Functional inequalities and applications

Luca Tamanini

Université Paris Ouest Nanterre La Défense
SISSA

July 5, 2015

Contents

These notes are based on the course *Inégalités fonctionnelles et applications* held by Ivan Gentil in Grenoble for the workshop *Inter'Actions en Mathématiques* (26 - 29 May 2015) and they are conceived as an introduction to the Γ-calculus developed by Bakry and Émery. As main reference for the course, we cite the book *Analysis and Geometry of Markov diffusion operators* by D. Bakry, I. Gentil and M. Ledoux, denoted by [BGL] from now on.

1 Introduction

Functional inequalities are a fundamental aspect of mathematics, thanks to their connection with different domains, as for instance:

- probability and statistics: Talagrand and Ledoux proved in this sense mass concentration results;
- geometry: Perel'man's proof of the Poincaré conjecture uses geometric functional inequalities;
- PDEs: most of the applications are in the parabolic case, but some of them may be found even in the elliptic and non-linear case.

In honor of John Forbes Nash (1928-2015) and as first remarkable example, let us cite the functional inequality named after him: it states that in $\mathbb{R}^n$ equipped with the Lebesgue measure there exists a constant $C > 0$ such that for any $f \in L^1(\mathbb{R}^n) \cap W^{1,2}(\mathbb{R}^n)$

$$\|f\|_2^{1+2/n} \le C \|\nabla f\|_2 \|f\|_1^{2/n}$$

This means that $L^1(\mathbb{R}^n) \cap W^{1,2}(\mathbb{R}^n)$ can be continuously embedded into $L^2(\mathbb{R}^n)$; although this inequality was proven in 1958, the optimal constant C has been determined only in 1992.

1.1 A basic example

As a motivating example, let us consider the n-dimensional Gaussian space, namely $\mathbb{R}^n$ equipped with the Gaussian measure γ defined by

$$\gamma(dx) := \frac{e^{-|x|^2/2}}{(2\pi)^{n/2}} dx \tag{1.1}$$

Let us also introduce a one-parameter semigroup $(\mathrm{P}_t)_{t \ge 0}$ as follows: for any measurable function $f : \mathbb{R}^n \to \mathbb{R}$ we set

$$\mathrm{P}_t f(x) := \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1 - e^{-2t}} y) d\gamma(y) \tag{1.2}$$

In a probabilistic form, this definition can be rewritten as

$$\mathrm{P}_t f(x) = \mathbb{E}[f(e^{-t}x + \sqrt{1 - e^{-2t}} Y]$$

with the assumption that $Y \sim \mathscr{N}(0, \mathrm{Id})$, but P_t can be also introduced via PDEs and SDEs. Indeed, for the first approach observe that if $f \in C_c^\infty(\mathbb{R}^n)$, then

$$\nabla \mathrm{P}_t f(x) = e^{-t} \int_{\mathbb{R}^n} \nabla f(e^{-t}x + \sqrt{1 - e^{-2t}} y) d\gamma(y) \tag{1.3}$$

$$\Delta \mathrm{P}_t f(x) = \nabla \cdot \nabla \mathrm{P}_t f(x) = e^{-2t} \int_{\mathbb{R}^n} \Delta f(e^{-t}x + \sqrt{1 - e^{-2t}} y) d\gamma(y)$$

so that

$$\begin{aligned}\frac{\partial}{\partial t}\mathrm{P}_t f(x) &= \int_{\mathbb{R}^n}\Big(-e^{-t}x+\frac{e^{-2t}}{\sqrt{1-e^{-2t}}}y\Big)\cdot\nabla f(e^{-t}x+\sqrt{1-e^{-2t}}y)d\gamma(y)\\ &= -x\cdot\nabla\mathrm{P}_t f(x)+\int_{\mathbb{R}^n}\frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\nabla f(e^{-t}x+\sqrt{1-e^{-2t}}y)\cdot\nabla\Big(-\frac{e^{-|y|^2/2}}{(2\pi)^{n/2}}\Big)dy\\ &= -x\cdot\nabla\mathrm{P}_t f(x)+e^{-2t}\int_{\mathbb{R}^n}\Delta f(e^{-t}x+\sqrt{1-e^{-2t}}y)d\gamma(y)\\ &= -x\cdot\nabla\mathrm{P}_t f(x)+\Delta\mathrm{P}_t f(x)\end{aligned}$$

This means that $u(t,x):=\mathrm{P}_t f(x)$ is the only solution to the PDE

$$\begin{cases}\dfrac{\partial u}{\partial t}=\mathrm{L}\,u\\ u|_{t=0}=f\end{cases}\tag{1.4}$$

where the operator L is given by $\mathrm{L}=\Delta-x\cdot\nabla$. For an SDE representation, it is sufficient to observe that

$$\mathrm{P}_t f(x)=\mathbb{E}[f(X_t)\,|\,X_0=x]$$

where the process $(X_t)_{t\geq 0}$ is the unique solution to the following SDE:

$$\begin{cases}dX_t=\sqrt{2}dB_t-X_tdt\\ X_0=x\end{cases}$$

In the literature, P_t is known as the *Ornstein-Uhlenbeck semigroup* (named after the Dutch physicists Leonard Ornstein and George Eugene Uhlenbeck) and the solution X_t to the SDE above is called the *Ornstein-Uhlenbeck process*. Now let us point out some interesting properties enjoyed by the Gaussian measure and, mostly, by P_t.

(a) For any measurable function $f:\mathbb{R}^n\to\mathbb{R}$, $f\in L^p(\mathbb{R}^n,\gamma)$ for any $p\in[1,+\infty]$.

(b) If $f\geq 0$, then $\mathrm{P}_t f\geq 0$ as well.

(c) $(\mathrm{P}_t)_{t\geq 0}$ is a family of linear operators such that $\mathrm{P}_t 1=1$, where 1 denotes the constant function identically equal to 1.

Properties (b) and (c) simply come from the fact that γ is a probability measure, but it is worth expressing them in terms of P_t for the forthcoming theory, where P_t will be completely abstract. They are also strongly related to the maximum principle enjoyed by (1.4). Thanks to these properties, the Ornstein-Uhlenbeck semigroup satisfies Jensen's inequality, in the sense that if $\varphi:\mathbb{R}\to\mathbb{R}$ is a convex function, then $\varphi(\mathrm{P}_t f)\leq\mathrm{P}_t(\varphi(f))$. In particular, taking into account the fact that by (1.3) we have $\nabla\mathrm{P}_t f=e^{-t}\mathrm{P}_t(\nabla f)$ and by choosing $\varphi(\cdot)=|\cdot|$, we deduce that

$$|\nabla\mathrm{P}_t f|(x)\leq e^{-t}\mathrm{P}_t(|\nabla f|)(x)\tag{1.5}$$

for any $x\in\mathbb{R}^n$.

(d) $(\mathrm{P}_t)_{t\geq 0}$ is a one-parameter semigroup, i.e. $\mathrm{P}_0=\mathrm{Id}$ and for any $s,t\geq 0$ it holds $\mathrm{P}_{t+s}=\mathrm{P}_t\circ\mathrm{P}_s$.

(e) $(\mathrm{P}_t)_{t\geq 0}$ is a contractive semigroup on $L^\infty(\mathbb{R}^n,dx)$, that is $\|\mathrm{P}_t f\|_{L^\infty}\leq\|f\|_{L^\infty}$.

(f) The following facts hold

$$\lim_{t\downarrow 0}\mathrm{P}_t f=f,\qquad\lim_{t\to+\infty}\mathrm{P}_t f=\int_{\mathbb{R}^n}f\,d\gamma$$

where the convergence has to be meant in $L^2(\mathbb{R}^n,\gamma)$.

The fact that $(\mathrm{P}_t)_{t\geq 0}$ is really a semigroup relies on the uniqueness of the solutions to (1.4). By the semigroup property, any regularity degree at $t = 0$ is authomatically extended to the whole half-line $\mathbb{R}_+$: this means that the continuity of P_t at $t = 0$, given by property (f), holds for any $t \geq 0$ and thus $(\mathrm{P}_t)_{t\geq 0}$ is a continuous semigroup. Furthermore, (f) also tells us that the Ornstein-Uhlenbeck semigroup can be regarded as an interpolation between f and its average w.r.t. γ for any function f.

(g) As shown by (1.4), $\mathrm{L} := \Delta - x \cdot \nabla$ is the generator of the Ornstein-Uhlenbeck semigroup and it commutes with P_t.

(h) For any $f, g \in C_c^\infty(\mathbb{R}^n)$, the following integration by parts formula (IP-O.U. for short) holds

$$\int_{\mathbb{R}^n} f\,\mathrm{L}\,g\,d\gamma = -\int_{\mathbb{R}^n} \nabla f \cdot \nabla g\,d\gamma \tag{IP-O.U.}$$

(i) For any $f \in C_c^\infty(\mathbb{R}^n)$ it holds

$$\int_{\mathbb{R}^n} \mathrm{P}_t\,f\,d\gamma = \int_{\mathbb{R}^n} f\,d\gamma$$

that is, no mass dissipation occurs and this tells us that L admits γ as invariant measure.

Observe that the Laplacian is self-adjoint w.r.t. the Lebesgue measure, but in the Gaussian case it is not; indeed, by (IP-O.U.) formula we have that

$$\int_{\mathbb{R}^n} f\,\mathrm{L}\,g\,d\gamma = -\int_{\mathbb{R}^n} \nabla f \cdot \nabla g\,d\gamma = \int_{\mathbb{R}^n} g\,\mathrm{L}\,f\,d\gamma$$

and this shows that Δ has to be modified with a drift term $b(x) = -x$ in order to get a self-adjoint operator. Moreover, property (i) is a straightforward consequence of (h), because by (IP-O.U.) formula we see that

$$\int_{\mathbb{R}^n} \mathrm{L}\,f\,d\gamma = 0 \tag{1.6}$$

whence the mass conservation principle stated at (i). After this preliminary part, let us prove two fundamental and well-known results: Hölder and log-Sobolev inequalities. As the reader will immediately find out, the adopted technique is conceptually easy but very powerful.

1.1.1 Hölder inequality

Our aim is to prove that for any functions $f, g \geq 0$ and for any $\vartheta \in [0, 1]$ it holds

$$\int_{\mathbb{R}^n} f^\vartheta g^{1-\vartheta} d\gamma \leq \left(\int_{\mathbb{R}^n} f\,d\gamma\right)^\vartheta \left(\int_{\mathbb{R}^n} g\,d\gamma\right)^{1-\vartheta}$$

If we are able to show that for any $t \in [0, +\infty[$ the following local functional inequality is satisfied

$$\mathrm{P}_t(f^\vartheta g^{1-\vartheta}) \leq (\mathrm{P}_t\,f)^\vartheta (\mathrm{P}_t\,g)^{1-\vartheta} \tag{1.7}$$

(the adjective *local* is due to the fact that the inequality holds for any x), then Hölder inequality will follow, because it will be sufficient to pass to the limit as $t \to +\infty$ and apply property (f). Thus, let us prove (1.7) and, as a first step, for t fixed and $s \in [0, t]$ let us define

$$\Lambda(s) := \mathrm{P}_s((\mathrm{P}_{t-s}\,f)^\vartheta (\mathrm{P}_{t-s}\,g)^{1-\vartheta})$$

Notice that

$$\Lambda(0) = (\mathrm{P}_t\,f)^\vartheta (\mathrm{P}_t\,g)^{1-\vartheta}, \quad \Lambda(t) = \mathrm{P}_t(f^\vartheta g^{1-\vartheta})$$

so that if we prove that Λ is decreasing, the conclusion will follow. For sake of simplicity, let us set $F := \log \mathrm{P}_{t-s} f$ and $G := \log \mathrm{P}_{t-s} g$ and observe that Λ can be rewritten as $\Lambda(s) = \mathrm{P}_s(e^{\vartheta F + (1-\vartheta)G})$. Then, notice that

$$\frac{\partial F}{\partial s} = -\frac{\mathrm{L}\,\mathrm{P}_{t-s} f}{\mathrm{P}_{t-s} f} = -\frac{\mathrm{L}\, e^F}{e^F} = -\Delta F - |\nabla F|^2 + x \cdot \nabla F$$

and an analogous computation holds for G too, hence for any linear combination of F and G. With this remark and with the further notation $H := \vartheta F + (1-\vartheta)G$ we are finally ready to compute $\Lambda'(s)$.

$$\begin{aligned}
\Lambda'(s) &= \mathrm{L}(\mathrm{P}_s(e^H)) + \mathrm{P}_s(\partial_s e^H) = \mathrm{P}_s(\mathrm{L}(e^H)) + \mathrm{P}_s(\partial_s e^H) = \mathrm{P}_s(\mathrm{L}(e^H) + \partial_s e^H) \\
&= \mathrm{P}_s((\Delta H + |\nabla H|^2 - x \cdot \nabla H - \vartheta \Delta F - \vartheta |\nabla F|^2 + \vartheta x \cdot \nabla F - (1-\vartheta)\Delta G \\
&\quad -(1-\vartheta)|\nabla G|^2 + (1-\vartheta) x \cdot \nabla G) e^H) \\
&= \mathrm{P}_s((|\vartheta \nabla F + (1-\vartheta)\nabla G|^2 - \vartheta |\nabla F|^2 - (1-\vartheta)|\nabla G|^2) e^H) \le 0
\end{aligned}$$

For the second identity we simply used the fact that L commutes with the semigroup, while the final inequality is due to the convexity of $|\cdot|^2$ together with the fact that P_s preserves the sign, as stated in property (b). Therefore $\Lambda(t) \le \Lambda(0)$, whence Hölder inequality for the Gaussian measure and, as straightforward byproduct, Cauchy-Schwarz inequality as well.

It is worth stressing that, in the computations we did, we did not use the explicit definition of P_t, but some of the properties (a)-(i). Hence this approach can be fairly well generalized to an abstract setting, as we will see in the next section.

1.1.2 Log-Sobolev inequality

In order to state the functional inequality we are interested in, let us first define the relative entropy functional. To this aim, let μ be a measure on $\mathbb{R}^n$ and let $\varphi \ge 0$ be any μ-integrable function: then the entropy of $\varphi\mu$ relative to μ is defined as

$$\mathrm{Ent}_\mu(\varphi) := \int_{\mathbb{R}^n} \varphi \log\left(\frac{\varphi}{\int_{\mathbb{R}^n} \varphi \, d\mu}\right) d\mu$$

the integral being possibly infinite; it is easy to see that $\mathrm{Ent}_\mu(\varphi) = 0$ if and only if φ is constant, provided that μ is a probability measure. The result we are looking at is the following.

Theorem 1.1 (Gross, 1975)**.** *For any function $f \in W^{1,2}(\mathbb{R}^n)$ it holds*

$$\mathrm{Ent}_\gamma(f^2) \le 2 \int_{\mathbb{R}^n} |\nabla f|^2 \, d\gamma \tag{1.8}$$

The constant 2 appearing at the right-hand side in the inequality above is optimal.

