. Bakry and M. Emery in [BÉ85]. This criterium gives conditions on an infinitesimal generator $\mathbf{L}$ such that all the computations done for the Ornstein-Uhlenbeck semigroup could be applied to $\mathbf{L}$.

Let a function $\psi \in \mathcal{C}^2(\mathbb{R}^n)$, and define the infinitesimal generator:

$$\mathbf{L}f = \Delta f - \nabla\psi \cdot \nabla f, \tag{13}$$

for all smooth functions f.

Assume that $\int e^{-\psi} dx < +\infty$ and define the probability measure $d\mu_\psi(x) = \frac{e^{-\psi}dx}{Z_\psi} dx$, where $Z_\psi = \int e^{-\psi} dx$. It is easy to see that the operator $\mathbf{L}$ satisfies for all smooth functions f and g on $\mathbb{R}^n$,

$$\int f\mathbf{L}g d\mu_\psi = \int g\mathbf{L}f d\mu_\psi = -\int \nabla f \cdot \nabla g d\mu_\psi, \tag{14}$$

and $\int \mathbf{L}f d\mu_\psi = 0$. We recover the same property as for the Ornstein-Uhlenbeck semigroup, see (5). As for the Ornstein-Uhlenbeck semigroup, $\mathbf{L}$ is symmetric in $L^2(d\mu_\psi)$ and the probability measure μ_ψ is also invariant with respect to $\mathbf{L}$.

Let define the *Carré du champ*, for all smooth functions f,

$$\Gamma(f, f) = \frac{1}{2}\big(\mathbf{L}(f^2) - 2f\mathbf{L}f\big), \tag{15}$$

we note usually $\Gamma(f)$ instead of $\Gamma(f, f)$. The carré du champ is a quadratic form and the bilinear form associated is given by

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

If we iterate the process one gets the Γ_2-operator, for all smooth functions f,

$$\Gamma_2(f, f) = \frac{1}{2}(\mathbf{L}(\Gamma(f)) - 2\Gamma(f, \mathbf{L}f)). \tag{16}$$

Definition 3.1 *We say that the linear operator* $\mathbf{L}$*, satisfies the* Γ_2*-criterium* $CD(\rho, +\infty)$ *with some* $\rho \in \mathbb{R}$ *if for all smooth functions* f

$$\Gamma_2(f) \geqslant \rho\Gamma(f). \tag{17}$$

Remark 3.2 *Since for all smooth functions* f, $\mathbf{L}f = \Delta f - \nabla\psi \cdot \nabla f$*, a straight forward computation gives,*

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

and

$$\Gamma_2(f) = \|\mathrm{Hess(f)}\|^2_{H.S.} + < \nabla f, \mathrm{Hess}(\psi)\nabla \mathrm{f} >,$$

where the Hilbert-Schmidt norm is given by $\|\mathrm{Hess(f)}\|^2_{H.S.} = \sum_{i,j}\left(\frac{\partial^2}{\partial x_i \partial x_j} f\right)^2$.

Then the linear operator $\mathbf{L}$ *defined in* (13) *satisfies the* Γ_2*-criterium* $CD(\rho, +\infty)$ *with some* $\rho \in \mathbb{R}$ *if for all* $x \in \mathbb{R}^n$

$$\mathrm{Hess}(\psi)(\mathrm{x}) \geqslant \rho \mathrm{Id}, \tag{18}$$

in the sense of the symmetric matrix, i.e. for all $Y \in \mathbb{R}^n$*,*

$$< Y, \mathrm{Hess}(\psi)(\mathrm{x})\mathrm{Y} > \geqslant \rho|\mathrm{Y}|^2,$$

where $< \cdot, \cdot >$ *is the Euclidean scalar product.*

Theorem 3.3 *Let $\psi \in \mathcal{C}^2(\mathbb{R}^n)$ and assume that there exists $\rho > 0$ such that the linear operator (13) satisfies a Γ_2-criterium $CD(\rho, +\infty)$, then the probability measure μ_ψ satisfies a Poincaré inequality*

$$\mathbf{Var}_{\mu_\psi}(f) \leq \frac{1}{\rho} \int |\nabla f|^2 d\mu_\psi, \tag{19}$$

and a logarithmic Sobolev inequality

$$\mathbf{Ent}_\gamma(f) \leq \frac{1}{2\rho} \int \frac{|\nabla f|^2}{f} d\gamma, \tag{20}$$

for all smooth and non-negative functions f.

