Logarithmic Sobolev inequality for diffusion semigroups

Ivan Gentil*

September 20, 2010

Abstract

Through the main example of the Ornstein-Uhlenbeck semigroup, the Bakry-Emery criterion is presented as a main tool to get functional inequalities as Poincaré or logarithmic Sobolev inequalities. Moreover an alternative method using the optimal mass transportation, is also given to obtain the logarithmic Sobolev inequality.

Mathematics Subject Classification (2000) : Primary 35B40, 35K10, 60J60.

Keywords : Logarithmic Sobolev inequality, Poincaré inequality, Ornstein-Uhlenbeck semigroup, Bakry-Emery criterion.

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. 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 instance [BÉ85, Bak06]. We will focus on the example of the Ornstein-Uhlenbeck semigroup and on the Γ_2 -criterium.

In section 2 we investigate the main example of the Ornstein-Uhlenbeck semigroup whereas in section 3 we show how the Γ_2-crierium implies such inequalities. In section 4, we will explain an alternative method to get a logarithmic Sobolev inequality under curvature assumption. It is called the *mass transportation method* and has been introduced recently, see [CE02, OV00, CENV04, Vil09]. By this way we will also obtain an another inequality called the *Talagrand inequality or $\mathcal{T}_2$ inequality.*

2 The Ornstein-Uhlenbeck semigroup and the Gaussian measure

In the general setting if $(X_t)_{t\geqslant 0}$ is a Markov process on $\mathbb{R}^n$ then the family of operators :

$$\mathbf{P}_t(f)(x) = E(f(X_t)),$$

where $X_0 = x$ and a smooth function f, defined is Markov semigroup on $\mathbb{R}^n$. There are two main examples. The first one is the heat semigroup which is associated to the Brownian motion on

*Ceremade, UMR CNRS 7534, Université Paris-Dauphine.

$\mathbb{R}^n$. In this course we will study the second one which is the Ornstein-Uhlenbeck semigroup. We will see that 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, for example $C_c^\infty(\mathbb{R}^n)$.

2.1 Definition and general properties

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

$$\mathbf{P}_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}}y)d\gamma(y), \tag{1}$$

where

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

$\nabla\gamma(y) = -y\,\gamma(y)$

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 *Let $(X_t)_{t\geqslant 0}$ be a Markov process, solution of the stochastic differential equation*

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

Since the stochastic differential equation is linear, there is an explicit solution

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

and equation (1) *is known as the Mehler Formula. Moreover Itô's formula gives that for all continuous and bounded functions f on $\mathbb{R}^n$*

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

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

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

(ii) For all functions $f \in C_b(\mathbb{R}^n)$, 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) $\|\mathbf{P}_t f\|_\infty \leq \|f\|_\infty$.

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

Proof

$\lhd$ We will give only some indication~~s~~ 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. $\rhd$

Proposition 2.4 *For all smooth functions f 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, $\boxed{\mathbf{L}f = \Delta f - x\cdot\nabla f.}$

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

$$\mathbb{L}f(x) = \lim_{h\downarrow 0} \frac{P_h f(x) - f(x)}{h}$$

Proof

$\lhd$ If 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

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

recall here $\nabla\gamma(y) = -y\gamma(y)$

which finishes the proof.

Using the same computation one can prove the commutation property between $\mathbf{P}_t$ and the generator $\mathbf{L}$. $\rhd$

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 $\mathbf{L}\mathbf{P}_t = \mathbf{P}_t\mathbf{L}$ 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 probability 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,

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

for all smooth functions f and g 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

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

Relies on the fact $\partial_i \gamma(x) = -x_i \gamma(x)$

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

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. $\rhd$

As we have seen in the proof of 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 Ornstein-Uhlenbeck semigroup. Inequality (7) is the key formula to get classical inequalities.

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 $\mathbb{R}^n$,*

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

POINCARÉ

The term $\mathbf{Var}_\gamma(f)$ is the variance of f under 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

$\lhd$ Let f be a smooth function on $\mathbb{R}^n$ then $\mathbf{P}_0 f = f$ and $\mathbf{P}_\infty f = \int f d\gamma$ (see (3)), therefore the

With $V^f(t) \triangleq \int_{\mathbb{R}^n} (P_t f(x))^2 d\gamma(x)$ we have $\mathrm{Var}_\gamma(f) = V^f(0) - V^f(\infty) = -\int_0^\infty \dot{V}^f(t)\, dt$

In particular,

$$\dot{V}^f(t) = \frac{d}{dt} V^f(t) = 2\int_{\mathbb{R}^n} \frac{d}{dt} P_t f(x) . P_t f(x) d\gamma(x) = 2\int_{\mathbb{R}^n} L P_t f(x) . P_t f(x) . d\gamma(x) = -\int_{\mathbb{R}^n} |\nabla P_t f|^2(x)\, d\gamma(x)$$

The quantity $\underline{I_\gamma(f) \triangleq \int \frac{|\nabla f|^2}{f} d\gamma}$ is the Fisher Information of f.