From an historical point of view, in 1938 Sobolev proved his embedding theorem; it states that, on $\mathbb{R}^n$ equipped with the Lebesgue measure dx, $W^{1,p}(\mathbb{R}^n) \subset L^{p^*}(\mathbb{R}^n)$ with continuous embedding, where p^* is the Sobolev conjugate of p, given by

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}$$

Namely, for $p = 2$, we have

$$\|f\|_{\frac{2n}{n-2}} \le C \|\nabla f\|_2$$

for a suitable constant $C > 0$ independent of f. However, such an inequality is no longer true on $(\mathbb{R}^n, \gamma)$ and it was only in 1975 that Gross proved Theorem 1.1; such result tells us that if

$\nabla f \in L^2(\gamma)$, then $f \in L^2 \log L(\gamma)$ and we can not expect higher regularity, since for any $a \in \mathbb{R}^n$ the map $x \mapsto e^{a\cdot x}$ is an extremal function, in the sense that for it equality is attained at (1.8) and the computation can be easily performed. After this digression, let us set $g := f^2$ and let us also observe that

$$\mathrm{P}_t(|\nabla g|)^2 = \mathrm{P}_t\left(\frac{|\nabla g|}{\sqrt{g}}\sqrt{g}\right)^2 \leq \mathrm{P}_t\left(\frac{|\nabla g|^2}{g}\right)\mathrm{P}_t(g) \tag{1.9}$$

where the second passage comes from the Cauchy-Schwarz inequality we proved in the previous subsection. With these premises, the proof of (1.8) goes as follows.

$$\begin{aligned}
\mathrm{Ent}_\gamma(f^2) &= -\int_0^{+\infty} \frac{d}{dt}\int_{\mathbb{R}^n} \mathrm{P}_t(g)\log \mathrm{P}_t(g)\,d\gamma dt \\
&= -\int_0^{+\infty}\int_{\mathbb{R}^n} \Big(\mathrm{L}\,\mathrm{P}_t(g)\log \mathrm{P}_t(g) + \mathrm{L}\,\mathrm{P}_t(g)\Big)d\gamma dt \\
&\overset{(1.6)}{=} -\int_0^{+\infty}\int_{\mathbb{R}^n} \mathrm{L}\,\mathrm{P}_t(g)\log \mathrm{P}_t(g)\,d\gamma dt \overset{(IP-O.U.)}{=} \int_0^{+\infty}\int_{\mathbb{R}^n} \nabla \mathrm{P}_t(g)\cdot\nabla \log \mathrm{P}_t(g)\,d\gamma dt \\
&= \int_0^{+\infty}\int_{\mathbb{R}^n} \frac{|\nabla \mathrm{P}_t(g)|^2}{\mathrm{P}_t(g)}d\gamma dt \overset{(1.5)}{\leq} \int_0^{+\infty}\int_{\mathbb{R}^n} e^{-2t}\frac{\mathrm{P}_t(|\nabla g|)^2}{\mathrm{P}_t(g)}d\gamma dt \\
&\overset{(1.9)}{\leq} \int_0^{+\infty} e^{-2t}\int_{\mathbb{R}^n} \mathrm{P}_t\left(\frac{|\nabla g|^2}{g}\right)d\gamma dt \overset{(i)}{=} \int_0^{+\infty} e^{-2t}\int_{\mathbb{R}^n} \frac{|\nabla g|^2}{g}d\gamma dt = \frac{1}{2}\int_{\mathbb{R}^n}\frac{|\nabla g|^2}{g}d\gamma \\
&= 2\int_{\mathbb{R}^n}|\nabla f|^2 d\gamma
\end{aligned}$$

The log-Sobolev inequality for the Gaussian measure is thus proved and, again, we did not use the explicit definition of the Ornstein-Uhlenbeck semigroup.

On $(\mathbb{R}^n, \gamma)$ it turns out that this inequality is equivalent to 2-Poincaré's one for γ, namely

$$\int_{\mathbb{R}^n} f^2\,d\gamma - \left(\int_{\mathbb{R}^n} f\,d\gamma\right)^2 \leq \int_{\mathbb{R}^n}|\nabla f|^2\,d\gamma \tag{1.10}$$

and also to the following asymptotic behaviour

$$\mathrm{Ent}_\gamma(\mathrm{P}_t\,\varphi) \leq e^{-t/2}\,\mathrm{Ent}_\gamma(\varphi), \quad \forall t \geq 0,\ \forall \varphi \geq 0$$

This means that the log-Sobolev inequality provides us with an exponential decay in relative entropy; more precisely, $\mathrm{P}_t\,\varphi$ converges exponentially fast in relative entropy to the equilibrium $\int_{\mathbb{R}^n}\varphi\,d\gamma$. This is perfectly coherent with the spectral properties of the Poincaré inequality, which we will briefly describe for sake of information. From an historical point of view, Poincaré proved the inequality named after him in 1880 for the Lebesgue measure and it was soon understood that this result was strongly connected to the spectral properties of the Laplacian; for the Gaussian measure, the same relationship holds: the left-hand side in (1.10), known in the literature as *spectral gap*, is the smallest non-zero eigenvalue of L and provides us with the (exponential) rate of convergence of P_t to the equilibrium in the $L^2(\gamma)$-norm.

2 Semigroups and the Bakry-Émery criterion

After the example of the Ornstein-Uhlenbeck semigroup in $\mathbb{R}^n$, relative to the Gaussian measure γ, one could study the heat semigroup on $(\mathbb{R}^n, dx)$ or on the n-dimensional unit sphere $\mathbb{S}^n$ equipped with the Haar measure (for $n \geq 3$ the situation is not well-understood yet), but the reader should guess that this approach is not particularly satisfying and not general at all. For this reason, we will undertake a completely different path.

2.1 The importance of being Markov

As a first ingredient, let $(E, \mathscr{F}, \mu)$ be a triple, where E is a Polish space, $\mathscr{F}$ a σ-algebra and μ a σ-finite measure. In this framework, let us consider a family $(\mathrm{P}_t)_{t\geq 0}$ of operators defined on some set of real-valued measurable functions on $(E, \mathscr{F})$ and let us point out a preliminary list of conditions:

(H1) for every $t \geq 0$, $\mathrm{P}_t : L^\infty(E, \mu) \to L^\infty(E, \mu)$;

(H2) $\mathrm{P}_t\, 1 = 1$, where 1 denotes the constant function identically equal to 1;

(H3) $\mathrm{P}_0 = \mathrm{Id}$;

(H4) if $f \geq 0$, then $\mathrm{P}_t\, f \geq 0$;

(H5) for every $t, s \geq 0$, $\mathrm{P}_{t+s} = \mathrm{P}_t \circ \mathrm{P}_s$;

(H6) μ is an invariant measure, that is for every bounded non-negative measurable function f and for every $t \geq 0$ it holds

$$\int_E \mathrm{P}_t\, f \, d\mu = \int_E f \, d\mu \tag{2.1}$$

In other words, $\mathrm{P}_t^*\, \mu = \mu$ for every $t \geq 0$;

(H7) for every $f \in L^2(E, \mu)$, $\mathrm{P}_t\, f \to f$ in $L^2(E, \mu)$ as $t \to 0$

While (H1)-(H6) are in some sense natural, the last assumption is the most arbitrary, because we choose a precise type of continuity.

Definition 2.1. *A family of operators* $(\mathrm{P}_t)_{t\geq 0}$ *defined on* $L^\infty(E, \mu)$ *satisfying the conditions (H1)-(H7) is called a* Markov semigroup of operators.

It is worth saying that hypotheses (H1)-(H7) entail a certain number of interesting properties. First of all, P_t can be extended to a bounded operator on $L^p(E, \mu)$ for any $p \geq 1$ and this is due to the following fact: (H2) and (H4) determine the Markov behaviour of P_t and allow us to invoke Jensen's inequality, so that for every convex function $\varphi : \mathbb{R} \to \mathbb{R}$ we have $\varphi(\mathrm{P}_t\, f) \leq \mathrm{P}_t(\varphi(f))$; this immediatley implies that $|\mathrm{P}_t\, f|^p \leq \mathrm{P}_t(|f|^p)$ for any $p \in [1, +\infty[$, whence

$$\| \mathrm{P}_t\, f\|_p^p = \int_E |\mathrm{P}_t\, f|^p \, d\mu \leq \int_E \mathrm{P}_t(|f|^p) \, d\mu \overset{(H6)}{=} \int_E |f|^p \, d\mu = \|f\|_p^p$$

This inequality passes to the limit as $p \to +\infty$ and so we can say that P_t is a contraction on $L^p(E, \mu)$ for any $t \geq 0$, whence the possibility to extend it. In particular, the invariance definition (2.1) extends to any $f \in L^1(E, \mu)$. Secondly, P_t can be represented by probability kernels corresponding to the transition probabilities of the associated Markov process. This amounts to say that for every bounded measurable function $f : E \to \mathbb{R}$ it holds

$$\mathrm{P}_t\, f(x) = \int_E f(y) p_t(x, dy) \tag{2.2}$$

for any $t \geq 0$ and for μ-a.e. $x \in E$, where $p_t(x, dy)$ is a probability kernel for any $t \geq 0$, that is $p_t(x, \cdot)$ is a probability measure for any $x \in E$ and $x \mapsto p_t(x, A)$ is measurable for any $A \in \mathscr{F}$. In general, p_t need not be absolutely continuous w.r.t. μ. Thanks to this representation formula, relying on abstract measure theory results and following from (H2), (H4) and from the fact that P_t is continuous on $L^1(E, \mu)$, the operators P_t can be extended to measurable functions $f : E \to [0, +\infty]$. It is worth stressing that the conditions we have just cited for the existence of probability kernels are necessary, in order to avoid the following problem: one could be tempted to define

$$p_t(x, A) = \int_A p_t(x, dy) := \mathrm{P}_t(\mathbb{1}_A)$$

for any $A \in \mathscr{F}$, but unfortunately if P_t does not fulfill (H2), (H4) and the continuity on $L^1(E,\mu)$, this definition is bad, because $x \mapsto p_t(x,A)$ needs not be measurable. Thirdly, the set of assumptions we formulated ensures the application of the Hille-Yosida theorem, so that there exist a dense linear subspace $\mathscr{D}(\mathrm{L}) \subset L^2(E,\mu)$ and a linear operator $\mathrm{L} : \mathscr{D}(\mathrm{L}) \to L^2(E,\mu)$ such that for any $f \in \mathscr{D}(\mathrm{L})$

$$\mathrm{L}\, f = \lim_{t \to 0} \frac{\mathrm{P}_t\, f - f}{t}$$

where the limit has to be understood in the $L^2(E,\mu)$ sense. The operator L is called the *infinitesimal generator* of the semigroup $(\mathrm{P}_t)_{t \geq 0}$; in the general theory of operator semigroups, the existence of a bounded infinitesimal generator allows to write

$$\mathrm{P}_t = e^{t\,\mathrm{L}} := \sum_{k=0}^{\infty} \frac{L^k}{k!} t^k$$

and the exponential notation is compatible with the notation for matrix exponentials, in the case E is a finite-dimensional space and thus P_t can be represented as a matrix. Moreover, notice that the semigroup property and the linearity of each P_t imply

$$\frac{\mathrm{P}_{t+s} - \mathrm{P}_t}{s} = \mathrm{P}_t \circ \frac{\mathrm{P}_s - \mathrm{Id}}{s} = \frac{\mathrm{P}_s - \mathrm{Id}}{s} \circ \mathrm{P}_t$$

whence

$$\frac{d}{dt}\, \mathrm{P}_t\, f = \mathrm{L}\, \mathrm{P}_t\, f = \mathrm{P}_t(\mathrm{L}\, f)$$

for any $t \geq 0$, $f \in \mathscr{D}(\mathrm{L})$ in the $L^2(E,\mu)$ sense: this means that P_t and L commute. Finally, even if obvious, it is worth saying that constant functions belong to the kernel of L because of (H2).

In view of the forthcoming definition, let $\mathscr{A} \subset \mathscr{D}(\mathrm{L})$ be a suitable algebra of functions dense in $L^p(E,\mu)$ for $p \in [1, +\infty[$; in a given context, the choice of $\mathscr{A}$ will be mostly natural: for instance, in the case of a manifold $\mathscr{A}$ will be the algebra of smooth compactly supported functions. In particular, this ensures us first that if $f, g \in \mathscr{A}$, then $fg \in \mathscr{A}$ as well; secondly, even if we do not know what $\mathscr{D}(\mathrm{L})$ is, by density it is sufficient to work with functions belonging to $\mathscr{A}$.

Definition 2.2. *Let $(\mathrm{P}_t)_{t \geq 0}$ be a Markov semigroup on $(E, \mathscr{F}, \mu)$. The bilinear map*

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

defined for every $f, g \in \mathscr{A}$ is called the carré du champ operator *of the Markov generator* L. *The quadratic form associated to Γ will be still denoted in the same way, i.e. $\Gamma(f) := \Gamma(f,f)$.*

As a first remark, Γ is positive on $\mathscr{A}$, in the sense that

$$\Gamma(f) \geq 0, \quad \forall f \in \mathscr{A} \tag{2.3}$$

In fact, by Jensen's inequality we know that $\mathrm{P}_t(f^2) \geq \mathrm{P}_t(f)^2$ for every $f \in \mathscr{A}$ and in $t = 0$ equality actually holds; thus, by differentiation at $t = 0$ the inequality is preserved and this means that $\mathrm{L}(f^2) \geq 2f\,\mathrm{L}\,f$, whence (2.3) by definition of Γ. But in order to work properly with the carré du champ operator we need to assume some further conditions, which will be added to the previous (H1)-(H7):

(H8) the semigroup $(\mathrm{P}_t)_{t \geq 0}$ is symmetric w.r.t. μ or, equivalently, μ is reversible for $(\mathrm{P}_t)_{t \geq 0}$; this means that for every $f, g \in L^2(E,\mu)$ and for every $t \geq 0$ it holds

$$\int_E f\,\mathrm{P}_t\, g\, d\mu = \int_E g\,\mathrm{P}_t\, f\, d\mu \tag{2.4}$$

(H9) for every $f \in \mathscr{A}$, if $\Gamma(f) = 0$, then f is μ-a.e. constant;

(H10) P_t is a diffusion, namely for any twice differentiable function $\phi : \mathbb{R}^k \to \mathbb{R}$ and for every $f_1, ..., f_k \in \mathscr{A}$ the following identity is satisfied:

$$\mathrm{L}(\phi(f_1, ..., f_k)) = \sum_{i=1}^{k} \partial_i \phi(f_1, ..., f_k) \, \mathrm{L} \, f_i + \sum_{i,j=1}^{k} \partial_{ij} \phi(f_1, ..., f_k) \Gamma(f_i, f_j) \tag{2.5}$$

Observe that (H8) can be expressed in terms of the generator of the process, since if we differentiate (2.4) at $t = 0$ then we get

$$\int_E f \, \mathrm{L} \, g \, d\mu = \int_E g \, \mathrm{L} \, f \, d\mu$$

for every $f, g \in \mathscr{A}$ and this means that L is a symmetric operator, although unbounded on $L^2(E, \mu)$. As for the Ornstein-Uhlenbeck semigroup, this identity immediately implies that

$$\int_E \mathrm{L} \, f \, d\mu = 0 \tag{2.6}$$

for any $f \in \mathscr{A}$ (cf. (1.6)). In addition, (H8) also entails an analogue of the integration by parts formula (IP-O.U.) we have already seen for the Ornstein-Uhlenbeck semigroup, namely

$$\int_E f \, \mathrm{L} \, g \, d\mu = -\int_E \Gamma(f, g) \, d\mu, \quad \forall f, g, \in \mathscr{A} \tag{IP}$$

and also in this case we will refer to it as *integration by parts formula* (IP for short). As regards (H9), it is worth stressing that at a first glance this assumption has a geometric meaning, in the sense that it forces E to be connected (if it were not the case, then any locally constant function f would satisfy $\Gamma(f) = 0$, but it would not necessarily satisfy the fact to be constant), but that is not all what we can say. Indeed, it has also ergodic consequences, because for any $f \in L^2(E, \mu)$

- if $\mu(E) < +\infty$, then
$$\lim_{t \to +\infty} \mathrm{P}_t \, f = \int_E f \, d\mu$$
the limit being understood in the $L^2(E, \mu)$ sense;
- if $\mu(E) = +\infty$, then $\mathrm{P}_t \, f \to 0$ in $L^2(E, \mu)$.