Lemma 3.4 *Let $(\mathbf{P_t})_{t\geqslant 0}$ be the Markov semigroup associated to the infinitesimal generator* $\mathbf{L}$. *Assume that $\rho > 0$ then $(\mathbf{P_t})_{t\geqslant 0}$ is μ_ψ-ergodic that means for all smooth functions f*

$$\lim_{t\to+\infty} \mathbf{P_t} f(x) = \int f d\mu_\psi,$$

in $f \in L^2(d\mu_\psi)$ and μ_ψ almost surely.

Lemma 3.5 *Let φ be a $\mathcal{C}^2$ function, then for all smooth functions f,*

$$\mathbf{L}\varphi(f) = \varphi'(f)\mathbf{L}f + \varphi''(f)\Gamma(f) \text{ and } \Gamma(\log \mathrm{f}) = \frac{1}{\mathrm{f}^2}\Gamma(\mathrm{f}), \tag{21}$$

moreover one has

$$\Gamma_2(\log f) = \frac{1}{f^2}\Gamma_2(f) - \frac{1}{f^3}\Gamma(f, \Gamma(f)) + \frac{1}{f^4}(\Gamma(f))^2 \tag{22}$$

Proof of the Theorem 3.3

$\triangleleft$ First we will prove the first inequality (19). As for the Ornsten-Uhlenbeck semigroup, one gets if $(\mathbf{P_t})_{t\geqslant 0}$ is the Markov semigroup associated to the infinitesimal generator $\mathbf{L}$,

$$\begin{aligned} \mathbf{Var}_{\mu_\psi}(f) &= -\int_0^{+\infty} \frac{d}{dt} \int (\mathbf{P_t} f)^2 d\mu_\psi dt \\ &= -2 \int_0^{+\infty} \int \mathbf{L}\mathbf{P_t} f \mathbf{P_t} f d\mu_\psi dt \end{aligned}$$

Since μ_ψ is invariant,

$$\int 2\mathbf{P_t} f \mathbf{L}\mathbf{P_t} f d\mu_\psi = \int \left(2\mathbf{P_t} f \mathbf{L}\mathbf{P_t} f - \mathbf{L}(\mathbf{P_t} f)^2\right) d\mu_\psi = -2 \int \Gamma(\mathbf{P_t} f) d\mu_\psi,$$

which gives

$$\mathbf{Var}_{\mu_\psi}(f) = \int_0^{+\infty} 2 \int \Gamma(\mathbf{P_t} f) d\mu_\psi dt. \tag{23}$$

Let now consider for all $t > 0$,

$$\Phi(t) = 2 \int \Gamma(\mathbf{P_t} f) d\mu_\psi,$$

The time derivative of Φ is equal to

$$\Phi'(t) = 4 \int \Gamma(\mathbf{P_t} f, \mathbf{L}\mathbf{P_t} f) d\mu_\psi =$$

$$2 \int (2\Gamma(\mathbf{P_t} f, \mathbf{L}\mathbf{P_t} f) - \mathbf{L}(\Gamma(\mathbf{P_t} f))) d\mu_\psi = -4 \int \Gamma_2(\mathbf{P_t} f) d\mu_\psi.$$

The Γ_2-criterium implies that $\Phi'(t) \leq -2\rho\Phi(t)$ which gives $\Phi(t) \leq e^{-t2\rho}\Phi(0)$. The last inequality with (23) implies

$$\mathbf{Var}_{\mu_\psi}(f) \leq \int_0^{+\infty} e^{-t2\rho}dt \int 2\Gamma(f)d\mu_\psi = \frac{1}{\rho}\int \Gamma(f)d\mu_\psi dt.$$