The proof of Thm 2.7 is based on the representation:

$$\underline{\mathrm{Ent}_\gamma(f) = \int_0^\infty I_\gamma(P_t f)\,dt}$$

More generally,

$$\underline{\mathrm{Ent}_\gamma(f) - \mathrm{Ent}_\gamma(P_T f) = \int_0^T I_\gamma(P_t f)\,dt}$$

Let Z be a RV with distr. γ, and $h(x) = x\log x$ a convex function. Then $\mathrm{Ent}_\gamma(f) = \mathbb{E}[h(Z)] - h(\mathbb{E}(Z)) \geq 0$

On the other hand, with $E(t) \triangleq \int_{\mathbb{R}^n} h(P_t f(x))\,d\gamma(x)$ we have $\mathrm{Ent}_\gamma(f) = E(0) - E(\infty) = -\int_0^\infty \dot{E}(t)\,dt$

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 (5), (7), Cauchy-Schwarz inequality and the invariance property of the standard Gaussian distribution (4).
On can check that in all stages of the proof, smooth functions satisfying $\nabla f = C$ are the unique function such that the two inequalities become equalities. ▷

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

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

LOG-SOBOLEV

The term $\mathbf{Ent}_\gamma(f)$ is known as the entropy of f under 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
◁ Let us mimic the proof of the Poincaré inequality, let f be a smooth and non-negative function on $\mathbb{R}^n$ then

$$\begin{aligned}
\mathbf{Ent}_\gamma(f) &= -\int_0^{+\infty} \frac{d}{dt}\int \mathbf{P_t} f \log \mathbf{P_t} f d\gamma dt = -\int_0^\infty \underbrace{\frac{d}{dt}\left(\int_{\mathbb{R}^n} h(P_t f(x))\,d\gamma(x)\right)}_{\int_{\mathbb{R}^n} LP_t f(x)\cdot \underbrace{h'(P_t f(x))}_{P_t f(x) + 1} d\gamma(x)} dt \\
&= -\int_0^{+\infty}\int \mathbf{L}\mathbf{P_t} f \log \mathbf{P_t} f d\gamma dt \\
&= \int_0^{+\infty}\int \nabla \mathbf{P_t} f \cdot \nabla \log \mathbf{P_t} f d\gamma dt \\
&= \int_0^{+\infty}\int \frac{|\nabla \mathbf{P_t} f|^2}{\mathbf{P_t} f} d\gamma dt, \\
&\leq \int_0^{+\infty}\int e^{-2t}\frac{(\mathbf{P_t}|\nabla f|)^2}{\mathbf{P_t} f} d\gamma dt
\end{aligned}$$

Now recall (5)

where we have used the same argument as for Poincaré inequality. Now Cauchy-Schwarz inequality or the convexity of the map

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

$$P_t(|\nabla f|) = P_t\left(\frac{|\nabla f|}{\sqrt{f}}\cdot\sqrt{f}\right) \leq \sqrt{P_t\left(\frac{|\nabla f|^2}{f}\right)\cdot P_t(f)}$$

for $x, y > 0$, implies

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

then one gets

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

One obtains extremal functions in the same way than for Poincaré inequality. $\rhd$

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

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

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 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

$\lhd$ For the first assertion, an elementary computation gives that

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

From the computations on p. 4;
$\mathrm{Var}_\gamma(P_t f) = V^{P_t f}(0) - V^{P_t f}(\infty) = V^f(t) - \left(\int f d\gamma\right.$

then the Poincaré inequality and Grönwall lemma implies (10). 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,

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

$\rhd$

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.

2. $\Gamma_2(f,g) \triangleq \mathbf{L}(\Gamma(f,g)) - \Gamma(\mathbf{L}f,g) - \Gamma(f,\mathbf{L}g)$

More general hypercontractive diffusions.

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 $\mathbf{L}$.

Recall here Ex. 5.6.18 p. 361 in [KS]

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

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

For the OU semig. $\psi(x) = \frac{x^2}{2}$, $\mu_\psi = \gamma$

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). The generator $\mathbf{L}$ 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,

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

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

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

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

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

We assume in this section that there exits a set of function $\mathcal{A}$, dense in $L^2(d\mu)$, such that all computations can be done in this class of function. In the previous section, the set $\mathcal{A}$ was $C_c^\infty(\mathbb{R}^n)$ and one of the main problem is to describe this class of functions. It can be done under the Γ_2-criterium $CD(\rho,+\infty)$ (see the definition below), we refer to [ABC+00, Bak06] and references therein to get more informations.