As a last remark, (H9) can be clearly stated as an *if and only if* assumption, because for a μ-a.e. constant function f we know that $\mathrm{L} \, f = 0$, whence $\Gamma(f) = 0$ by definition of Γ. About (H10) notice that this hypothesis rules out jump processes, such as Poisson and Lévy ones; nevertheless, it is still possible to develop a (different) Γ-calculus on discrete spaces, although the theory is far from being satisfactory. In addition, (H10) has important structural consequences, because it tells us that L is a second order differential operator if $E = \mathbb{R}^n$. In fact, notice that if we take $f_1, ..., f_n$ as the coordinate functions $x_1, ..., x_n$, then (2.5) reads as

$$\mathrm{L} \, \phi = \sum_{i=1}^{n} \partial_i \phi \, \mathrm{L} \, x_i + \sum_{i,j=1}^{n} \partial_{ij} \phi \Gamma(x_i, x_j)$$

and this means that L can be expressed as

$$\mathrm{L} = \sum_{i=1}^{n} b_i(x) \partial_i + \sum_{i,j=1}^{n} \sigma_{ij}(x) \partial_{ij}$$

for suitable functions b_i, σ_{ij}. As a final remark, a particular case of (H10) is given by $k = 1$, so that $\phi : \mathbb{R} \to \mathbb{R}$ and (2.5) becomes

$$\mathrm{L}(\phi(f)) = \phi'(f)\,\mathrm{L}\,f + \phi''(f)\Gamma(f) \tag{2.7}$$

In turn, this identity implies that

$$\Gamma(\phi(f), g) = \phi'(f)\Gamma(f, g), \quad \Gamma(\phi(f)) = (\phi'(f))^2\Gamma(f) \tag{2.8}$$

Both (2.7) and (2.8) will be used several times in the computations we are going to perform.

Definition 2.3. *Assume that conditions (H1)-(H10) are satisfied. Then* (E, Γ, μ) *is called a* Markov triple.

After we have single out the notion of Markov triple and the tools we will work with, let us present a list of concrete examples, in order to make clear how the carré du champ operator may look like and how important the role played by the measure μ is.

Example 2.4. Let $E = \mathbb{R}^n$, let μ be the Gaussian measure given at (1.1) and let us define the carré du champ operator by $\Gamma(f, g) := \nabla f \cdot \nabla g$, so that the associated quadratic form is $\Gamma(f) = |\nabla f|^2$. In this case, the semigroup associated to the triple (E, Γ, μ) is the Ornstein-Uhlenbeck one, defined at (1.2). As already seen, the generator of the semigroup is given by $\mathrm{L} = \Delta - x \cdot \nabla$ and since $\mu = \gamma$ is a probability measure, then the system is ergodic and

$$\lim_{t\to+\infty} \mathrm{P}_t\,f = \int_{\mathbb{R}^n} f\,d\gamma$$

for any $f \in L^2(\mathbb{R}^n, \gamma)$. $\lozenge$

Example 2.5. Let $E = \mathbb{R}^n$, μ be the Lebesgue measure and let the carré du champ operator be defined as before. In this case, P_t is the heat semigroup, given by

$$\mathrm{P}_t\,f(x) := \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} f(y) e^{-\frac{|x-y|^2}{4t}}\,dy$$

whose generator is the well-known Laplace operator Δ. $\lozenge$

Example 2.6. A slightly more general situation was proposed by Kolmogorov. Still on $E = \mathbb{R}^n$, let us consider $\mu = e^{-V}dx$, where dx denotes the Lebesgue measure and $V : \mathbb{R}^n \to \mathbb{R}$ is a sufficiently nice function; if we define again Γ by $\Gamma(f, g) := \nabla f \cdot \nabla g$, then the generator of the resulting semigroup is $\mathrm{L} = \Delta - \nabla V \cdot \nabla$: thus, letting Γ unchanged but modifying the measure μ produces a drift term in the generator of the underlying process. As a final remark, recall that the PDE associated to our framework is (1.4); hence, a lack of regularity on V may cause a blow up of the solution in finite time, so that it is not possible to define a semigroup. In order to avoid this obstacle, one can assume for instance that V is coercive, namely

$$\lim_{|x|\to+\infty} V(x) = +\infty$$

and in addition $|\nabla V|^2 - \Delta V$ is also lower bounded. $\lozenge$

Example 2.7. For a more general case, let $E = \mathbb{R}^n$ and let us consider the following SDE

$$dX_t = \sqrt{2}\sigma(X_t)dB_t + b(X_t)dt \tag{2.9}$$

where B denotes the standard Brownian motion with generator Δ. The solutions of the SDE are clearly Markovian, because the coefficients σ_{ij} and b_i do not depend on the whole trajectory, but only on X_t. Hence, we can talk about the Markov generator of the solution $(X_t)_{t\geq 0}$ to (2.9), which is given by

$$\mathrm{L}\, f(x) = \sum_{i,j=1}^{n} a_{ij}(x)\partial_{ij} f(x) + \sum_{i=1}^{n} b_i(x)\partial_i f(x)$$

where $A = (a_{ij})_{i,j} = \sigma\sigma^*$. Under suitable assumptions on σ_{ij} and b_i, it is possible to show that there exists a reversible measure μ of the form $e^{-V}dx$ for the semigroup associated to L. This provides us with the second ingredient of the Markov triple. For the last element observe that the dynamics described by L is completely equivalent to a carré du champ operator given by $\Gamma(f,g) = \nabla f \cdot A\nabla g$ and associated quadratic form $\Gamma(f)\nabla f \cdot A\nabla f$. Thus $\mathbb{R}^n$ is not endowed with the classical Euclidean structure, but with a different one, described by the matrix A. $\diamondsuit$

Example 2.8. As a final example, let (M,g) be a complete connected smooth Riemannian manifold and let us define $\Gamma(f,f') := \langle \nabla f, \nabla f'\rangle_g$. If we endow M with the volume measure, then the generator L of the underlying process is the Laplace-Beltrami operator Δ_g; on the other hand, if we equip M with $\mu = e^{-V}$ vol, then $\mathrm{L} = \Delta_g - \nabla V \cdot \nabla$. $\diamondsuit$

In all these examples, the reader can observe that Γ represents only the diffusion part of the underlying process. Indeed, it is not affected by the drift term $\nabla V \cdot \nabla$ appearing in the generator L in Example 2.6 and Example 2.8 (in this case when the manifold is equipped with a weighted version of the volume measure); this is not surprising, because by (H10) we can see that Γ is only involved in the second order part of L.

2.2 The Bakry-Émery criterion

The abstract theory we have just begun can be further developed in recursive way as follows: for any $f, g \in \mathscr{A}$ and $n \in \mathbb{N}$ we set

$$\Gamma_0(f,g) := fg, \quad \Gamma_{n+1}(f,g) := \frac{1}{2}\Big(\mathrm{L}\,\Gamma_n(f,g) - \Gamma_n(f,\mathrm{L}\,g) - \Gamma_n(g,\mathrm{L}\,f)\Big)$$

In this way we have a family of bilinear maps and, at first glance, we can see that $\Gamma_1 = \Gamma$. As a second remark, it is worth expressing Γ_2 explicitly, because it has a strong relation with Γ and depending on such relation we are going to introduce the notion of curvature-dimension bound (cf. Definition 2.9); Γ_2 is also called *iterated carré du champ operator* and an immediate computation shows that

$$\Gamma_2(f,g) = \frac{1}{2}\Big(\mathrm{L}\,\Gamma(f,g) - \Gamma(f,\mathrm{L}\,g) - \Gamma(g,\mathrm{L}\,f)\Big), \quad \Gamma_2(f) = \frac{1}{2}\Big(\mathrm{L}\,\Gamma(f) - 2\Gamma(f,\mathrm{L}\,f)\Big)$$

On the contrary, for $n \geq 3$ the geometric meaning of Γ_n is not known yet, so that from now on we will work only with Γ and Γ_2 and we start with the following definition, formulated by Bakry and Émery in 1985.

Definition 2.9. *Let (E,Γ,μ) be a Markov triple, $\rho \in \mathbb{R}$ and $n \in [1,+\infty]$. If $n \in [1,+\infty[$, we say that the semigroup $(\mathrm{P}_t)_{t\geq 0}$ satisfies the* curvature-dimension condition *$CD(\rho,n)$ provided that for every $f \in \mathscr{A}$ it holds*

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

In the case $n = +\infty$, the semigroup $(\mathrm{P}_t)_{t\geq 0}$ satisfies the curvature-dimension condition *$CD(\rho,\infty)$ if*

$$\Gamma_2(f) \geq \rho\Gamma(f)$$

for every $f \in \mathscr{A}$.

From a geometric point of view, the $CD(\rho, n)$ condition has to be meant as a bound from below on the curvature of the space; more precisely, in the case E is a smooth Riemannian manifold equipped with a semigroup satisfying $CD(\rho, n)$, then the Ricci curvature of E is bounded from below by ρ. On the contrary, the condition on n is not an upper bound on the topological dimension of the space, as one could naïvely think, but it is rather linked to the nature of L. In order to better understand this aspect, let us study in a deeper way some of the previous examples and let us see what kind of curvature-dimension condition they satisfy.

Example 2.10. Let us consider the heat semigroup on $(\mathbb{R}^n, dx)$. As we have already seen in Example 2.5, $\mathrm{L} = \Delta$ and $\Gamma(f, g) = \nabla f \cdot \nabla g$; hence, we can easily compute the iterated carré du champ operator, which is given by

$$\Gamma_2(f) = \frac{1}{2}\Delta|\nabla f|^2 - \nabla f \cdot \nabla \Delta f = \|\nabla^2 f\|_{HS}^2 \tag{2.10}$$

where $\nabla^2 f$ denotes the Hessian matrix of f and $\|\nabla^2 f\|_{HS}$ its Hilbert-Schmidt norm. Now, by Cauchy-Schwarz inequality we see that

$$\|\nabla^2 f\|_{HS}^2 = \sum_{i,j=1}^{n} (\partial_{ij} f)^2 \geq \sum_{i=1}^{n} (\partial_i^2 f)^2 \geq \frac{1}{n}(\Delta f)^2 \tag{2.11}$$

and coupling this inequality with (2.10) we see that the heat semigroup satisfies $CD(0, n)$. Clearly, in this case no curvature term appears because $(\mathbb{R}^n, dx)$ is a flat space and the bound on the dimension corresponds to a bound on the topological dimension of $\mathbb{R}^n$. $\Diamond$

Example 2.11. In the case of Kolmogorov's example (cf. Example 2.6), $\mathrm{L} = \Delta - \nabla V \cdot \nabla$ and $\Gamma(f, g) = \nabla f \cdot \nabla g$, whence we deduce that

$$\Gamma_2(f) = \|\nabla^2 f\|_{HS}^2 + \nabla f \cdot \nabla^2 V \nabla f \tag{2.12}$$

If $\nabla^2 V \geq \rho\,\mathrm{Id}$ for some $\rho \in \mathbb{R}$, then the second term at the right-hand side above can be easily estimated as follows: $\nabla f \cdot \nabla^2 V \nabla f \geq \rho|\nabla f|^2$. On the contrary, if we handle the first term by means of Cauchy-Schwarz inequality as in (2.11), we do not get an estimation in terms of L but only Δ; actually,

$$\|\nabla^2 f\|_{HS}^2 \geq \frac{1}{n}(\mathrm{L}\, f)^2$$

is false for every $1 \leq n < +\infty$. The only thing we can say is that $\|\nabla^2 f\|_{HS} \geq 0$, so that (2.12) becomes

$$\Gamma_2(f) \geq \rho\Gamma(f)$$

and this means that $CD(\rho, \infty)$ holds. Thus, although the topological dimension of $\mathbb{R}^n$ is exactly n, no $CD(\rho, m)$ condition holds for finite m: for this reason, n should be regarded as an upper bound on the dimension of the operator L. In the case of the Ornstein-Uhlenbeck semigroup, $\rho = 1$. $\Diamond$

Example 2.12. As a final example, let again (M, g) be a complete connected smooth Riemannian manifold of dimension n equipped with the volume measure. As previously seen, the carré du champ operator is given by $\Gamma(f, f') = \langle \nabla f, \nabla f' \rangle_g$ and L is the Laplace-Beltrami operator Δ_g, so that

$$\Gamma_2(f) = \Delta_g |\nabla f|_g^2 - \nabla f \cdot \nabla \Delta_g f$$

and, by virtue of the Bochner-Lichnerowicz-Wetzenbock formula, this expression can be rewritten as

$$\Gamma_2(f) = \|\nabla^2 f\|_{HS}^2 + \mathrm{Ric}(\nabla f, \nabla f)$$

If we assume that M has Ricci curvature bounded from below by ρ, then by the same argument of (2.11) we see that

$$\Gamma_2(f) \geq \rho\Gamma(f) + \frac{1}{n}(\Delta_g f)^2$$

and this means that the heat semigroup on M satisfies $CD(\rho, n)$. In the case M is the n-dimensional unit sphere S^n, then $CD(n-1, n)$ holds, because S^n has constant Ricci curvature equal to $n-1$. On the other hand, in the case M is the n-dimensional hyperbolic space H^n, the satisfied curvature-dimension condition is the $CD(-n+1, n)$ one, because H^n has constant Ricci curvature equal to $-(n-1)$. $\Diamond$

In all these examples the reader can see that the curvature is not only bounded from below, but also globally constant: for this reason, these examples represent the class of model spaces we will refer to.