Let now prove the logarithmic Sobolev inequality for the measure μ_ψ. Let f be a non-negative and smooth function on $\mathbb{R}^n$,

$$\begin{aligned}\mathbf{Ent}_{\mu_\psi}(f) &= -\int_0^{+\infty} \frac{d}{dt}\int \mathbf{P_t} f \log \mathbf{P_t} f d\mu_\psi dt \\ &= -\int_0^{+\infty} \int \mathbf{L}\mathbf{P_t} f \log \mathbf{P_t} f d\mu_\psi dt\end{aligned}$$

Since $\mathbf{L}$ is symmetric and by lemma 3.5 one gets

$$\int \mathbf{L}\mathbf{P_t} f \log \mathbf{P_t} f d\mu_\psi = \int \mathbf{P_t} f \mathbf{L} \log \mathbf{P_t} f d\mu_\psi = -\int \frac{\Gamma(\mathbf{P_t} f)}{\mathbf{P_t} f} d\mu_\psi = -\int \Gamma(\log \mathbf{P_t} f)\mathbf{P_t} f d\mu_\psi,$$

which gives

$$\mathbf{Ent}_{\mu_\psi}(f) = \int_0^{+\infty} \int \Gamma(\log \mathbf{P_t} f)\mathbf{P_t} f d\mu_\psi dt. \tag{24}$$

As for Poincaré inequality, let consider for all $t > 0$,

$$\Phi(t) = \int \frac{\Gamma(\mathbf{P_t} f)}{\mathbf{P_t} f} d\mu_\psi$$

where $\mathbf{P_t} f = g$. The time derivative of Φ is equal to

$$\Phi'(t) = \int \left(2\frac{\Gamma(\mathbf{L}g, g)}{g} - \frac{\mathbf{L}g\Gamma(g)}{g^2}\right)\mu_\psi = \int \left(2\frac{\Gamma(\mathbf{L}g, g)}{g} - \frac{\mathbf{L}g\Gamma(g)}{g^2} - \mathbf{L}\left(\frac{\Gamma(g)}{g}\right)\right)\mu_\psi.$$

Since

$$\mathbf{L}\left(\frac{\Gamma(g)}{g}\right) = 2\Gamma\left(\Gamma(g), \frac{1}{g}\right) + \frac{1}{g}\mathbf{L}\Gamma(g) + \mathbf{L}\left(\frac{1}{g}\right)\Gamma(g),$$

by Lemma 3.5 one has

$$\Phi'(t) = -2\int \Gamma_2(\log \mathbf{P_t} f)\mathbf{P_t} f d\mu_\psi.$$

The Γ_2-criterium implies that $\Phi'(t) \leq -2\rho\Phi(t)$ which gives $\Phi(t) \leq e^{-2\rho t}\Phi(0)$. This inequality with (24) implies that

$$\mathbf{Ent}_{\mu_\psi}(f) \leq \int_0^{+\infty} e^{-2\rho t}dt \int \Gamma(\log f) f d\mu_\psi = \frac{1}{2\rho}\int \Gamma(\log f) f d\mu_\psi = \frac{1}{2\rho}\int \frac{|\nabla f|}{f} d\mu_\psi.$$

$\triangleright$

The meaning of this result is : if μ_ψ is more log-concave than the Gaussian distribution then μ_ψ satisfies both inequalities.

Remark 3.6 *The Γ_2-criterium is in fact a more general criterium. The definition of a diffusion semigroup could be a Markov semigroup such that for all smooth functions φ, the equations* (21) *and* (22) *hold for the generator associated to the semigroup.*

In fact that means that the infinitesimal generator $\mathbf{L}$ of the Markov semigroup is given by,

$$\forall x \in \mathbb{R}^n,\ \mathbf{L}f(x) = \sum_{i,j} D_{i,j}(x)\partial_{i,j} f(x) - \sum_i a_i(x)\partial_i f(x),$$

where $D(x) = (D_{i,j}(x))_{i,j}$ is a symmetric and non-negative matrix and $a(x) = (a_i(x))_i$ is a vector.