Bakry-Emery Criterion

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 functions* $f \in \mathcal{A}$

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

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

$$\Gamma(f) = |\nabla f|^2, \qquad \Gamma(f,g) = \nabla f \cdot \nabla g \qquad (17.a)$$

and

$$\Gamma_2(f) = \|\mathrm{Hess(f)}\|^2_{H.S.} + < \nabla f, \mathrm{Hess}(\psi)\nabla \mathrm{f} >,$$ (17b)

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* $x \in \mathbb{R}^n$

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

(Strict Convexity)

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

$$\boxed{\langle \mathrm{Y}, \mathrm{Hess}(\psi)(\mathrm{x})\mathrm{Y} \rangle \geqslant \rho |\mathrm{Y}|^2,}$$

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

$\xi' \partial^2 \psi(x) \xi \geqslant \rho |\xi|^2$
Strict Convexity of $\psi(\cdot)$, strict log-Concavity of $e^{-\psi(\cdot)}$.

Theorem 3.3 *Let* $\psi \in C^2(\mathbb{R}^n)$ *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*

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

Equality with $\rho = 1$ for Gaussian.

for all $f \in \mathcal{A}$ *and a logarithmic Sobolev inequality*

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

for all smooth and non-negative functions $f \in \mathcal{A}$.

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

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

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

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

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

moreover one has

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

Proof of the Theorem 3.3
$\triangleleft$ First we prove the first inequality (19). As for the Ornsten-Uhlenbeck semigroup, one gets if $(\mathbf{P}_\mathrm{t})_{t \geqslant 0}$ is the Markov semigroup associated to the infinitesimal generator $\mathbf{L}$, for all functions $f \in \mathcal{A}$,

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

Since μ_ψ is invariant,

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

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 $t > 0$,

$$\Phi(t) = 2\int \Gamma(\mathbf{P_t}f)d\mu_\psi = 2\int_{\mathbb{R}^n} \Gamma(P_t f, P_t f)\, d\mu_\psi .$$

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 (us) 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

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

See the Bobkov et al. (2001) paper, for a re-interpretation of this argument.

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 on $\mathbb{R}^n$ (or on a manifold on a local chart) 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 [BG10].
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 $CD(0,n)$. On can observe that the Ornstein-Uhlenbeck semigroup does not satisfies a $CD(r,m)$ criterium for any $r, m > 0$.

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 functions $f \in \mathcal{A}$.
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 functions $f \in \mathcal{A}$.

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 is also a consequence of the Talagrand inequality described in the next section.

4 The logarithmic Sobolev and transport inequalities by transportation method

We will see how Brenier's Theorem can be used in this context to give a new proof of the logarithmic Sobolev inequality, the method is called mass transportation method.
We will illustrate this method for the Gaussian measure but it could be generalized for a large class of measures, this will be discussed later. The method come from [OV00, CE02] and has been generalized for many Euclidean inequalities as Sobolev and Gagliardo-Nirenberg inequalities, see [AGK04, CENV04, Naz06].

The Wasserstein distance between two probability measures μ and ν is defined by

$$W_2(\mu,\nu) = \left(\inf \int |x-y|^2 d\pi(x,y)\right)^{1/2}. \qquad (27)$$

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

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

Such probability is called a coupling of (μ,ν).

- Brenier's theorem says that that there exits an optimal deterministic coupling of (μ,ν) : there exists a convex map Φ satisfying

$$\int h(\nabla\Phi) d\nu = \int h d\mu,$$

for all bounded functions h. Moreover

$$W_2^2(\mu,\nu) = \int |\nabla\theta|^2 d\nu,$$

where $\theta(x) = \Phi(x) - \frac{1}{2}|x|^2$. This result has been proved by Brenier, $\nabla\Phi$ is called the Brenier map between ν and μ, see [Vil09].

We apply this result in the Gaussian case. Let f be a smooth and positive function such that $\int f d\gamma = 1$, Brenier's theorem implies that there exists a convex map Φ satisfying

$\mu = \gamma$
$d\nu = f\, d\gamma$

$$\int h(\nabla\Phi) f d\gamma = \int h d\gamma, \qquad (28)$$

for all bounded and measurable functions h. Moreover

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

where $\theta(x) = \Phi(x) - \frac{1}{2}|x|^2$.

If now Φ is a $\mathcal{C}^2(\mathbb{R}^n)$ function, then coming from (28), the Monge-Ampère equation holds : $f d\gamma$-a.e.

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

Determines the function $\theta(\cdot)$, and from it the function $\Phi(\cdot)$.

After taking the logarithm, we get

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

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

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

The integration by parts implies

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

which is the optimal logarithmic Sobolev inequality (9).