2.2.1 Gradient bounds

After this preliminary part on the Γ-calculus, let us now deepen the study of the subject and try to single out characterizations and consequences of a pure curvature condition, namely $CD(\rho, \infty)$. This means that for the moment no dimension bound is involved, but this simplifying hypothesis is not assumed for laziness: the $CD(\rho, n)$ condition is not well understood yet and only few comments will be given later. As a first result, let us show that $CD(\rho, \infty)$ can be equivalently restated in terms of a commutation inequality between Γ and P_t.

Theorem 2.13. *Let (E, Γ, μ) be a Markov triple and let $\rho \in \mathbb{R}$. Then the following facts are equivalent:*

(i) $CD(\rho, \infty)$ holds;

(ii) for any $f \in \mathscr{A}$ and $t \geq 0$ the following weak commutation *inequality holds*

$$\Gamma(\mathrm{P}_t f) \leq e^{-2\rho t}\, \mathrm{P}_t(\Gamma(f)) \tag{WC}$$

Proof. Let us show the two implications.

<u>(i) ⇒ (ii)</u> Let us fix $t \geq 0$ and, for $s \in [0, t]$, let us define

$$\Lambda(s) := \mathrm{P}_s(\Gamma(\mathrm{P}_{t-s} f))$$

For sake of simplicity, set $g := \mathrm{P}_{t-s} f$, so that Λ can be rewritten as $\Lambda(s) = \mathrm{P}_s(\Gamma(g))$. Notice also that $(\Gamma(g))'(0) = -2\Gamma(g, \mathrm{L}\, g)$ because Γ is a quadratic form. With this remark, the first derivative of Λ can be computed as follows

$$\Lambda'(s) = \mathrm{P}_s(\mathrm{L}\,\Gamma(g) - 2\Gamma(g, \mathrm{L}\, g)) = 2\,\mathrm{P}_s(\Gamma_2(g)) \geq 2\rho\,\mathrm{P}_s(\Gamma(g)) = 2\rho\Lambda(s)$$

where the inequality follows from (i). This means that $(\Lambda(s)e^{-2\rho s})' \geq 0$, whence $\Lambda(t)e^{-2\rho t} \geq \Lambda(0)$ by integration and thus the conclusion by definition of Λ.

<u>(ii) ⇒ (i)</u> Since equality holds in (WC) for $t = 0$, the inequality is preserved if we differentiate (WC) in $t = 0$. Therefore

$$2\Gamma(f, \mathrm{L}\, f) \leq -2\rho\Gamma(f) + \mathrm{L}\,\Gamma(f)$$

whence the $CD(\rho, \infty)$ condition. $\square$

The $CD(\rho, \infty)$ condition is actually equivalent to an even stronger commutation inequality between Γ and P_t, but for the forthcoming Theorem 2.15 we first need the next lemma, whose proof is omitted. The interested reader can find it in the already cited book [BGL].

Lemma 2.14. *Let (E, Γ, μ) be a Markov triple and let $\rho \in \mathbb{R}$. Then the following facts are equivalent:*

(i) $CD(\rho, \infty)$ holds;

(ii) for any $f \in \mathscr{A}$ the following inequality is satisfied:

$$\Gamma_2(f) \geq \rho\Gamma(f) + \frac{1}{4}\frac{\Gamma(\Gamma(f))}{\Gamma(f)}$$

With this technical tool, we are finally ready to present the following result, due to Bakry.

Theorem 2.15 (Bakry)**.** *Let (E, Γ, μ) be a Markov triple and let $\rho \in \mathbb{R}$. Then the following facts are equivalent:*

(i) $CD(\rho, \infty)$ holds;

(ii) for any $f \in \mathscr{A}$ and $t \geq 0$ the following strong commutation *inequality holds*

$$\sqrt{\Gamma(\mathrm{P}_t f)} \leq e^{-\rho t}\, \mathrm{P}_t(\sqrt{\Gamma(f)}) \tag{SC}$$

Proof. Let us prove the two implications, starting with the easiest one.
(ii) $\Rightarrow$ (i) Observe that the strong commutation implies the weak one, because by Jensen's inequality $\mathrm{P}_t(\sqrt{\Gamma(f)}) \leq \sqrt{\mathrm{P}_t(\Gamma(f))}$, and by Theorem 2.13 (WC) is equivalent to $CD(\rho, \infty)$.
(i) $\Rightarrow$ (ii) Fix $t \geq 0$ and, for $s \in [0, t]$, let us define

$$\Lambda(s) := \mathrm{P}_s(\sqrt{\Gamma(g)})$$

where $g := \mathrm{P}_{t-s} f$; by (H10) observe that

$$\mathrm{L}\sqrt{H} = \frac{1}{2\sqrt{H}}\,\mathrm{L}\,H - \frac{1}{4H^{3/2}}\Gamma(H)$$

Hence, in the case $H = \Gamma(g)$, the identity above can be used in the computation of Λ' as follows.

$$\begin{aligned}
\Lambda'(s) &= \mathrm{P}_s\left(\mathrm{L}\sqrt{\Gamma(g)} - \frac{\Gamma(g, \mathrm{L}\,g)}{\sqrt{\Gamma(g)}}\right) = \mathrm{P}_s\left(\frac{1}{2\sqrt{\Gamma(g)}}\,\mathrm{L}\,\Gamma(g) - \frac{1}{4\Gamma(g)^{3/2}}\Gamma(\Gamma(g)) - \frac{\Gamma(g, \mathrm{L}\,g)}{\sqrt{\Gamma(g)}}\right) \\
&= \mathrm{P}_s\left(\frac{1}{\sqrt{\Gamma(g)}}\left(\Gamma_2(g) - \frac{1}{4}\frac{\Gamma(\Gamma(g))}{\Gamma(g)}\right)\right) \geq \rho\,\mathrm{P}_s(\sqrt{\Gamma(g)}) = \rho\Lambda(s)
\end{aligned}$$

This inequality entails that $(\Lambda(s)e^{-\rho s})' \geq 0$, whence $\Lambda(t)e^{-\rho t} \geq \Lambda(0)$ and, by definition of Λ, (SC) is proven. ☐

Notice that by (1.5) the Ornstein-Uhlenbeck semigroup satisfies (SC) with $\rho = 1$: hence $CD(1, \infty)$ holds, as already claimed in Example 2.11 and now finally shown. A third commutation inequality between Γ and P_t, known as *mild commutation*, is possible: for every non-negative $f \in \mathscr{A}$ and $t \geq 0$ it holds

$$\frac{\Gamma(\mathrm{P}_t f)}{\mathrm{P}_t f} \leq e^{-2\rho t}\,\mathrm{P}_t\left(\frac{\Gamma(f)}{f}\right) \tag{MC}$$

As suggested by the name, (MC) entails (WC) and is implied by (SC); this can be easily seen: on the one hand (SC) $\Rightarrow$ (MC) because by Cauchy-Schwarz inequality

$$\frac{\Gamma(\mathrm{P}_t f)}{\mathrm{P}_t f} \leq \frac{e^{-2\rho t}}{\mathrm{P}_t f}\,\mathrm{P}_t(\sqrt{\Gamma(f)})^2 = \frac{e^{-2\rho t}}{\mathrm{P}_t f}\,\mathrm{P}_t\left(\sqrt{\frac{\Gamma(f)}{f}}\sqrt{f}\right)^2 \leq e^{-2\rho t}\,\mathrm{P}_t\left(\frac{\Gamma(f)}{f}\right)$$

whereas on the other hand (MC) $\Rightarrow$ (WC), because

$$\textit{ciao}$$

But since $CD(\rho, \infty)$ is equivalent to (WC) and (SC) by Theorem 2.13 and Theorem 2.15 respectively, all the three commutation inequalities are equivalent. Thus it is not possible to fully appreciate the difference between the three commutation inequalities when no dimension condition is involved; such a difference only appears when one tries to introduce a dimension term: in (WC) and (MC) this is actually possible and one can study the relationship between $CD(\rho, n)$ and the modified weak and mild commutations, but on the contrary this is impossible for (SC), where no dimension term can be introduced.

2.2.2 Applications to Poincaré and log-Sobolev

Now let us turn our attention to the applications of Γ-calculus to functional inequalities, in particular local/global Poincaré and log-Sobolev inequalities. In order to state the following result, let us first anticipate the required notation: for a probability measure μ, $f \in \mathscr{A}$ and $t \geq 0$ we set

$$\begin{aligned}
\mathrm{Var}_\mu(f) &:= \int_E f^2 \, d\mu - \left(\int_E f \, d\mu \right)^2 = \int_E \left(f - \int_E f \, d\mu \right)^2 d\mu \\
\mathrm{Var}_{\mathrm{P}_t}(f) &:= \mathrm{P}_t(f^2) - (\mathrm{P}_t \, f)^2 \\
\mathrm{Ent}_{\mathrm{P}_t}(f) &:= \mathrm{P}_t(f \log f) - \mathrm{P}_t \, f \log \mathrm{P}_t \, f
\end{aligned}$$

With this premise, the next theorem makes perfectly sense.

Theorem 2.16. *Let (E, Γ, μ) be a Markov triple with the further assumption that μ is a probability measure and let $\rho \in \mathbb{R}$. Then the following facts are equivalent:*

(i) $CD(\rho, \infty)$ holds;

(ii) (SC) is satisfied;

(iii) (local Poincaré inequality) for any $f \in \mathscr{A}$ and $t \geq 0$ we have

$$\mathrm{Var}_{\mathrm{P}_t}(f) \leq \frac{1 - e^{-2\rho t}}{\rho} \, \mathrm{P}_t(\Gamma(f))$$

(iv) (inverse local Poincaré inequality) for any $f \in \mathscr{A}$ and $t \geq 0$ it holds

$$\mathrm{Var}_{\mathrm{P}_t}(f) \geq \frac{e^{2\rho t} - 1}{\rho} \Gamma(\mathrm{P}_t \, f)$$

(v) (local log-Sobolev inequality) for any non-negative $f \in \mathscr{A}$ and $t \geq 0$ we have

$$\mathrm{Ent}_{\mathrm{P}_t}(f) \leq \frac{1 - e^{-2\rho t}}{2\rho} \, \mathrm{P}_t \left(\frac{\Gamma(f)}{f} \right)$$

(vi) (inverse local log-Sobolev inequality) for any non-negative $f \in \mathscr{A}$ and $t \geq 0$ it holds

$$\mathrm{Ent}_{\mathrm{P}_t}(f) \geq \frac{e^{2\rho t} - 1}{2\rho} \frac{\Gamma(\mathrm{P}_t \, f)}{\mathrm{P}_t \, f}$$

Before starting the proof, let us stress that, in order to give a meaning to the inequalities above in the case $\rho = 0$, we replace $e^{2\rho t} - 1$ and $1 - e^{-2\rho t}$ by the first order Taylor expansion in ρ, i.e. $2\rho t$. In the proof below we are going to perform all the computations with $\rho \neq 0$; the zero-curvature case is clearly simpler.

Proof. The equivalence between (i) and (ii) has already been proved in Theorem 2.15. In addition, (v) ⇒ (iii) and (vi) ⇒ (iv) because, roughly speaking, relative entropy has more information than variance. More precisely, variance is the second order expansion of the relative entropy functional and the first non-zero term of the expansion, in the sense that for any probability measure ν, for any bounded function g and for ε small enough

$$\mathrm{Ent}_\nu(1 + \varepsilon g) = \frac{\varepsilon^2}{2} \mathrm{Var}_\nu(g) + o(\varepsilon^2) \tag{2.13}$$

Moreover

$$\mathrm{P}_t \left(\frac{\Gamma(1 + \varepsilon g)}{1 + \varepsilon g} \right) = \mathrm{P}_t \left(\frac{\varepsilon^2 \Gamma(g)}{1 + \varepsilon g} \right) = \varepsilon^2 \, \mathrm{P}_t(\Gamma(g)) + o(\varepsilon^2) \tag{2.14}$$

because, as already noticed, Γ vanishes on constant functions and enjoys the polarization formula $\Gamma(h + h') = \Gamma(h) + 2\Gamma(h, h') + \Gamma(h')$, since it is a quadratic form. Thus if (v) holds, it is sufficient to consider it with $f = 1 + \varepsilon g$, so that by (2.13) and (2.14) property (iii) follows. By the same argument (vi) entails (iv).

As a next step, let us prove that (iii) implies (i). Since equality holds in the local Poincaré inequality for $t = 0$, the inequality is preserved if we differentiate both sides and evaluate them at $t = 0$; unfortunately, this procedure is not sufficient to conclude, because in $t = 0$ the derivatives of the two sides agree. Hence we need to differentiate one more time. For the left-hand side, observe that

$$\begin{aligned}
\left.\frac{d^2}{dt^2}\right|_{t=0} \mathrm{Var}_{\mathrm{P}_t}(f) \quad &= \quad \mathrm{L}\,\mathrm{L}(f^2) - 2(\mathrm{L}\,f)^2 - 2f\,\mathrm{L}\,\mathrm{L}\,f \\
&\overset{(2.7)}{=} \quad 2\,\mathrm{L}(f\,\mathrm{L}\,f) + 2\,\mathrm{L}\,\Gamma(f) - 2(\mathrm{L}\,f)^2 - 2f\,\mathrm{L}\,\mathrm{L}\,f \\
&\overset{(H10)}{=} \quad 2((\mathrm{L}\,f)^2 + f\,\mathrm{L}\,\mathrm{L}\,f + 2\Gamma(f, \mathrm{L}\,f)) + 2\,\mathrm{L}\,\Gamma(f) - 2(\mathrm{L}\,f)^2 - 2f\,\mathrm{L}\,\mathrm{L}\,f \\
&= \quad 4\Gamma(f, \mathrm{L}\,f) + 2\,\mathrm{L}\,\Gamma(f)
\end{aligned}$$

while for the right-hand side we have

$$\left.\frac{d^2}{dt^2}\right|_{t=0} \frac{1 - e^{-2\rho t}}{\rho} \mathrm{P}_t(\Gamma(f)) = -4\rho\Gamma(f) + 4\,\mathrm{L}\,\Gamma(f)$$