Then the conditions $\Gamma_2(f) \geqslant \rho\Gamma(f)$ for some $\rho > 0$ implies that there exists an invariant measure μ of the semigroup and μ satisfies the Poincaré and a logarithmic Sobolev inequality with the same constant as before. One of the difficulties of this general case is to find tractable conditions on functions D and a such that the Γ_2-criterium holds. Some others examples can be found in [BG09].

Let us also note that the Γ_2-criterium $CD(\rho,\infty)$ is a particular case of the $CD(\rho,n)$ criterium where $n \in \mathbb{N}^*$:

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

for all smooth functions f. For example, the Ornstein-Uhlenbeck semigroup satisfies the $CD(1,\infty)$ criterium and the heat equation $\mathbf{L} = \Delta$ satisfies the $CD(0,n)$.

Theorem 3.7 *As for the Ornstein-Uhlenbeck semigroup, the Poincaré inequality* (19) *is equivalent to the following inequality*

$$\mathbf{Var}_{\mu_\psi}(\mathbf{P_t}f) \leq e^{-\frac{2}{\rho}t}\mathbf{Var}_{\mu_\psi}(f), \tag{25}$$

for all smooth functions $f \in L^2(d\mu_\psi)$.

And in the same way, the logarithmic Sobolev inequality (20) *is equivalent to*

$$\mathbf{Ent}_{\mu_\psi}(\mathbf{P_t}f) \leq e^{-2t}\mathbf{Ent}_{\mu_\psi}(f), \tag{26}$$

for all non-negative and smooth functions $f \in L\log L(d\mu_\psi)$ (it means that $\mathbf{Ent}_{\mu_\psi}(f) < +\infty$).

The logarithmic Sobolev inequality has two main applications. The first one the asymptotic behaviour in term of entropy, this is the result of Theorem 3.7. The second application is about concentration inequality, a probability measure μ satisfying a logarithmic Sobolev inequality has the same tail as the Gaussian distribution.

This properties can also be found in the Talagrand inequality described in the next section.

4 The Logarithmic Sobolev and transportation inequalities by transportation method

Let us see how the Brenier's Theorem and the Wasserstein distance can be used in this context. The method come from [OV00, CE02] and has been generalized for many inequalities in [AGK04, CENV04, Naz06].

Let recall the Wasserstein distance between two probability measures μ and ν,

$$T_2(\mu,\nu) = \left(\inf\left\{\int |x-y|^2 d\pi(x,y)\ ;\ \pi \in P(\mu,\nu)\right\}\right)^{1/2}. \tag{27}$$

where the infimum is running over all probability measure π on $\mathbb{R}^n \times \mathbb{R}^n$ with respective marginals μ and ν: for all functions g and h

$$\int (g(x)+h(y))d\pi(x,y) = \int g d\mu + \int h d\nu.$$

Let f be non-negative such that $\int f d\gamma = 1$. Let $\nabla\Phi$ be the Brenier map between $f d\gamma$ and γ: for all bounded and measurable functions h,

$$\int h(\nabla\Phi) f d\gamma = \int h d\gamma,$$

and

$$W_2^2(f d\gamma, d\gamma) = \int |\nabla\theta|^2 f d\gamma,$$

where $\theta(x) = \Phi(x) - \frac{1}{2}|x|^2$. Denoting by Id the identity matrix, we have $Id + \text{Hess}(\theta) > 0$. The Monge-Ampére equation holding $fd\gamma$-a.e. is:

$$f(x)e^{-|x|^2/2} = \det(\mathrm{Id} + \mathrm{Hess}(\theta))\mathrm{e}^{-|\mathrm{x}+\nabla\theta(\mathrm{x})|^2/2}. \tag{28}$$