Hence we have proved, using Brenier's map, the logarithmic Sobolev inequality for the Gaussian measure with the optimal constant. As we can see in the proof, one has assumed that Φ is a $\mathcal{C}^2$ function. It can be obtained using Caffarelli's regularity theory : it needs another assumptions, f has to be smooth with a compact and convex support. We skip it for simplicity of the description of the method, many informations can be bound in [Vil09]

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

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

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

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

In that case the Monge-Ampère equation gives

$$\mathrm{det(Id+Hess(\theta))f(x+\nabla\theta(x))e^{-|x+\nabla\theta(x)|^2/2} = e^{-|x|^2/2}}. \tag{30}$$

Which implies

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

where $\mathbf{L}$ is the Ornstein-Uhlenbeck generator. Then

$$\begin{aligned}\mathbf{Ent}_\gamma(f) &= \int f \log f d\gamma \\ &= \int \log f (\nabla\Phi) d\gamma \\ &\geqslant \int -\mathbf{L}\theta d\gamma + \int \frac{1}{2}|\nabla\theta(x)|^2 d\gamma \\ &= \int \frac{1}{2}|\nabla\theta(x)|^2 d\gamma = \frac{1}{2}W_2^2(f d\gamma, d\gamma)\end{aligned}$$

We have proved that for all functions f such that $f d\gamma$ is a probability measure, one has

$$\boxed{W_2(f d\gamma, d\gamma) \leq \sqrt{2\mathbf{Ent}_\gamma(f)}.} \tag{31}$$

This inequality, called *Talagrand inequality for the Gaussian distribution (or $\mathcal{T}_2$ inequality)*, has been proved by Talagrand in [Tal96].
As for Poincaré and logarithmic Sobolev inequalities, we say that a probability measure μ satisfies a Talagrand inequality if there exists $C \geqslant 0$ such that,

$$W_2(f d\mu, d\mu) \leq \sqrt{C\mathbf{Ent}_\mu(f)}, \tag{32}$$

for all functions f such that $f d\mu$ 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 and 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 constant $1/(2\rho)$. This is an alternative proof of Theorem 3.3. Actually this method is not useful to obtain directly a Poincaré inequality.

Of course, as for Ornstein-Uhlenbeck semigroup, the mass transportation method gives also a Talagrand inequality (32) for the measure μ_ψ :

$$W_2(f d\mu_\psi, d\mu_\psi) \leq \sqrt{\frac{1}{\rho}\mathbf{Ent}_{\mu_\psi}(f)},$$

for all probability measure $f d\mu_\psi$.

In fact the general result holds,

Theorem 4.1 (Otto-Villani) *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 $C \geqslant 0$.
Then μ satisfies a Talagrand inequality

$$W_2(f d\mu, d\mu) \leq \sqrt{2C\mathbf{Ent}_\mu(f)},$$

for all probability measure $f d\mu$.

Log-Sobolev implies Talagrand (implies Poincaré)

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

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.

[BG10] F. Bolley and I. Gentil. Phi-entropy inequalities for diffusion semigroups. *J. Math. Pures Appl.*, 93(5):449–473, 2010.

[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. Springer-Verlag, Berlin, 2009.

Ceremade, UMR CNRS 7534
Université Paris-Dauphine

Justification of (17.a): From (16), we have

$$2\Gamma(f,f) = \mathbb{L}(f^2) - 2f.\mathbb{L}(f) = \Delta f^2 - \nabla\psi.\nabla f^2 - 2f(\Delta f - \nabla\psi.\nabla f).$$

Now $\Delta(f^2) = \sum_i \frac{\partial^2 f^2}{\partial x_i^2} = \sum_i \frac{\partial}{\partial x_i}\left(2f\frac{\partial f}{\partial x_i}\right) = 2f\,\Delta f + 2|\nabla f|^2$

whereas $\nabla\psi.\nabla f^2 = 2f.\nabla\psi.\nabla f$, thus: $\Gamma(f,f) = |\nabla f|^2$.

By the same token,

$$2\Gamma(f,g) = \mathbb{L}(fg) - f\,\mathbb{L}(g) - g\,\mathbb{L}(f)$$

$$= \Delta(fg) - \nabla\psi.\nabla(fg) - f(\Delta g - \nabla\psi.\nabla g) - g(\Delta f - \nabla\psi.\nabla f)$$

$$= 2\nabla f.\nabla g + f\Delta g + g\Delta f - \nabla\psi(f\nabla g + g\nabla f)$$

$$- f\Delta g + f\,\nabla\psi.\nabla g - g\Delta f + g\,\nabla\psi.\nabla f$$

$$= 2\nabla f.\nabla g.$$

Justification of (17.b): From (16), (17.a) we obtain

$$2\Gamma_2(f,f) = \mathbb{L}(\Gamma(f,f)) - 2\Gamma(f,\mathbb{L}f) = \mathbb{L}(|\nabla f|^2) - 2\nabla f.\nabla(\mathbb{L}f)$$

$$= \mathbb{L}\left(\sum_{j=1}^n\left(\frac{\partial f}{\partial x_j}\right)^2\right) - 2\sum_{j=1}^n \frac{\partial f}{\partial x_j}.\frac{\partial}{\partial x_j}(\mathbb{L}f)$$