Since, as just said, the first derivatives of the two sides agree at $t = 0$, the inequality is inherited by the second derivatives, so that

$$4\Gamma(f, \mathrm{L}\,f) + 2\,\mathrm{L}\,\Gamma(f) \leq -4\rho\Gamma(f) + 4\,\mathrm{L}\,\Gamma(f)$$

and the $CD(\rho, \infty)$ condition immediately follows by the very definition of Γ_2. In the same fashion, the reader can prove that (iv) entails (i). For this reason, the proof will be achieved if we are able to show that $CD(\rho, \infty)$ entails (v) and (vi). To this aim, fix $t \geq 0$ and for $s \in [0, t]$ let us define

$$\Lambda(s) := \mathrm{P}_s(g \log g)$$

where $g := \mathrm{P}_{t-s} f$; observe that $\mathrm{Ent}_{\mathrm{P}_t}(f) = \Lambda(t) - \Lambda(0)$. As usual in Γ-calculus arguments, the next step consists in the computation of Λ'.

$$\Lambda'(s) \quad = \quad \mathrm{P}_s(\mathrm{L}(g \log g) - \log g\,\mathrm{L}\,g - \mathrm{L}\,g) = \mathrm{P}_s(2\Gamma(g, \log g) + g\,\mathrm{L}\log g - \mathrm{L}\,g)$$

$$= \quad \mathrm{P}_s\left(\frac{2}{g}\Gamma(g) + g\left(\frac{1}{g}\,\mathrm{L}\,g - \frac{1}{g^2}\Gamma(g)\right) - \mathrm{L}\,g\right) = \mathrm{P}_s\left(\frac{\Gamma(g)}{g}\right) = \mathrm{P}_s(g\Gamma(\log g)) \quad (2.15)$$

The second identity is motivated by the very definition of Γ, because for any $h, h' \in \mathscr{A}$

$$\mathrm{L}(hh') = 2\Gamma(h,h') + h\,\mathrm{L}\,h' + h'\,\mathrm{L}\,h \quad (2.16)$$

so that it suffices to take $h = g$ and $h' = \log g$, while the following ones are due to (H10), more precisely to (2.7) and (2.8). However, unlike all the previous demonstrations, in this case the computation of Λ' is not sufficient and now comes the intuition of Bakry: we need to compute Λ''.

$$\begin{aligned}
\Lambda''(s) &= \mathrm{P}_s\left(\mathrm{L}\left(\frac{\Gamma(g)}{g}\right) - 2\frac{\Gamma(g,\mathrm{L}\,g)}{g} + \frac{\Gamma(g)\,\mathrm{L}\,g}{g^2}\right) \\
&= \mathrm{P}_s\left(2\Gamma(\Gamma(g), 1/g) + \frac{1}{g}\,\mathrm{L}\,\Gamma(g) + \Gamma(g)\,\mathrm{L}(1/g) - 2\frac{\Gamma(g,\mathrm{L}\,g)}{g} + \frac{\Gamma(g)\,\mathrm{L}\,g}{g^2}\right) \\
&= \mathrm{P}_s\left(-\frac{2}{g^2}\Gamma(\Gamma(g),g) + \frac{1}{g}\,\mathrm{L}\,\Gamma(g) + \frac{2}{g^3}\Gamma(g)^2 - 2\frac{\Gamma(g,\mathrm{L}\,g)}{g}\right) \\
&= \mathrm{P}_s\left(\frac{2}{g}\Gamma_2(g) - \frac{2}{g^2}\Gamma(\Gamma(g),g) + \frac{2}{g^3}\Gamma(g)^2\right) = 2\,\mathrm{P}_s(g\Gamma_2(\log g)) \\
&\overset{(i)}{\geq} 2\rho\,\mathrm{P}_s(g\Gamma(\log g))
\end{aligned}$$

The second identity is again motivated by (2.16) with $h = \Gamma(g)$ and $h' = 1/g$, the third one by (2.7) and (2.8), while the fourth one comes from the definition of Γ_2. Plugging (2.15) into the computations above we finally get

$$\Lambda'' \geq 2\rho\Lambda'$$

and this means that $(\Lambda'(s)e^{-2\rho s})' \geq 0$. If we first integrate this inequality on $[s,t]$, we get

$$\Lambda'(s) \leq \Lambda'(t)e^{-2\rho(t-s)}$$

whence, by a further integration on $[0,t]$,

$$\Lambda(t) - \Lambda(0) \leq \Lambda'(t)\frac{1-e^{-2\rho t}}{2\rho}$$

In this way, recalling that $\mathrm{Ent}_{\mathrm{P}_t}(f) = \Lambda(t) - \Lambda(0)$ and $\Lambda'(t) = \mathrm{P}_t(\Gamma(f)/f)$, (v) is proved. On the other hand, if we first integrate $(\Lambda'(s)e^{-2\rho s})' \geq 0$ on $[0,s]$, whence

$$\Lambda'(s) \geq e^{2\rho s}\Lambda'(0)$$

and then integrate again on $[0,t]$, so that

$$\Lambda(t) - \Lambda(0) \geq \Lambda'(0)\frac{e^{2\rho t}-1}{2\rho}$$

then (vi) is proved, because $\Lambda'(0) = \Gamma(\mathrm{P}_t\,f)/\mathrm{P}_t\,f$. All the stated equivalences are thus established. □

The proof being achieved, let us now point out a few of comments. First of all, observe that (iii) - (vi) are local inequalities because all the right-hand sides depend on x (on the contrary, the forthcoming Poincaré inequality (2.17) and log-Sobolev inequality (2.18) are global, because the right-hand sides are integrated, but these two inequalities will be treated later). Secondly, the inverse local Poincaré inequality shows the regularizing property of P_t, because if $f \in L^2(E,\mu)$

then (iv) gives a pointwise upper bound on $\Gamma(\mathrm{P}_t f)$ which only depends on t and f; in the case $\Gamma(f) = |\nabla f|^2$, the usual notion of regularization is recovered. Moreover, it is worth underlying the equivalence between $CD(\rho, \infty)$ and (v) + (vi), because log-Sobolev type inequalities are richer than spectral ones (as for instance Poincaré inequality) and this has already been seen in the previous proof: indeed by (2.13) we know that relative entropy has more information than variance. Finally, the previous theorem has two immediate consequences; the first one is the following.

Corollary 2.17. *Let (E, Γ, μ) be a Markov triple satisfying $CD(\rho, \infty)$ for $\rho > 0$; let us further assume that μ is a probability measure. Then the following facts hold:*

(a) for any $f \in \mathscr{A}$ we have

$$\mathrm{Var}_\mu(f) \leq \frac{1}{\rho} \int_E \Gamma(f)\, d\mu \tag{2.17}$$

that is a Poincaré inequality of constant $1/\rho$;

(b) for any non-negative $f \in \mathscr{A}$ we have

$$\mathrm{Ent}_\mu(f^2) \leq \frac{2}{\rho} \int_E \Gamma(f)\, d\mu \tag{2.18}$$

that is a log-Sobolev inequality of constant $2/\rho$.

Proof. For (a) it is sufficient to consider (iii) of Theorem 2.16 and let $t \to +\infty$. For (b), the same procedure applies to (v). □

As a straightforward byproduct, if we take $E = \mathbb{R}^n$, $\Gamma(f, g) = \nabla f \cdot \nabla g$ for any $f, g \in \mathscr{A}$, $\mu(dx) = ce^{-V} dx$ (where c is the normalization constant, so that μ is a probability measure) and we assume that there exists $\rho > 0$ such that $\nabla^2 V \geq \rho\, \mathrm{Id}$, then Corollary 2.17 entails Poincaré and log-Sobolev inequalities for the weighted Lebesgue measure. In particular, the latter can be rewritten in a more familiar way as follows

$$\mathrm{Ent}_\mu(f) \leq \frac{1}{2\rho} \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f}\, d\mu$$

where the reader can recognize the Fisher information of f at the right-hand side. Since the Ornstein-Uhlenbeck semigroup is a particular case of the situation we have just described and satisfies $CD(1, \infty)$, (2.18) precisely becomes (1.8), so that we have another proof of Theorem 1.1. Let us now pass to the second consequence of Theorem 2.16.

Corollary 2.18. *In the same framework of the previous corollary the following facts are true:*

(c) (2.17) is equivalent to

$$\mathrm{Var}_\mu(\mathrm{P}_t f) \leq e^{-2\rho t}\, \mathrm{Var}_\mu(f) \tag{2.19}$$

for every $f \in L^2(\mathbb{R}^n, \mu)$ and $t \geq 0$;

(d) (2.18) is equivalent to

$$\mathrm{Ent}_\mu(\mathrm{P}_t f) \leq e^{-\rho t}\, \mathrm{Ent}_\mu(f) \tag{2.20}$$

for every non-negative $f \in L \log L(\mathbb{R}^n, \mu)$ and $t \geq 0$.

Proof. Let us prove (c) and, as a first step, let us define $\Lambda(t) := \mathrm{Var}_\mu(\mathrm{P}_t f)$, assume (2.17) and notice that

$$\Lambda'(t) = 2 \int_E \mathrm{P}_t f\, \mathrm{L}\, \mathrm{P}_t f\, d\mu \overset{(IP)}{=} -2 \int_E \Gamma(\mathrm{P}_t f)\, d\mu \overset{(2.17)}{\leq} -2\rho\, \mathrm{Var}_\mu(\mathrm{P}_t f) = -2\rho \Lambda(t)$$

whence $(\Lambda(t)e^{2\rho t})' \leq 0$. This implies that $\Lambda(t)e^{2\rho t} \leq \Lambda(0)$ and so (2.19) by definition of Λ. For the opposite implication, observe that for $t = 0$ equality holds in (2.19), so that the inequality is still true if we differentiate (2.19) and evaluate it at $t = 0$. Since

$$(\mathrm{Var}_\mu(\mathrm{P}_t\, f))'\Big|_{t=0} = \Lambda'(0) = -2\int_E \Gamma(f)\, d\mu$$

and $(e^{-2\rho t}\,\mathrm{Var}_\mu(f))'(0) = -2\rho\,\mathrm{Var}_\mu(f)$, we exactly get (2.17). *Mutatis mutandis*, for (d) the argument is exactly the same for both implications and thus will be omitted. □

About Corollary 2.17, it is worth saying that Poincaré inequality can be proved in a different way via spectral techniques, but this procedure is much more difficult (and will not be treated at all here), so that we have one more evidence of the strength of Γ-calculus. For the reader who is not at ease with Poincaré inequality, its spectral nature is contained in the constant at the right-hand side, in this case $1/\rho$, because it represents the spectral gap of L, i.e. the distance between 0 and the least positive eigenvalue of L.

About Corollary 2.18, it provides us not only with two asymptotic behaviours for $\mathrm{P}_t\, f$, but also with the respective convergence rates, because (2.19) tells us how fast $\mathrm{P}_t\, f$ converges in variance to the equilibrium, while (2.20) gives us the same information in terms of entropy convergence. As a last remark on log-Sobolev type inequalities, the adopted definition of relative entropy is crucial; indeed, given a function $\phi : \mathbb{R} \to \mathbb{R}$ a more general relative ϕ-entropy can be defined as follows

$$\mathrm{Ent}_\nu^\phi(f) := \int_E \phi(f)\, d\nu - \phi\bigg(\int_E f\, d\nu\bigg)$$

but the only nice cases are given by $\phi(x) = x^2$ and $\phi(x) = x\log x$ and in these notes only the latter is considered.

Up to now, the only $CD(\rho,\infty)$ condition has been studied, i.e. no dimension has appeared yet. Hence it is natural to wonder how to restate for $CD(\rho, n)$ what we have seen for $CD(\rho,\infty)$, but as already anticipated, the introduction of the dimension in the theory of Bakry and Émery is not easy at all and not well understood yet. However, when present, a dimension condition strengthen the information we can get and a first example is given by the next theorem, which has to be compared with Corollary 2.17.

Theorem 2.19. *Let (E,Γ,μ) be a Markov triple satisfying $CD(\rho, n)$ with $\rho > 0$ and $1 \leq n < +\infty$; let us further assume that μ is a probability measure. Then the following facts are true:*

(i) for any $f \in \mathscr{A}$ we have

$$\mathrm{Var}_\mu(f) \leq \frac{n-1}{\rho n}\int_E \Gamma(f)\, d\mu \tag{2.21}$$

that is a Poincaré inequality of constant $(n-1)/\rho n$;

(ii) for any non-negative $f \in \mathscr{A}$ we have

$$\mathrm{Ent}_\mu(f^2) \leq 2\frac{n-1}{\rho n}\int_E \Gamma(f)\, d\mu \tag{2.22}$$

that is a log-Sobolev inequality of constant $2(n-1)/\rho n$.

Observe that $(n-1)/n < 1$, so that (2.21) implies (2.17) and (2.22) implies (2.18); this is perfectly coherent with the (intuitive) fact that $CD(\rho, n)$ entails $CD(\rho,\infty)$, but that is not all what we can say. These inequalities are better than the ones we got in Corollary 2.17, because they are optimal for the sphere, in the sense that now equality is attained, and so we have also better spectral estimates. For the proof we first need the following result.

Proposition 2.20. *Let (E, Γ, μ) be a Markov triple with the further assumption that μ is a probability measure. Then the Poincaré inequality of constant c*

$$\mathrm{Var}_\mu(f) \leq c \int_E \Gamma(f)\, d\mu, \quad \forall f \in \mathscr{A} \tag{2.23}$$

is equivalent to

$$\int_E \Gamma(f)\, d\mu \leq c \int_E \Gamma_2(f)\, d\mu, \quad \forall f \in \mathscr{A} \tag{2.24}$$

Proof. Let us prove the two implications.

As a first step, let us assume (2.23) and take $f \in \mathscr{A}$ such that $\int_E f\, d\mu = 0$: as a consequence $\mathrm{Var}_\mu(f)$ coincides with the squared $L^2(E, \mu)$-norm of f and this hypothesis is not restrictive, because modifications of f by constants do not affect Γ and Γ_2, so that for $\tilde{f} := f - \int_E f\, d\mu$ we have $\Gamma(f) = \Gamma(\tilde{f})$ and $\Gamma_2(f) = \Gamma_2(\tilde{f})$. Then observe that, by Cauchy-Schwarz inequality and by (2.23), we have

$$\int_E \Gamma(f)\, d\mu \overset{(IP)}{=} -\int_E f \,\mathrm{L}\, f\, d\mu \leq \sqrt{\int_E f^2\, d\mu}\sqrt{\int_E (\mathrm{L}\, f)^2\, d\mu} \leq \sqrt{c\int_E \Gamma(f)\, d\mu}\sqrt{\int_E (\mathrm{L}\, f)^2\, d\mu}$$

In order to get the conclusion, notice that by definition of Γ_2, by (IP) and by (2.6)

$$\int_E \Gamma_2(f)\, d\mu = \frac{1}{2}\int_E \mathrm{L}\,\Gamma(f)\, d\mu - \int_E \Gamma(f, \mathrm{L}\, f)\, d\mu = \int_E (\mathrm{L}\, f)^2\, d\mu$$

Plugging this identity into the previous inequality we deduce that

$$\sqrt{\int_E \Gamma(f)\, d\mu} \leq \sqrt{c\int_E \Gamma_2(f)\, d\mu}$$

whence (2.24) by squaring. On the contrary, if we assume (2.24) and we take any $f \in \mathscr{A}$, then we shall define $\Lambda(t) := \mathrm{Var}_\mu(\mathrm{P}_t\, f)$ and observe that, for any $s \in [0, t]$,

$$\Lambda'(s) = -2\int_E \Gamma(\mathrm{P}_s\, f)\, d\mu$$

As in the proof of Theorem 2.16, the computation of Λ' is not sufficient, so that we have to determine Λ'' as well:

$$\Lambda''(s) = 4\int_E \Gamma_2(\mathrm{P}_s\, f)\, d\mu \overset{(2.24)}{\geq} \frac{4}{c}\int_E \Gamma(\mathrm{P}_s\, f)\, d\mu = -\frac{2}{c}\Lambda'(s)$$

Thus $(\Lambda'(s)e^{2s/c})' \geq 0$, whence $\Lambda'(t) \geq e^{-2t/c}\Lambda'(0)$ by integration on $[0, t]$; a further integration on $[0, +\infty[$ provides us with $\Lambda(0) \leq c\Lambda'(0)$, whence (2.23) by definition of Λ and by the computation of Λ' we performed. □

Thanks to this characterization, we can come back to Theorem 2.19 and prove it.