After taking the logarithm, we have:

$$\begin{aligned}
\log f(x) &= -\frac{1}{2}|x+\nabla\theta(x)|^2 + \frac{1}{2}|x|^2 + \log\det(Id + \text{Hess}(\theta)) \\
&= -x\cdot\nabla\theta(x) - \frac{1}{2}|\nabla\theta(x)|^2 + \log\det(Id + \text{Hess}(\theta)) \\
&\leq -x\cdot\nabla\theta(x) - \frac{1}{2}|\nabla\theta(x)|^2 + \Delta\theta(x),
\end{aligned}$$

where we used $\log(1+t) \leq t$ whenever $1+t > 0$. We integrate with respect to $fd\gamma$:

$$\mathbf{Ent}_\gamma(f) \leq \int f(\Delta\theta - x\cdot\nabla\theta)d\gamma - \int \frac{1}{2}|\nabla\theta(x)|^2 f d\gamma.$$

By integration by parts (14) we get:

$$\begin{aligned}
\mathbf{Ent}_\gamma(f) &\leq -\int \nabla\theta\cdot\nabla f d\gamma - \int \frac{1}{2}|\nabla\theta(x)|^2 f d\gamma \\
&\leq -\frac{1}{2}\int \left|\sqrt{f}\nabla\theta + \frac{\nabla f}{\sqrt{f}}\right|^2 d\gamma + \frac{1}{2}\int \frac{|\nabla f|^2}{f}d\gamma \\
&\leq \frac{1}{2}\int \frac{|\nabla f|^2}{f}d\gamma,
\end{aligned}$$

which is inequality (9).

Hence we have proved, using the Brenier's map, the logarithmic Sobolev inequality for the Gaussian measure.

Let us see what can be done if now $\nabla\Phi$ be the Brenier map between $d\gamma$ and $fd\gamma$ that is for all bounded and measurable functions h:

$$\int hf d\gamma = \int h(\nabla\Phi)d\gamma,$$

and if $x + \nabla\theta(x) = \nabla\theta$ then

$$W_2^2(fd\gamma, d\gamma) = \int |\nabla\theta|^2 d\gamma.$$

In that case the Monge-Ampère equation gives

$$\det(\mathrm{Id} + \mathrm{Hess}(\theta))\mathrm{f}(\mathrm{x} + \nabla\theta(\mathrm{x}))\mathrm{e}^{-|\mathrm{x}+\nabla\theta(\mathrm{x})|^2/2} = \mathrm{e}^{-|\mathrm{x}|^2/2}. \tag{29}$$

Which implies

$$\begin{aligned}
\log f(x+\nabla\theta(x)) &= \frac{1}{2}|x+\nabla\theta(x)|^2 - \frac{1}{2}|x|^2 - \log\det(Id + \text{Hess}(\theta)) \\
&= x\cdot\nabla\theta(x) + \frac{1}{2}|\nabla\theta(x)|^2 - \log\det(Id + \text{Hess}(\theta)) \\
&\geqslant x\cdot\nabla\theta(x) + \frac{1}{2}|\nabla\theta(x)|^2 - \Delta\theta(x) \\
&= -L\theta + \frac{1}{2}|\nabla\theta(x)|^2.
\end{aligned}$$

Then

$$\begin{aligned}
\mathbf{Ent}_\gamma(f) &= \int f\log f d\gamma \\
&= \int \log f(\nabla\Phi)d\gamma \\
&\geqslant \int -L\theta d\gamma + \int \frac{1}{2}|\nabla\theta(x)|^2 d\gamma \\
&= \textstyle\int \frac{1}{2}|\nabla\theta(x)|^2 d\gamma = \frac{1}{2}T_2^2(fd\gamma, d\gamma)
\end{aligned}$$

We prove that for all function f such that $fd\gamma$ is a probability measure, one has

$$T_2(fd\gamma, d\gamma) \leq \sqrt{2\mathbf{Ent}_\gamma(f)}. \tag{30}$$

This inequality is called *Talagrand inequality for the Gaussian distribution (or $\mathcal{T}_2$ inequality)* and has been proved by Talagrand in [Tal96].