Now

$$\mathbb{L}f = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} - \sum_{i=1}^n \frac{\partial\psi}{\partial x_i}\frac{\partial f}{\partial x_i},$$ therefore

$$\frac{\partial}{\partial x_j}(\mathbb{L}f) = \sum_{i=1}^n \frac{\partial^3 f}{\partial x_j \partial x_i^2} - \sum_{i=1}^n \frac{\partial^2 \psi}{\partial x_i \partial x_j} \frac{\partial f}{\partial x_i} - \sum_{i=1}^n \frac{\partial \psi}{\partial x_i} \frac{\partial^2 f}{\partial x_i \partial x_j}$$

and

$$-2 \cdot \sum_{j=1}^n \frac{\partial f}{\partial x_j} \cdot \frac{\partial}{\partial x_j}(\mathbb{L}f) = 2 \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 \psi}{\partial x_i \partial x_j} \cdot \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} +$$

$$+ 2 \sum_i \sum_j \frac{\partial^2 f}{\partial x_i \partial x_j} \cdot \frac{\partial \psi}{\partial x_i} \frac{\partial f}{\partial x_j} - 2 \sum_i \sum_j \frac{\partial f}{\partial x_j} \frac{\partial^3 f}{\partial x_j \partial x_i^2}.$$

On the other hand,

$$\mathbb{L}(\Delta f) = \sum_{i=1}^n \frac{\partial^2 (\Delta f)}{\partial x_i^2} - \sum_{i=1}^n \frac{\partial \psi}{\partial x_i} \frac{\partial}{\partial x_i}(\Delta f)$$

$$= \sum_{i=1}^n \sum_{j=1}^n \frac{\partial}{\partial x_i}\left(2 \frac{\partial f}{\partial x_j} \frac{\partial^2 f}{\partial x_i \partial x_j}\right) - \sum_{i=1}^n \sum_{j=1}^n \frac{\partial \psi}{\partial x_i} \left(2 \frac{\partial f}{\partial x_j} \frac{\partial^2 f}{\partial x_i \partial x_j}\right)$$

$$= 2 \sum_i \sum_j \frac{\partial^2 f}{\partial x_i \partial x_j} \frac{\partial^2 f}{\partial x_i \partial x_j} + 2 \sum_i \sum_j \frac{\partial f}{\partial x_j} \frac{\partial^3 f}{\partial x_i^2 \partial x_j}$$

$$- 2 \sum_i \sum_j \frac{\partial \psi}{\partial x_i} \frac{\partial f}{\partial x_j} \frac{\partial^2 f}{\partial x_i \partial x_j}$$

We deduce

$$\Gamma_2(f,f) = \sum_i \sum_j \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)^2 + \sum_i \sum_j \frac{\partial^2 \psi}{\partial x_i \partial x_j} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j}$$

$$= \left\| \mathcal{D}^2 f \right\|_{H.S}^2 + \left\langle \nabla f, \mathcal{D}^2 \psi \cdot \nabla f \right\rangle$$

<u>Proof of Lemma 3.5:</u> We have $\mathbb{L}(\varphi(f)) = \Delta(\varphi(f)) - \nabla\psi . \nabla\varphi(f)$

therefore:
$$\mathbb{L}(\varphi(f)) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\varphi'(f)\frac{\partial f}{\partial x_i}\right) - \varphi'(f)\, \nabla\psi . \nabla f$$
$$= \sum_{i=1}^{n}\left\{\varphi'(f)\frac{\partial^2 f}{\partial x_i^2} + \varphi''(f) . \left(\frac{\partial f}{\partial x_i}\right)^2\right\} - \varphi'(f)\, \nabla\psi . \nabla f$$
$$= \varphi'(f)\left(\Delta f - \nabla\psi . \nabla f\right) + \varphi''(f) . |\nabla f|^2$$
$$= \varphi'(f) . \mathbb{L}(f) + \varphi''(f) . |\nabla f|^2 .$$

On the other hand, from (17.a) we have
$$\Gamma(\log f) = |\nabla \log f|^2 = \frac{|\nabla f|^2}{f^2} = \frac{\Gamma(f)}{f^2} .$$

In particular, from the above:
$$\mathbb{L}(\log f) = \frac{\mathbb{L}(f)}{f} - \frac{\Gamma(f)}{f^2},$$
therefore with $g = \log f$ the expression (16) gives
$$\Gamma_2(\log f) = \Gamma_2(g) = \frac{1}{2}\mathbb{L}(\Gamma(g)) - \Gamma(g, \mathbb{L}g)$$
$$= \frac{1}{2}\mathbb{L}\left(\frac{\Gamma(f)}{f^2}\right) - \Gamma\left(\log f, \frac{\mathbb{L}(f)}{f} - \frac{\Gamma(f)}{f^2}\right) .$$