Proof. It is sufficient to observe that

$$\int_E \Gamma_2(f)\, d\mu \geq \int_E \left(\rho\Gamma(f) + \frac{1}{n}(\mathrm{L}\, f)^2\right) d\mu = \rho\int_E \Gamma(f)\, d\mu + \frac{1}{n}\int_E \Gamma_2(f)\, d\mu$$

whence

$$\frac{n-1}{n}\int_E \Gamma_2(f)\, d\mu \geq \rho\int_E \Gamma(f)\, d\mu$$

and thus the conclusion. □

3 Harnack's inequality

A first glance on the role played by dimension bounds in Γ-calculus has been already seen in Theorem 2.19, where Poincaré and log-Sobolev inequalities have been refined. Now, we are going to approach a completely new functional inequality, due to Carl Gustav Axel von Harnack (1851-1888) and named after him, whose proof relies on the $CD(0,n)$ condition.

3.1 Classical statements

First of all, let us state Harnack's original result, recalling that a function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be harmonic provided that $\Delta f = 0$.

Theorem 3.1 (Harnack, 1887)**.** *Let $\Omega \subset \mathbb{R}^n$ and let A be a bounded, open and connected domain such that $A \subset \overline{A} \subset \Omega$. Then there exists a constant $C = C_A$, depending only on A, such that for every non-negative harmonic function f it holds*

$$\frac{1}{C} f(y) \le f(x) \le C f(y)$$

for every $x, y \in A$.

From an intuitive point of view, the two-sided uniform bound is possible because, at a given point x, any harmonic function can be expressed as an integral of such function over any sphere centered at x; more precisely, the invoked formula is

$$f(x) = \frac{1}{R^{n-1}\omega_{n-1}} \int_{|y-x|=R} f(y) dy$$

where ω_{n-1} denotes the area of the unit sphere in $\mathbb{R}^n$, which is nothing but the mean value theorem for harmonic functions. For one-parameter Markov semigroups, we have seen that hypotheses (H1)-(H7) entail the representation formula (2.2), so that also in this case we have an integral representation for $\mathrm{P}_t f$, although now the integration is over the whole space. Thus, one can conjecture that a parabolic version of Harnack's inequality holds and, indeed, this is actually the case in the sense of the following theorem (cf. Evans' book *Partial Differential Equations* for a more precise statement).

Theorem 3.2. *Let $\Omega \subset \mathbb{R}^n$, let A be a bounded, open and connected domain such that $A \subset \overline{A} \subset \Omega$, let $(\mathrm{P}_t)_{t\ge 0}$ be a Markov semigroup of operators and let $0 \le t_1 < t_2$. Then there exists a constant $C = C_{A,(\mathrm{P}_t),t_1,t_2}$ such that for every non-negative function f it holds*

$$\mathrm{P}_{t_1} f(x) \le C\, \mathrm{P}_{t_2} f(y) \tag{3.1}$$

for every $x, y \in A$.

Observe that if we set $u(t,x) := \mathrm{P}_t f(x)$, then u is the only solution to the parabolic equation $\partial_t u = \mathrm{L}\, u$, whence the name *parabolic Harnack's inequality*; in addition, (3.1) can be restated as

$$\sup_{x\in A} u(t_1, x) \le \inf_{y\in A} u(t_2, y)$$

The interest in this new inequality is due to the fact that it allows to control the present state at a given point x in terms of the future state at any point y belonging to a suitable set A *a priori* fixed. The result can be actually generalized to smooth Riemannian manifolds (and indeed we are going to prove it in this framework), but it is not known on Polish spaces yet.

3.2 The proof

For the Riemannian case, the standard Γ-calculus approach is the following. Let (M, g) be a smooth Riemannian manifold, fix $t \geq 0$, $s > 0$, $x, y \in M$, let $(x_u)_{u\in[s,t]}$ be a constant speed geodesic from x to y and for $u \in [s,t]$ let us define $\Lambda(u) := \mathrm{P}_u f(x_u)$. Then

$$\begin{aligned}
\Lambda'(u) &= \mathrm{P}_u(\mathrm{L} f)(x_u) + \nabla \mathrm{P}_u f(x_u)\cdot \dot{x}_u \geq \mathrm{P}_u(\mathrm{L} f)(x_u) - |\nabla \mathrm{P}_u f(x_u)||\dot{x}_u| \\
&= \mathrm{P}_u(\mathrm{L} f)(x_u) - |\nabla \mathrm{P}_u f(x_u)|\frac{d(x,y)}{t-s} \\
&\geq \mathrm{P}_u(\mathrm{L} f)(x_u) - \frac{|\nabla \mathrm{P}_u f(x_u)|^2}{2\lambda \mathrm{P}_u f(x_u)} - \frac{d^2(x,y)}{2(t-s)^2}\lambda \mathrm{P}_u f(x_u)
\end{aligned}$$

where the last step is motivated by the following modified version of Young's inequality:

$$ab \leq \frac{\lambda a^2}{2} + \frac{b^2}{2\lambda} \tag{3.2}$$

However, we are now stopped because we are not able to estimate the last term in the chain of inequalities above by means of the results we presented by now. Hence we need to study

$$\mathrm{P}_u(\mathrm{L} f)(x_u) - \frac{|\nabla \mathrm{P}_u f(x_u)|^2}{2\lambda \mathrm{P}_u f(x_u)}$$

and in order to get a guess on the final result, let us consider a special example: the heat semigroup on $\mathbb{R}^n$. As already seen in Example 2.5, the heat semigroup is given by

$$\mathrm{P}_t f(x) = \int_{\mathbb{R}^n} f(y)p_t(x,y)dy, \quad p_t(x,y) = \frac{1}{(4\pi t)^{n/2}}e^{-\frac{|x-y|^2}{4t}}$$

and observe that the gradient of the heat kernel p_t is given by

$$\nabla_x p_t(x,y) = -\frac{x-y}{2t}p_t(x,y)$$

so that

$$\Delta_x p_t(x,y) = -\frac{n}{2t}p_t(x,y) - \frac{x-y}{2t}\nabla_x p_t(x,y) = -\frac{n}{2t}p_t(x,y) + \frac{|\nabla_x p_t(x,y)|^2}{p_t(x,y)} \tag{3.3}$$

This consideration on p_t can be used to estimate the Laplacian of a general solution to the heat equation, because

$$\begin{aligned}
\Delta \mathrm{P}_t f(x) &= \int_{\mathbb{R}^n} f(y)\Delta_x p_t(x,y)dy = -\frac{n}{2t}\int_{\mathbb{R}^n} f(y)p_t(x,y)dy + \int_{\mathbb{R}^n} f(y)\frac{|\nabla_x p_t(x,y)|^2}{p_t(x,y)}dy \\
&= -\frac{n}{2t}\mathrm{P}_t f(x) + \mathrm{P}_t f(x)\int_{\mathbb{R}^n} f(y)\frac{|\nabla_x p_t(x,y)|^2}{p_t(x,y)^2}\frac{p_t(x,y)}{\mathrm{P}_t f(x)}dy \\
&\geq -\frac{n}{2t}\mathrm{P}_t f(x) + \mathrm{P}_t f(x)\left|\int_{\mathbb{R}^n} f(y)\frac{|\nabla_x p_t(x,y)|}{p_t(x,y)}\frac{p_t(x,y)}{\mathrm{P}_t f(x)}dy\right|^2 \\
&= -\frac{n}{2t}\mathrm{P}_t f(x) + \frac{|\nabla \mathrm{P}_t f(x)|^2}{\mathrm{P}_t f(x)}
\end{aligned}$$

This is exactly the kind of result we are looking for and it is known in the literature as *Li-Yau inequality*. Pay attention to the fact that the inequality above is motivated by Jensen's inequality: indeed

$$f(y)\frac{p_t(x,y)}{\mathrm{P}_t f(x)}dy$$

is a probability measure on $\mathbb{R}^n$ (it is precisely here that we use the assumption that $f \geq 0$) and this allows us to invoke such result. Moreover, thanks to (3.3) we see that equality can actually be attained and this means that the inequality is sharp on $\mathbb{R}^n$, but to be honest heat kernels are extremal for the Li-Yau inequality also on Riemannian manifolds w.r.t. the Laplace-Beltrami operator. Now, let us try to generalize what we have just seen to $CD(0,n)$ spaces.

Theorem 3.3 (Li-Yau, 1986)**.** *Let (E,Γ,μ) be a Markov triple satisfying $CD(0,n)$. Then for every non-negative f and $t>0$ it holds*

$$\mathrm{P}_t(\mathrm{L}\, f) - \frac{\Gamma(\mathrm{P}_t\, f)}{\mathrm{P}_t\, f} \geq -\frac{n}{2t}\,\mathrm{P}_t\, f \tag{3.4}$$

Before the proof, let us point out that a modified Li-Yau inequality holds also in the case of negative curvature, in the sense that for every non-negative f and $t>0$ we have

$$\frac{\Gamma(\mathrm{P}_t\, f)}{\mathrm{P}_t\, f} \leq \phi_{\rho,t}\left(\frac{\mathrm{L}\,\mathrm{P}_t\, f}{\mathrm{P}_t\, f}\right)$$

for a suitable function $\phi_{\rho,t}$ dependent on time and curvature. However, $\phi_{\rho,t}$ does not have the same nice properties as for $\rho \geq 0$.

Proof. As a preliminary consideration, (3.4) is easily seen to be equivalent to

$$1 - \frac{2t}{n\,\mathrm{P}_t\, f}\left(\frac{\Gamma(\mathrm{P}_t\, f)}{\mathrm{P}_t\, f} - \mathrm{L}\,\mathrm{P}_t\, f\right) \geq 0 \tag{3.5}$$

so that the Li-Yau inequality will be proved in this form. Let f be as in the statement, fix $t>0$ and for $s \in [0,t]$ let us define

$$\Lambda(s) := \mathrm{P}_s(g \log g)$$

where $g = \mathrm{P}_{t-s}\, f$. The first derivative is given by $\Lambda'(s) = \mathrm{P}_s(g\Gamma(\log g))$, as already computed at (2.15); the fact that now μ needs not be a probability measure, as in Theorem 2.16, is not important, since in that case this assumption was requested to give a meaning to the variance. As in Theorem 2.16 however, the computation of Λ' is not sufficient; hence, we need to determine Λ'' too, but now the $CD(\rho,\infty)$ condition is replaced by the $CD(0,n)$ one. For this reason,

$$\begin{aligned} \Lambda''(s) &= 2\,\mathrm{P}_s(g\Gamma_2(\log g)) \geq \frac{2}{n}\,\mathrm{P}_s(g(\mathrm{L}\log g)^2) \overset{(2.7)}{=} \frac{2}{n}\,\mathrm{P}_s\left(g\left(\frac{1}{g}\,\mathrm{L}\, g - \frac{1}{g^2}\Gamma(g)\right)^2\right) \\ &= \frac{2}{n}\,\mathrm{P}_s\left(\frac{1}{g}\left(\mathrm{L}\, g - \frac{1}{g}\Gamma(g)\right)^2\right) \end{aligned}$$

the first step being motivated by what we have already seen in the proof of Theorem 2.16. In order to go on, notice that, by Cauchy-Schwarz inequality, for any $s \geq 0$, for any function H and non-negative function G it holds

$$\mathrm{P}_s\left(\frac{H^2}{G}\right) \geq \frac{\mathrm{P}_s(H)^2}{\mathrm{P}_s(G)}$$

Therefore

$$\Lambda''(s) \geq \frac{2}{n\,\mathrm{P}_s\, g}\left(\mathrm{P}_s(\mathrm{L}\, g) - \mathrm{P}_s\left(\frac{\Gamma(g)}{g}\right)\right)^2 \tag{3.6}$$

Now observe that, by definition of g, $\mathrm{P}_s\, g = \mathrm{P}_t\, f$, $\mathrm{P}_s(\mathrm{L}\, g) = \mathrm{P}_t(\mathrm{L}\, f)$ and

$$\mathrm{P}_s\left(\frac{\Gamma(g)}{g}\right) = \mathrm{P}_s(g\Gamma(\log g)) = \Lambda'(s)$$

so that, plugging these identities into (3.6) we get

$$\Lambda''(s) \geq \frac{2}{n\,\mathrm{P}_t\, f}\Big(\mathrm{P}_t\,\mathrm{L}\, f - \Lambda'(s)\Big)^2$$

If we set $\varphi(s) := \Lambda(s) - s\,\mathrm{P}_t\,\mathrm{L}\, f$, then this inequality becomes

$$\varphi''(s) \geq \frac{2}{n\,\mathrm{P}_t\, f}(\varphi'(s))^2 \tag{3.7}$$

and this implies that

$$\psi(s) := \exp\Big(- \frac{2}{n\,\mathrm{P}_t\, f}\varphi(s)\Big)$$

is concave. Indeed, $\psi' = -2\varphi'\psi/(n\,\mathrm{P}_t\, f)$ whence

$$\psi''(s) = \frac{2}{n\,\mathrm{P}_t\, f}\Big(- \varphi'' + \frac{2}{n\,\mathrm{P}_t\, f}(\varphi')^2\Big)\psi \overset{(3.7)}{\leq} 0$$

By concavity we can then say that $t\psi'(0) \geq \psi(t) - \psi(0)$, i.e.

$$-\frac{2t}{n\,\mathrm{P}_t\, f}(\Lambda'(0) - \mathrm{L}\,\mathrm{P}_t\, f)\exp\Big(-\frac{2}{n\,\mathrm{P}_t\, f}\Lambda(0)\Big) \geq \exp\Big(-\frac{2}{n\,\mathrm{P}_t\, f}(\Lambda(t) - t\,\mathrm{L}\,\mathrm{P}_t\, f)\Big) - \exp\Big(-\frac{2}{n\,\mathrm{P}_t\, f}\Lambda(0)\Big)$$

and by purely algebraic manipulations we finally get

$$1 - \frac{2t}{n\,\mathrm{P}_t\, f}\Big(\frac{\Gamma(\mathrm{P}_t\, f)}{\mathrm{P}_t\, f} - \mathrm{L}\,\mathrm{P}_t\, f\Big) \geq \exp\Big(- \frac{2}{n\,\mathrm{P}_t\, f}(\Lambda(t) - \Lambda(0) - t\,\mathrm{L}\,\mathrm{P}_t\, f)\Big) \tag{3.8}$$

Since the right-hand side is always positive, (3.5) follows and thus the conclusion. □

It is worth saying that, in the original proof of Li and Yau, Γ-calculus is not involved, because all computations are made only in the smooth case, i.e. on Riemannian manifolds. In fact, Li and Yau knew the result they were looking for, because they had the heat equation as model example and so they explicitly proved the conjecture. On the contrary, the proof we propose is abstract and holds on Polish spaces supporting a Markov triple. Thanks to this result, it is now possible to prove parabolic Harnack's inequality in a form which is even more general than the one we presented at Theorem 3.2; the statement is then the following.