As for Poincaré and logarithmic Sobolev inequalities, we says that a probability measure μ satisfies a Talagrand inequality if there exists $C \geqslant 0$ such that,

$$T_2(fd\mu, d\mu) \leq \sqrt{C\mathbf{Ent}_\mu(f)}, \tag{31}$$

for all functions f such that $fd\mu$ is a probability measure,

4.1 Remarks and extensions

This method can also be used is the context of the section 3. Assume that ψ is uniformly convex, satisfying

$$\mathrm{Hess}(\psi) \geqslant \rho \mathrm{I},$$

with some $\rho > 0$. The mass transportation method implies that the measure

$$d\mu_\psi(x) = \frac{e^{-\psi}dx}{Z_\psi}dx$$

satisfies the logarithmic Sobolev inequality (20) with the same constant $1/(2\rho)$. This is an alternative proof of Theorem 3.3. Let us remark that the method is not useful to get directly a Poincaré inequality.

Of course, as for Ornstein-Uhlenbeck semigroup, the mass transportation method gives also a talagrand inequality (31).

$$T_2(fd\mu_\psi, d\mu_\psi) \leq \sqrt{\frac{1}{\rho}\mathbf{Ent}_{\mu_\psi}(f)},$$

for all probability measure $fd\mu_\psi$.

In fact we have the general result,

Theorem 4.1 (Otto-Villani 2000) *Let μ be a probability measure on $\mathbb{R}^n$ satisfying a logarithmic Sobolev inequality*

$$\mathbf{Ent}_\mu(f^2) \leq C \int |\nabla f|^2 d\mu,$$

for all smooth functions f and for some constant $C > 0$.

Then μ satisfies a Talagrand inequality

$$T_2(fd\mu, d\mu) \leq \sqrt{2C\mathbf{Ent}_\mu(f)},$$

for all probability measure $fd\mu$.

The original proof comes from [OV00] and an easier one, using Hamilton-Jacobi equation, can be seen in [BGL01]. These two inequalities are quite similar but it has been proved in [CG06, Goz07] that they are not equivalent.

The main application of the Talagrand inequality is the same as for the logarithmic Sobolev inequality, a probability measure satisfying a Talagrand inequality has the same tail as the Gaussian distribution.

References

[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*. Société Mathématique de France, Paris, 2000.

[AGK04] M. Agueh, N. Ghoussoub, and X. Kang. Geometric inequalities via a general comparison principle for interacting gases. *Geom. Funct. Anal.*, 14(1):215–244, 2004.

[Bak06] D. Bakry. Functional inequalities for Markov semigroups. In *Probability measures on groups: recent directions and trends*, pages 91–147. Tata Inst. Fund. Res., Mumbai, 2006.

[BÉ85] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, Lecture Notes in Math. 1123, pages 177–206. Springer, Berlin, 1985.

[BG09] F. Bolley and I. Gentil. Phi-entropy inequalities for diffusion semigroups. *Prépublication*, 2009.

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

[CE02] D. Cordero-Erausquin. Some applications of mass transport to Gaussian-type inequalities. *Arch. Ration. Mech. Anal.*, 161(3):257–269, 2002.

[CENV04] D. Cordero-Erausquin, B. Nazaret, and C. Villani. A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities. *Adv. Math.*, 182(2):307–332, 2004.

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

[Goz07] N. Gozlan. Characterization of Talagrand's like transportation-cost inequalities on the real line. *J. Funct. Anal.*, 250(2):400–425, 2007.

[Naz06] B. Nazaret. Best constant in Sobolev trace inequalities on the half-space. *Nonlinear Anal.*, 65(10):1977–1985, 2006.

[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.

[Tal96] M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.

[Vil09] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

Université Paris-Dauphine
CEREMADE
Place du Maréchal De Lattre De Tassigny
75116 PARIS - France
gentil@ceremade.dauphine.fr