In order to prove (22), let us first cast it in the equivalent form

$$(22)' \qquad (\Gamma(f))^2 - f.\,\Gamma(f,\Gamma(f)) = f^4\,\Gamma_2(\log f) - f^2\,\Gamma_2(f).$$

According to (17.b), we have

$$\Gamma_2(\log f) = \sum_i\sum_j \left(D^2_{ij}\log f\right)^2 + \sum_i\sum_j D_i\log f.\,D^2_{ij}\psi.\,D_j\log f$$

But

$$D_j\log f = \frac{D_j f}{f}, \qquad D^2_{ij}\log f = \frac{D^2_{ij}f.\,f - D_i f.\,D_j f}{f^2}$$

and

$$\left(D^2_{ij}\log f\right)^2 = \frac{1}{f^4}\left(f^2.\,(D^2_{ij}f)^2 + (D_i f)^2(D_j f)^2 - 2f.\,D_i f.\,D^2_{ij}f.\,D_j f\right).$$

Therefore,

$$f^4\,\Gamma_2(\log f) = f^2.\sum_i\sum_j (D^2_{ij}f)^2 + \Big(\sum_i (D_i f)^2\Big)^2 - 2f.\sum_i\sum_j D_i f\,D^2_{ij}f\,D_j f$$

$$+ f^4.\sum_i\sum_j \frac{D_i f}{f}.\,D_{ij}\psi.\,\frac{D_j f}{f}$$

$$= f^2\left(\sum_i\sum_j (D^2_{ij}f)^2 + \sum_i\sum_j D_i f.\,D^2_{ij}\psi.\,D_j f\right)$$

$$+ |\nabla f|^4 - 2f.\,(\nabla f)'\,D^2 f\,(\nabla f)$$

We have shown, in conjunction with (17.a) and with (17.b) once again, that the right-hand side of (22)' equals

(22)'' $$f^4\,\Gamma_2(\log f) - f^2\,\Gamma_2(f) = (\Gamma(f))^2 - f\cdot 2(\nabla f)'\,\partial^2 f\,(\nabla f).$$

Now from (17.a) once again, we deduce

$$\Gamma(f) = |\nabla f|^2 = \sum_{j=1}^{n} (\partial_j f)^2, \quad \partial_i \Gamma(f) = 2\sum_{j=1}^{n} \partial_{ij}^2 f\cdot \partial_j f$$

as well as

$$\Gamma(f, \Gamma f) = \nabla f\cdot \nabla\Gamma(f) = \sum_{i=1}^{n} \partial_i f\cdot \partial_i \Gamma(f)$$

$$= 2\sum_i\sum_j \partial_i f\cdot \partial_{ij}^2 f\cdot \partial_j f\,,$$

and now (22)' follows from (22)''.

Justification for (5): We have

$$\int_{\mathbb{R}^n} f(x)\, \mathbb{L} g(x)\, \gamma(x)\, dx = \int_{\mathbb{R}^n} f(x)\,\big(\Delta g(x) - x' Dg(x)\big)\,\gamma(x)\, dx$$

$$= \sum_{i=1}^{n} \int_{\mathbb{R}^n} f(x)\,\big(D^2_{ii} g(x) - x_i\, D_i g(x)\big)\,\gamma(x)\, dx.$$

Now integrate by parts for the first term

$$\int_{\mathbb{R}^n} f(x)\, D^2_{ii} g(x)\, \gamma(x)\, dx = -\int_{\mathbb{R}^n} D_i g(x) \cdot D_i\big(f(x)\gamma(x)\big)\, dx$$

$$= -\int_{\mathbb{R}^n} D_i g(x)\,\big[\gamma(x) D_i f(x) + f(x)\, D_i \gamma(x)\big]\, dx$$

$$= \int_{\mathbb{R}^n} D_i g(x)\,\big[-\gamma(x)\, D_i f(x) + f(x) \cdot x_i\, \gamma(x)\big]\, dx.$$

In conjunction with the second term, we obtain from this

$$\int_{\mathbb{R}^n} f(x)\, \mathbb{L} g(x)\, \gamma(x)\, dx = -\sum_{i=1}^{n} \int_{\mathbb{R}^n} D_i f(x) \cdot D_i g(x)\, \gamma(x)\, dx$$

$$= -\int_{\mathbb{R}^n} \langle Df(x), Dg(x)\rangle\, \gamma(x)\, dx$$

We have used again the identity

$$D_i \gamma(x) = -x_i\, \gamma(x).$$

Proof of (8): POINCARÉ INEQUALITY

$$\int_{\mathbb{R}^n} f^2(x)\gamma(x)\,dx - \left(\int_{\mathbb{R}^n} f(x)\gamma(x)\,dx\right)^2 =: \mathrm{Var}_\gamma(f) \le \int_{\mathbb{R}^n} |Df(x)|^2 \gamma(x)\,dx$$

with equality, if the smooth function $f(\cdot)$ satisfies $Df(\cdot) = C$ for some constant vector $C \in \mathbb{R}^n$.