Theorem 3.4. *Let (M, g) be a complete smooth Riemannian manifold equipped with the measure $\mu = e^{-V}\,\mathrm{vol}$, where* vol *denotes the volume measure, and let $(\mathrm{P}_t)_{t\geq 0}$ be the linear semigroup with generator $\mathrm{L} = \Delta_g - \nabla V \cdot \nabla$, as in Example 2.8. Assume that $(\mathrm{P}_t)_{t\geq 0}$ satisfies $CD(0, n)$. Then for every $x, y \in M$, $t \geq 0$, $s > 0$ and for every non-negative function $f \in L^p(\mu)$, $p \in [1, +\infty]$, it holds*

$$\mathrm{P}_t\, f(x) \leq \Big(\frac{t+s}{s}\Big)^{n/2}\exp\Big(\frac{d^2(x,y)}{4s}\Big)\mathrm{P}_{t+s}\, f(y)$$

Proof. Given s, t as in the statement and $x, y \in M$, let $(x_u)_{u\in[t,t+s]}$ be a constant speed geodesic from x to y and for $u \in [t, t+s]$ let us define $\Lambda(u) := \mathrm{P}_u\, f(x_u)$. As natural in Γ-calculus, the next step is the computation of Λ', which goes as follows; for sake of a light notation, the evaluation at x_u will be omitted.

$$\begin{aligned}\Lambda'(u) &= \mathrm{P}_u(\mathrm{L}\, f) + \nabla\,\mathrm{P}_u\, f\cdot \dot{x}_u \geq \mathrm{P}_u(\mathrm{L}\, f) - |\nabla\,\mathrm{P}_u\, f|\frac{d(x,y)}{s} \\ &= \mathrm{P}_u(\mathrm{L}\, f) - \frac{|\nabla\,\mathrm{P}_u\, f|}{\sqrt{\mathrm{P}_u\, f}}\frac{d(x,y)}{s}\sqrt{\mathrm{P}_u\, f} \geq \mathrm{P}_u(\mathrm{L}\, f) - \frac{\Gamma(\mathrm{P}_u\, f)}{\mathrm{P}_u\, f} - \frac{d^2(x,y)}{4s^2}\,\mathrm{P}_u\, f\end{aligned}$$

$$\overset{(3.4)}{\geq} \quad -\frac{n}{2u}\,\mathrm{P}_u f - \frac{d^2(x,y)}{4s^2}\,\mathrm{P}_u f = -\frac{n}{2u}\Lambda(u) - \frac{d^2(x,y)}{4s^2}\Lambda(u)$$

For the first inequality, just observe that $\dot{x}_u = d(x,y)/s$ constantly and apply Cauchy-Schwarz inequality, while for the second one we used (3.2). Finally, the application of Li-Yau inequality is motivated by the $CD(0,n)$ assumption. Thus we have just shown that

$$\frac{\Lambda'(u)}{\Lambda(u)} \geq -\frac{n}{2u} - \frac{d^2(x,y)}{4s^2}$$

and, by integration on $[t,t+s]$ w.r.t. u, this yields

$$\log\frac{\Lambda(t+s)}{\Lambda(t)} \geq -\frac{n}{2}\log\frac{t+s}{s} - \frac{d^2(x,y)}{4s}$$

By applying the exponential to both sides, we get the desired conclusion. □

3.2.1 A simultaneous comparison

The disadvantage of parabolic Harnack's inequality is that the present can be controlled only in terms of a (strictly) future state, i.e. no simultaneous comparisons are possible. This restriction was overcome in 1996 by Wang, who found a suitable modification of Harnack's inequality; the result is the following.

Theorem 3.5 (Wang, 1996)**.** *Let (M,g) be a complete smooth Riemannian manifold equipped with the measure $\mu = e^{-V}$ vol, where vol denotes the volume measure, and let $(\mathrm{P}_t)_{t\geq 0}$ be the linear semigroup with generator $\mathrm{L} = \Delta_g - \nabla V\cdot\nabla$, as in Example 2.8; let $\rho\in\mathbb{R}$. Then the following facts are equivalent:*

(i) $CD(\rho,\infty)$ holds;

(ii) for every $x,y\in M$, $p\in\,]1,+\infty[$ and for every non-negative function $f\in L^p(E,\mu)$ we have

$$(\mathrm{P}_t f)^p(x) \leq \exp\left(\frac{p^\rho d^2(x,y)}{2(p-1)(e^{2\rho t}-1)}\right)\mathrm{P}_t(f^p)(y)$$

if $\rho\neq 0$,

$$(\mathrm{P}_t f)^p(x) \leq \exp\left(\frac{d^2(x,y)}{2(p-1)t}\right)\mathrm{P}_t(f^p)(y) \tag{3.9}$$

otherwise

Proof. For sake of simplicity, let $p=2$ and $\rho=0$, so that our aim will be to show that

$$\mathrm{P}_t(f^2)(x) \leq \exp\left(\frac{d^2(x,y)}{2t}\right)\mathrm{P}_t(f^2)(y)$$

for every $x,y\in M$. Fix $t\geq 0$, $x,y\in M$, $f\in L^2(E,\mu)$ and let $(x_s)_{s\in[0,t]}$ be a constant speed geodesic from x to y; for $s\in[0,t]$ let us define $\Lambda(s) := \mathrm{P}_s(g^2)(x_s)$, where $g := \mathrm{P}_{t-s} f$, and observe that

$$\begin{aligned}
\Lambda'(s) &= \mathrm{P}_s(\mathrm{L}\, g^2 - 2g\,\mathrm{L}\, g) + \nabla\,\mathrm{P}_s(g^2)\cdot\dot{x}_s = \mathrm{P}_s(2\Gamma(g)) + \nabla\,\mathrm{P}_s(g^2)\cdot\dot{x}_s \\
&\geq \mathrm{P}_s(2\Gamma(g)) - |\nabla\,\mathrm{P}_s(g^2)|\frac{d(x,y)}{t} \overset{(SC)}{\geq} \mathrm{P}_s(2\Gamma(g))\,\mathrm{P}_s(|\nabla g^2|)\frac{d(x,y)}{t} \\
&= 2\,\mathrm{P}_s\left(\left|\nabla g - \frac{1}{2}g\frac{d(x,y)}{t}\right|^2\right) - \frac{1}{2}\,\mathrm{P}_s(g^2)\frac{d^2(x,y)}{t^2} \geq -\frac{1}{2}\,\mathrm{P}_s(g^2)\frac{d^2(x,y)}{t^2}
\end{aligned}$$

$$= \quad -\frac{1}{2}\Lambda(s)\frac{d^2(x,y)}{t^2}$$

the evalutaion at x_s have being omitted for sake of notation, whence

$$\frac{\Lambda'(s)}{\Lambda(s)} \geq -\frac{1}{2}\frac{d^2(x,y)}{t^2}$$

By integration on $[0,t]$ we obtain

$$\log\frac{\Lambda(t)}{\Lambda(0)} \geq -\frac{d^2(x,y)}{2t}$$

and by applying the exponential at both sides we finally get

$$\Lambda(0) \leq \exp\left(\frac{d^2(x,y)}{2t}\right)\Lambda(t)$$

which is precisely the thesis by the very definition of Λ. For $p \neq 2$ the proof is exactly the same and thus the equivalence between $CD(0,\infty)$ and parabolic Wang's inequality in the flat space (3.9) is established. □

Let us stress the fact that the proof of parabolic Wang's inequality for $\rho \neq 0$ requires a bigger effort, because the choice of a constant speed geodesic connecting two given points $x, y \in M$ does not work any longer. Thus, a suitable curvature-dependent speed geodesic has to be identified, but we will not deal with the case $\rho \neq 0$.

3.3 Final remarks

As conclusion of this section, we would like to emphasize a few of connections between the functional inequalities we presented (e.g. the inverse local log-Sobolev, Li-Yau and Harnack's inequalities) and the curvature-dimension condition. First of all, come back to the proof of the Li-Yau inequality and notice that, by definition of Λ, (3.8) can be rewritten as

$$1 - \frac{2t}{n\,\mathrm{P}_t f}\left(\frac{\Gamma(\mathrm{P}_t f)}{\mathrm{P}_t f} - \mathrm{L}\,\mathrm{P}_t f\right) \geq \exp\left(-\frac{2}{n\,\mathrm{P}_t f}(\mathrm{Ent}_{\mathrm{P}_t}(f) - t\,\mathrm{L}\,\mathrm{P}_t f)\right)$$

so that if we apply the logarithm to both sides, then we get

$$\mathrm{Ent}_{\mathrm{P}_t}(f) \geq t\,\mathrm{L}\,\mathrm{P}_t f - \frac{n\,\mathrm{P}_t f}{2}\log\left(1 - \frac{2t}{n\,\mathrm{P}_t f}\left(\frac{\Gamma(\mathrm{P}_t f)}{\mathrm{P}_t f} - \mathrm{L}\,\mathrm{P}_t f\right)\right)$$

Since $CD(0,n)$ implies $CD(0,n')$ for any $n' \geq n$, the inequality above still holds if we replace n by any $n' \geq n$, so that we can pass to the limit and get

$$\mathrm{Ent}_{\mathrm{P}_t}(f) \geq t\frac{\Gamma(\mathrm{P}_t f)}{\mathrm{P}_t f}$$

which is the inverse local log-Sobolev inequality in the case $\rho = 0$. This is not particularly surprising, because, as just said, $CD(0,n)$ entails $CD(0,\infty)$ and by Theorem 2.16 the latter is known to be equivalent to the inverse local log-Sobolev inequality. The relevance of the remark lies in the fact that it enables us to appreciate the gap between the (inverse) local log-Sobolev inequality and Li-Yau's one and the reader can guess that the first is stronger than the latter, as it is actually the case. Roughly speaking, this comes from the fact that in order to get the Li-Yau inequality we simply neglect the right-hand side in (3.8).

As regards Li-Yau and parabolic Harnack's inequalities, we have shown that the former implies the latter, but the converse implication also holds. The rough idea is the following: for $x \in M$ and $w \in T_xM$, let $y(s) := \exp_x(sw)$ be the geodesic from x to $\exp_x(w)$; along such a curve, parabolic Harnack's inequality holds; take the first order approximation of $y(s)$ and let $s \downarrow 0$: the Li-Yau inequality then follows.

Finally, in connection with log-Sobolev and parabolic Wang's inequalities, it is worth talking about a further topic: the Gaussian concentration of mass property. For the Gaussian measure γ on $\mathbb{R}^n$, whence the property name is taken, it is a well-known fact that if $A \subset \mathbb{R}^n$ is a Borel set such that $\gamma(A) \geq 1/2$, then $\gamma(A_r) \geq 1 - e^{-r^2/2}$ for every $r \geq 0$, where A_r denotes the r-enlarged set of A, i.e.

$$A_r := \{x \in \mathbb{R}^n \,:\, d(x, A) < r\}$$

This means that as r increases, the measure of A_r converges exponentially fast to 1. On a Polish space E with a probability measure μ, a Gaussian concentration of mass occurs if there exist two constant $C, c > 0$ such that for any Borel set A with $\mu(A) \geq 1/2$ it holds

$$\mu(A_r) \geq 1 - Ce^{-cr^2}, \quad \forall r \geq 0$$

The link between this new inequality and the log-Sobolev one is easily stated: the latter implies the former, but the converse is not always true. Hence, under $CD(\rho, \infty)$ condition with $\rho > 0$ we already know by part (b) of Corollary 2.17 that a log-Sobolev inequality with parameter $2/\rho$ actually holds, whence the Gaussian concentration of mass property too. In the case $\rho = 0$ we can no longer invoke Corollary 2.17, but by means of parabolic Wang's inequality it is possible to show that the log-Sobolev inequality and the Gaussian concentration of mass property are actually equivalent.

4 Relationship with the Wasserstein distance

Aim of this last part is to introduce the optimal transport problem in an abstract framework and point out some remarkable connections with the Γ-calculus, but for the first goal some preliminary notions are required. As a first step, let us consider a Polish space (E, d) and let us define $\mathscr{P}_2(E)$ as the space of probability measures on E with finite second moment, namely those probability measures μ such that for at least one (and hence all) $x_0 \in E$ it holds

$$\int_E d^2(x, x_0) d\mu(x) < +\infty$$

Then, given $\mu_0, \mu_1 \in \mathscr{P}_2(E)$ let us introduce the set of admissible plans $\mathfrak{Adm}(\mu_0, \mu_1)$ as

$$\mathfrak{Adm}(\mu_0, \mu_1) := \{\pi \in \mathscr{P}(E \times E) \,:\, \pi_0 = \mu_0,\, \pi_1 = \mu_1\}$$

where

$$\pi_0(A) := \pi(A \times E), \quad \pi_1(B) := \pi(E \times B)$$

for any Borel sets $A, B \subset E$. With these premises, the *Wasserstein distance* between any couple of measures $\mu_0, \mu_1 \in \mathscr{P}_2(E)$ is defined as

$$W_2(\mu_0, \mu_1) := \inf_{\pi \in \mathfrak{Adm}(\mu_0, \mu_1)} \sqrt{\int_{E \times E} d^2(x, y) d\pi(x, y)}$$

and a first fundamental result is the following duality formula, although the original result of Kantorovich and Rubinstein only concerned the Euclidean case.

Theorem 4.1 (Kantorovich-Rubinstein, 1958)**.** *Let $\mu_0, \mu_1 \in \mathscr{P}_2(E)$. Then we have*

$$\frac{1}{2} W_2^2(\mu_0, \mu_1) = \sup_{\varphi \in C_b(E)} \left\{ \int_E Q\varphi \, d\mu_0 - \int_E \varphi \, d\mu_1 \right\}$$

where $C_b(E)$ denotes the space of continuous and bounded functions on E and $Q\varphi$ is the inf-convolution of φ, defined as

$$Q\varphi(x) := \inf_{y\in E}\left\{\varphi(y) + \frac{d^2(x,y)}{2}\right\}$$

The inf-convolution appearing in this theorem is nothing but a particular case of a more general one-parameter family of operators $(Q_t)_{t\geq 0}$ on $C_b(E)$, defined by $Q_0 := \mathrm{Id}$ and

$$Q_t\varphi(x) = \inf_{y\in\mathbb{R}^n}\left\{\varphi(y) + \frac{d^2(x,y)}{2t}\right\} \tag{4.1}$$

for $t \in]0,+\infty[$. In the literature, (4.1) is known as *Hopf-Lax formula* and its importance comes from the fact that by means of $(Q_t)_{t\geq 0}$ we are able to produce subsolutions of the Hamilton-Jacobi equation, in the sense that $u(t,x) := Q_t\varphi(x)$ satisfies

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{|\nabla u|^2}{2} \leq 0 \\ u|_{t=0} = \varphi \end{cases}$$

for any $\varphi \in C_b(E)$. Under the further assumption that E is a geodesic space, Hopf-Lax formula turns out to define a one-parameter semigroup and $(Q_t\varphi)_{t\geq 0}$ is a solution to the Hamilton-Jacobi equation with initial condition given by φ, i.e.

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{|\nabla u|^2}{2} = 0 \\ u|_{t=0} = \varphi \end{cases} \tag{4.2}$$

Pay attention to the fact that in the previous section we were working with parabolic equations, where the regularization effect was clear, while now we are dealing with the Hamilton-Jacobi equation, whose solutions are not regular at all in general and may present shocks. The validity of (4.2) is perfectly clear in the smooth case, while for a general Polish space E much more effort is required in order to give a meaning to $|\nabla f|$, but all the theory of first order calculus on metric measure spaces will be omitted, as the discussion is very technical and non-trivial at all. For this reason, although the definition of Wasserstein distance we gave and the sketchy introduction to optimal transport are completely abstract, from now on we will be exclusively concerned with the Riemannian setting, so that the reader can think of E as a smooth Riemannian manifold M. This choice is also motivated by the fact that in the Riemannian case Karl-Theodor Sturm and Max von Renesse proved that the curvature-dimension condition is equivalent to a commutation inequality involving the heat semigroup $(\mathrm{P}_t)_{t\geq 0}$ associated to the Laplace-Beltrami operator and the Wasserstein distance. Their celebrated result is thus the following.

Theorem 4.2 (Sturm-von Renesse, 2005)**.** *Let (M,g) be a smooth Riemannian manifold equipped with the volume measure and let $(\mathrm{P}_t)_{t\geq 0}$ be the heat semigroup w.r.t. the Laplace-Beltrami operator Δ_g; let $\rho \in \mathbb{R}$. Then the following facts are equivalent:*

(i) $CD(\rho,\infty)$ holds;

(ii) for every non-negative functions f, g such that

$$\int_M f\, d\,\mathrm{vol} = \int_M g\, d\,\mathrm{vol} = 1$$

and for every $t \geq 0$ it holds

$$W_2(\mathrm{P}_t\, f\,\mathrm{vol}, \mathrm{P}_t\, g\,\mathrm{vol}) \leq e^{-\rho t} W_2(f\,\mathrm{vol}, g\,\mathrm{vol}) \tag{4.3}$$

In order to prove this theorem, we first need a different commutation inequality between the heat and the Hopf-Lax semigroups, whose proof is based on a Γ-calculus argument.

Theorem 4.3 (Bakry-Gentil-Ledoux, Kuwada, 2015)**.** *In the same framework of Theorem 4.2, the following facts are equivalent:*

(i) $CD(\rho,\infty)$ *holds;*

(ii) for every $t,s\geq 0$ *and for every measurable function* f *we have*

$$\mathrm{P}_t(Q_s f)\leq Q_{e^{2\rho t}s}(\mathrm{P}_t\, f) \tag{4.4}$$

Proof. Let us prove the two implications, but first let us point out that, for sake of simplicity, the computations will be performed only for $\rho=0$; with some effort the reader can achieve the proof also in the curved case.

(i) ⇒ (ii) As a preliminary consideration, observe that if we replace f by λf with $\lambda>0$ in (4.4), then the inequality becomes

$$\mathrm{P}_t(Q_s(\lambda f))\leq Q_s(\lambda\,\mathrm{P}_t\, f)$$

and, since $Q_s(\lambda f)=\lambda Q_{s\lambda}f$, the latter is equivalent to

$$\lambda\,\mathrm{P}_t(Q_{s\lambda}f)\leq \lambda Q_{s\lambda}(\mathrm{P}_t\, f)$$

For this reason, choosing $\lambda=1/s$ it is sufficient to prove that $\mathrm{P}_t(Q_1 f)\leq Q_1(\mathrm{P}_t\, f)$ in order to get the conclusion. To this aim, let $t\geq 0$, $x,y\in M$ and let $(x_s)_{s\in[0,1]}$ be a constant speed geodesic from x to y; for $s\in[0,1]$, let us define $\Lambda(s):=\mathrm{P}_t(Q_s f)(x_s)$ and let us compute Λ', recalling that the Hopf-Lax formula provides us with a solution of the Hamilton-Jacobi equation.

$$\begin{aligned}\Lambda'(s) &= \mathrm{P}_t\left(-\frac{|\nabla Q_s f|^2}{2}\right)(x_s)+\nabla\,\mathrm{P}_t(Q_s f)\cdot\dot{x}_s\\ &\leq \mathrm{P}_t\left(-\frac{|\nabla Q_s f|^2}{2}\right)(x_s)+|\nabla\,\mathrm{P}_t(Q_s f)|(x_s)d(x,y)\\ &\leq \mathrm{P}_t\left(-\frac{|\nabla Q_s f|^2}{2}\right)(x_s)+\frac{|\nabla\,\mathrm{P}_t(Q_s f)|^2(x_s)}{2}+\frac{d^2(x,y)}{2}\leq\frac{d^2(x,y)}{2}\end{aligned}$$

The last inequality comes from the fact that by (WC) for $\rho=0$ we deduce

$$|\nabla\,\mathrm{P}_t(Q_s f)|^2\leq \mathrm{P}_t(|\nabla Q_s f|^2)$$

Hence $\Lambda'(s)\leq d^2(x,y)/2$ and, by integration, $\Lambda(1)-\Lambda(0)\leq d^2(x,y)/2$, whence

$$\mathrm{P}_t(Q_1 f)(y)\leq \mathrm{P}_t\, f(x)+\frac{d^2(x,y)}{2},\quad \forall x,y\in M$$

by definition of Λ. By taking the infimum over $x\in M$, we finally get $\mathrm{P}_t(Q_1 f)(y)\leq Q_1(\mathrm{P}_t\, f)(y)$ for every $y\in M$ and thus the conclusion.

(ii) ⇒ (i) Observe that equality holds in (4.4) for $s=0$, so that the inequality is preserved for the derivative evaluated at $s=0$. Taking again (4.2) into account, this means that

$$\mathrm{P}_t\left(-\frac{|\nabla f|^2}{2}\right)\leq -\frac{|\nabla\,\mathrm{P}_t\, f|^2}{2}$$

and this is precisely the weak commutation (WC), which is equivalent to $CD(0,\infty)$. □

By virtue of this result, we are finally ready to demonstrate Theorem 4.2.

Proof. As a preliminary consideration, notice that since f, g are probability densities w.r.t. vol, so are $\mathrm{P}_t f$ and $\mathrm{P}_t g$ by (H6): thus we can consider the Wasserstein distance between them. Now, let us prove the two implications, under the simplifying assumption that $\rho = 0$; the case $\rho \neq 0$ is left to the reader as an exercise.

(i) ⇒ (ii) Let $\varphi \in C_b(M)$ and notice that

$$\begin{aligned}\int_M Q_1\varphi \,\mathrm{P}_t f \, d\,\mathrm{vol} - \int_M \varphi \,\mathrm{P}_t g \, d\,\mathrm{vol} &\overset{(H8)}{=} \int_M \mathrm{P}_t(Q_1\varphi) f \, d\,\mathrm{vol} - \int_M (\mathrm{P}_t \varphi) g \, d\,\mathrm{vol} \\ &\overset{(4.4)}{\leq} \int_M Q_1(\mathrm{P}_t \varphi) f \, d\,\mathrm{vol} - \int_M (\mathrm{P}_t \varphi) g \, d\,\mathrm{vol} \\ &\leq \frac{1}{2} W_2^2(f\,\mathrm{vol}, g\,\mathrm{vol})\end{aligned}$$

where the last inequality is motivated by the Kantorovich-Rubinstein duality formula. By taking the supremum over $\varphi \in C_b(M)$ and by virtue of the same result, we finally get the contraction inequality (4.3).

(ii) ⇒ (i) From a heuristic point of view, observe that if we are allowed to take $f\,\mathrm{vol} = \delta_x$ and $g\,\mathrm{vol} = \delta_y$ for given $x, y \in M$, then the $CD(0, \infty)$ condition can be deduced as follows: as a first step

$$\begin{aligned}\mathrm{P}_t(Q_1\varphi)(x) - \mathrm{P}_t \varphi(y) &= \int_M \mathrm{P}_t(Q_1\varphi) f \, d\,\mathrm{vol} - \int_M (\mathrm{P}_t \varphi) g \, d\,\mathrm{vol} \\ &\overset{(H8)}{=} \int_M Q_1\varphi \,\mathrm{P}_t f \, d\,\mathrm{vol} - \int_M \varphi \,\mathrm{P}_t g \, d\,\mathrm{vol} \\ &\leq \frac{1}{2} W_2^2(f\,\mathrm{vol}, g\,\mathrm{vol}) = \frac{d^2(x,y)}{2}\end{aligned}$$

whence

$$\mathrm{P}_t(Q_1\varphi)(x) \leq \mathrm{P}_t \varphi(y) + \frac{d^2(x,y)}{2}$$

Now it is sufficient to take the supremum over $y \in M$ in order to make Q_1 appear at the right-hand side and then deduce $\mathrm{P}_t(Q_1\varphi) \leq Q_1(\mathrm{P}_t f)$, which is known to be equivalent to the $CD(0, \infty)$ condition by Theorem 4.3. However, it is not possible to choose f, g in such a way that $f\,\mathrm{vol} = \delta_x$ and $g\,\mathrm{vol} = \delta_y$, but this obstacle can be easily avoided by approximating δ_x and δ_y by probability measures with compactly supported smooth densities. □

4.1 Recent developments

As a final consideration, observe that also in the last part of these notes the dimension has not entered yet. Some results in this direction have appeared only recently and, for sake of completeness, we would like to present them (without proof), so that the reader can appreciate once more the important role played by the dimension. First of all, in the case of non-negative curvature a result of Bakry, Gentil and Ledoux on the one hand and Kuwada on the other hand allows non-simultaneous contraction estimates in the Wasserstein distance with a correction term depending on the dimension, as stated below.

Theorem 4.4 (Kuwada, 2013; Bakry-Gentil-Ledoux, 2015)**.** *Let $n \geq 1$. In the framework of Theorem 4.2 the following facts are equivalent:*

(i) $CD(0, n)$ holds;

(ii) for every non-negative functions f, g such that

$$\int_M f \, d\,\mathrm{vol} = \int_M g \, d\,\mathrm{vol} = 1$$

and for every $t, s \geq 0$ *it holds*

$$W_2^2(\mathrm{P}_t\, f \,\mathrm{vol}, \mathrm{P}_s\, g \,\mathrm{vol}) \leq W_2^2(f \,\mathrm{vol}, g \,\mathrm{vol}) + 2n(\sqrt{t} - \sqrt{s})^2$$

A generalization of such result was proposed by Kuwada in 2013.

Theorem 4.5 (Kuwada, 2013)**.** *Let* $n \geq 1$*. In the framework of Theorem 4.2 the following facts are equivalent:*

(i) $CD(\rho, n)$ *holds;*

(ii) for every non-negative functions f, g *such that*

$$\int_M f \, d\,\mathrm{vol} = \int_M g \, d\,\mathrm{vol} = 1$$

and for every $t, s \geq 0$ *it holds*

$$W_2^2(\mathrm{P}_t\, f \,\mathrm{vol}, \mathrm{P}_s\, g \,\mathrm{vol}) \leq A(s,t,\rho) W_2^2(f \,\mathrm{vol}, g \,\mathrm{vol}) + B(s,t,\rho,n)$$

for suitable non-negative functions A, B*.*

On the other hand, as regards simultaneous contraction estimates (in the same spirit of Theorem 4.2) with a dimensional correction term, two results are worth to be mentioned and they are due to François Bolley, Ivan Gentil and Arnaud Guillin. In the first one, a refined contraction estimate in the Wasserstein distance for the heat semigroup on $\mathbb{R}^n$ is considered, whereas in the second paper the Wasserstein distance W_2 is replaced by the so-called *Markov transportation distance* T_2; by means of this new distance on measures, the former contraction estimate is generalized to the Riemannian setting, as stated below.

Theorem 4.6. *Let* $(\mathrm{P}_t)_{t \geq 0}$ *be the heat semigroup w.r.t.* Δ *on* $\mathbb{R}^n$ *equipped with the Lebesgue measure* dx*. Then for every non-negative functions* f, g *such that*

$$\int_{\mathbb{R}^n} f \, dx = \int_{\mathbb{R}^n} g \, dx = 1$$

and for every $t \geq 0$ *it holds*

$$W_2^2(\mathrm{P}_t\, f dx, \mathrm{P}_t\, g dx) \leq W_2^2(f dx, g dx) - \frac{2}{n} \int_0^t \Big(\mathrm{Ent}_{dx}(\mathrm{P}_s\, f) - \mathrm{Ent}_{dx}(\mathrm{P}_s\, g) \Big)^2 ds$$

Theorem 4.7. *Let* $n \geq 1$*. In the framework of Theorem 4.2 the following facts are equivalent:*

(i) $CD(\rho, n)$ *holds;*

(ii) for every non-negative functions f, g *such that*

$$\int_M f \, d\,\mathrm{vol} = \int_M g \, d\,\mathrm{vol} = 1$$

and for every $t \geq 0$ *it holds*

$$T_2^2(\mathrm{P}_t\, f \,\mathrm{vol}, \mathrm{P}_t\, g \,\mathrm{vol}) \leq e^{-2\rho t} T_2^2(f \,\mathrm{vol}, g \,\mathrm{vol}) - \frac{2}{n} \int_0^t e^{-2\rho(t-s)} \Big(\mathrm{Ent}_{\mathrm{vol}}(\mathrm{P}_s\, f) - \mathrm{Ent}_{\mathrm{vol}}(\mathrm{P}_s\, g) \Big)^2 ds$$