<u>Proof</u>: The crucial observation here is $\mathrm{Var}_\gamma(f) = V^f(0) - V^f(\infty)$, where

$$V^f(t) \triangleq \int_{\mathbb{R}^n} (P_t f(x))^2 \gamma(x)\,dx.$$

This is because $P_0 f(\cdot) = f(\cdot)$, $P_\infty f(\cdot) \equiv \int_{\mathbb{R}^n} f\gamma$. On the other hand,

$$\dot V^f(t) = 2\int_0^\infty \frac{\partial}{\partial t} P_t f(x) \cdot P_t f(x)\gamma(x)\,dx = 2\int_{\mathbb{R}^n} \mathbb{L} P_t f(x) \cdot P_t f(x)\gamma(x)\,dx$$

$$= -2\int_{\mathbb{R}^n} |DP_t f(x)|^2 \gamma(x)\,dx$$

from the basic identity (5).

Thus

$$\mathrm{Var}_\gamma(f) = V^f(0) - V^f(\infty) = -\int_0^\infty \dot V^f(t)\,dt = 2\int_0^\infty dt \int_{\mathbb{R}^n} |DP_t f(x)|^2 \gamma(x)\,dx$$

$$= 2\int_0^\infty e^{-2t}\,dt \int_{\mathbb{R}^n} |P_t Df(x)|^2 \gamma(x)\,dx.$$

We have used here the property

$$D_i P_t f(x) = e^{-t} \int_{\mathbb{R}^n} D_i f\left(x e^{-t} + y\sqrt{1-e^{-2t}}\right) \gamma(y)\,dy = e^{-t} P_t(D_i f)$$

or more compactly

$$D P_t f(x) = e^{-t} P_t(Df)(x) \qquad (6)$$

of the OU semigroup. Incidentally, for any norm $\|\cdot\|$ on $\mathbb{R}^n$ we obtain from this

$$\| D P_t f(x) \| \le e^{-t} \cdot P_t(\|Df\|)(x), \qquad (7)$$

as well as

$$\mathrm{Var}_\gamma(f) \le 2\int_0^\infty e^{-2t}\,dt \int_{\mathbb{R}^n} \left(P_t(|Df|)(x)\right)^2 \gamma(x)\,dx$$

with equality for $Df(\cdot) \equiv C$. But now with $g = |Df|$, we have by Cauchy-Schwarz

$$\left(P_t(|Df|)\right)^2 = \left(\mathbb{E}\, g(X_t)\right)^2 \le \mathbb{E}\left(g^2(X_t)\right) = P_t\left(|Df|^2\right),$$

with equality again for $Df(\cdot) \equiv C$. We deduce

$$\mathrm{Var}_\gamma(f) \le 2\int_0^\infty e^{-2t}\,dt \int_{\mathbb{R}^n} P_t\left(|Df|^2\right)(x)\, \gamma(x)\,dx$$

$$= 2\int_0^\infty e^{-2t}\,dt \int_{\mathbb{R}^n} |Df|^2(x)\, \gamma(x)\,dx = \int_{\mathbb{R}^n} |Df|^2(x)\gamma(x)$$

from the invariance property of $\gamma(\cdot)$.

Proof of (9): LOG - SOBOLEV INEQUALITY

For a positive, smooth function $f(\cdot)$ we introduce its ENTROPY

$$\mathrm{Ent}_\gamma(f) \triangleq \int_{\mathbb{R}^n} f(x) \log\left(\frac{f(x)}{\int_{\mathbb{R}^n} f(y)\gamma(y)\,dy} \right) \gamma(x)\,dx$$

$$= \int_{\mathbb{R}^n} f(x) \log f(x)\, d\gamma(x) - \int_{\mathbb{R}^n} f(x)\gamma(x)\,dx \,.\, \log\left(\int_{\mathbb{R}^n} f(z)\gamma(z)\,dz \right).$$

Introducing the r.v. $Z = f(\mathcal{U})$, where $\mathcal{U} \sim \gamma$, we have

$$\mathrm{Ent}_\gamma(f) = \mathbb{E}(Z \log Z) - \mathbb{E}(Z) . \log \mathbb{E}(Z) = \mathbb{E}[h(Z)] - h(\mathbb{E}(Z)) \geq 0$$

by Jensen, because the function $h(\xi) \triangleq \xi \log \xi$ is strictly convex.

We have for this positive quantity the

Logarithmic Sobolev Inequality

$$(9) \qquad \mathrm{Ent}_\gamma(f) \leq \frac{1}{2} \int_{\mathbb{R}^n} \frac{|Df(x)|^2}{f(x)} \gamma(x)\,dx$$

or equivalently:

$$\mathrm{Ent}_\gamma(f^2) \leq 2 \int_{\mathbb{R}^n} |Df(x)|^2 \gamma(x)\,dx.$$

The quantity

$$\boxed{I_\gamma(f) \triangleq \int_{\mathbb{R}^n} \frac{|Df(x)|^2}{f(x)} \gamma(x)\, dx}$$

is called the FISHER Information of the function $f(\cdot)$. And with

$$\underline{E^f(t) \triangleq \int_{\mathbb{R}^n} h(P_t f(x))\, \gamma(x)\, dx}$$

we have

$$E^f(0) = \int_{\mathbb{R}^n} h(f(x))\gamma(x)dx\,, \quad E^f(\infty) = h\Big(\int_{\mathbb{R}^n} f\gamma\, dx\Big) \quad \text{so} \quad \mathrm{Ent}_\gamma(f) = E^f(0) - E^f(\infty)$$

and

$$\dot{E}^f(t) = \int_{\mathbb{R}^n} \frac{\partial}{\partial t} P_t f(x) .\, h'(P_t f(x))\, \gamma(x)\, dx$$

$$= \int_{\mathbb{R}^n} \mathbb{L} g_t(x) \Big(\log g_t(x) + 1 \Big) \gamma(x)\,, \qquad g_t \triangleq P_t f$$

$$= \int_{\mathbb{R}^n} \mathbb{L} g_t(x) .\, \log g_t(x)\, \gamma(x)\, dx \qquad (\text{from } (5))$$

$$= -\int_{\mathbb{R}^n} D g_t(x) .\, D \log g_t(x)\, \gamma(x)\, dx \qquad (\text{from } (5) \text{ again})$$

$$= -\int_{\mathbb{R}^n} D P_t f(x) .\, \frac{D P_t f(x)}{P_t f(x)}\, \gamma(x)\, dx\,, \quad \text{so}$$

$$\frac{d}{dt} E^f(t) = -\int_{\mathbb{R}^n} \frac{|DP_t f(x)|^2}{P_t f(x)} \gamma(x)\, dx = -I_\gamma(P_t f).$$

In particular

$$\underline{\mathrm{Ent}_\gamma(f) - \mathrm{Ent}_\gamma(P_T f) = \int_0^T I_\gamma(P_t f)\, dt},$$

$$\mathrm{Ent}_\gamma(f) = E^f(0) - E^f(\infty) = \int_0^\infty I_\gamma(P_t f)\, dt$$

$$= \int_0^\infty \int_{\mathbb{R}^n} \frac{|DP_t f(x)|^2}{P_t f(x)} \gamma(x)\, dx\, dt$$

$$\leq \int_0^\infty e^{-2t}\, dt \int_{\mathbb{R}^n} \frac{\left(P_t(|Df|)(x)\right)^2}{P_t f(x)} \gamma(x)\, dx$$

by virtue of (7), as before. Now, Cauchy-Schwarz gives

$$P_t(|Df|) = P_t\left(\frac{|Df|}{\sqrt{f}} \sqrt{f}\right) \leq \sqrt{P_t\left(\frac{|Df|^2}{f}\right) P_t(f)}, \quad \text{or}$$

$$\frac{\left(P_t(|Df|)\right)^2}{P_t(f)} \leq P_t\left(\frac{|Df|^2}{f}\right), \quad \text{with equality if } \frac{Df(\cdot)}{f(\cdot)} \equiv C$$

and we arrive at

$$\mathrm{Ent}_\gamma(f) \leq \int_0^\infty e^{-2t}\, dt \int_{\mathbb{R}^n} P_t\left(\frac{|Df|^2}{f}\right) \gamma\, dx = \frac{1}{2} \int_{\mathbb{R}^n} \frac{|Df|^2}{f} \gamma\, dx$$

<u>Remark:</u> The Poincaré inequality is equivalent to

$$\text{(10)} \qquad \mathrm{Var}_\gamma (P_t f) \le e^{-2t}\, \mathrm{Var}_\gamma (f), \quad t \geq 0$$

and the Log-Sobolev inequality to

$$\text{(11)} \qquad \mathrm{Ent}_\gamma (P_t f) \le e^{-2t}\, \mathrm{Ent}_\gamma (f), \quad t \geq 0.$$

In other words, the speed of convergence to equilibrium, at least as far as the variance and entropy are concerned, for the OU semigroup is exponential.

• Indeed, the Log-Sobolev inequality amounts to $I_\gamma(f) \geq 2 . \mathrm{Ent}_\gamma(f)$; in conjunction with the (underlined) equation on the third line of page 25, this gives

$$\mathrm{Ent}_\gamma (P_T f) - \mathrm{Ent}_\gamma (f) = - \int_0^T I_\gamma (P_t f)\, dt \le -2 \int_0^T \mathrm{Ent}_\gamma (P_t f)\, dt .$$ The inequality (11) follows now as in Gronwall-Bellman.