γ_2 and Γ_2

in honour of Dominique Bakry and Michel Talagrand

Michel Ledoux

University of Toulouse, France

Abstract

This note is a wink to two of my mathematical heroes, Dominique Bakry and Michel Talagrand, linking in an anecdote example their famous Γ_2 and γ_2.

1 Introduction

My mathematical heroes, and collaborators and friends, Dominique Bakry and Michel Talagrand, introduced, in different areas, fundamental notions that they both named Γ_2 and γ_2. This note is dedicated to them, with a link between these two most impactful objects which gave rise to huge developments and applications (illustrated in particular in the monographs [1] and [8]).

The connection emphasized here is inspired by the works [4, 5] of E. Meckes on projections of random vectors and [6] of E. Meckes and M. Meckes on bounds on Kantorovich distances for empirical measures of random matrices. We thank them for this inspiring observation.

The note does not discuss any historical and technical aspects, and refers to the preceding reference books [1, 8] for details.

2 γ_2

The $\gamma_2(T, d)$ functional introduced by M. Talagrand is a measure of the size of a metric space (T, d) which was developed in the study of boundedness of random processes.

Given a set T, an admissible sequence is an increasing sequence $(\mathcal{A}_n)_{n\in\mathbb{N}}$ of partitions of T such that $\mathrm{Card}(\mathcal{A}_n) \leq 2^{2^n}$ for every $n \geq 1$ ($\mathrm{Card}(\mathcal{A}_0) \leq 1$). For every $t \in T$, denote by $A_n(t)$

the element of $\mathcal{A}_n$ which contains t. If d is a distance on T (not necessarily separating points), the diameter of $A_n(t)$ with respect to d is denoted by $D(A_n(t))$.

Let then

$$\gamma_2(T,d) \, = \, \inf \sup_{t \in T} \sum_{n \in \mathbb{N}} 2^{n/2} D\big(A_n(t)\big) \tag{1}$$

where the infimum is taken over all admissible sequences.

The $\gamma_2(T,d)$ functional extends and sharpens the earlier notions and tools of metric entropy and majorizing measures in the study of boundedness and continuity of random processes by the so-called chaining scheme (cf. [8]). It is actually an equivalent formulation of the majorizing measure bound introduced by my adviser X. Fernique [2]. On a probability space $(\Omega, \Sigma, \mathbb{P})$, let $X = (X_t)_{t \in T}$ be a family of real centered random variables (stochastic process), and set

$$M(X) \, = \, \sup \mathbb{E}\big(\sup_{t \in S} X_t \big)$$

where the supremum runs over all finite subsets $S \subset T$.

Say that X is sub-Gaussian with respect to the metric d if for all $s,t \in T$ and $u \geq 0$,

$$\mathbb{P}\Big(|X_s - X_t| \geq u \Big) \, \leq \, C \, \mathrm{e}^{-c\,u^2/2d(s,t)^2} \tag{2}$$

for some constants $C, c > 0$.

One first result in the chaining argument is that, for such a sub-Gaussian process,

$$M(X) \, \leq \, \frac{K}{\sqrt{c}} \, \gamma_2(T,d) \tag{3}$$

where $K = K(C)$ only depends on C.

This result applies in particular to a (centered) Gaussian process with the L^2-metric $d_X(s,t) = \mathbb{E}(|X_s - X_t|^2)^{1/2}$, $s,t \in T$, which satisfies (2) (with $C = 2$ and $c = 1$). But the remarkable achievement of M. Talagrand (1985) was actually to show that the upper-bound (3) may be reversed in this case

$$M(X) \, \geq \, \frac{1}{K} \, \gamma_2(T, d_X) \tag{4}$$

for some numerical $K > 0$.

It follows in particular that whenever $Y = (Y_t)_{t \in T}$ is a sub-Gaussian process with respect to the distance $d = d_X$ of a Gaussian process $X = (X_t)_{t \in T}$, then

$$M(Y) \, \leq \, \frac{K^2}{\sqrt{c}} \, M(X). \tag{5}$$

(It is of independent interest to find a direct proof of this comparison.)

3 Γ_2

The Γ_2 operator was introduced by D. Bakry in the mid-eighties in the study of Riesz transforms on manifolds and, with M. Émery, in the investigation of hypercontractive diffusions and logarithmic Sobolev inequalities as a functional tool to control the geometry and curvature of Laplacians and diffusion operators.

A Markov diffusion Triple (E, μ, Γ) in the sense of [1] consists of a state space E equipped with a diffusion semigroup $(P_t)_{t\geq 0}$ with infinitesimal generator L and carré du champ operator Γ, and invariant and reversible measure μ. The carré du champ operator Γ may be introduced from the generator L by

$$\Gamma(f,g) \,=\, \frac{1}{2}\big[\mathrm{L}(fg) - f\,\mathrm{L}g - g\,\mathrm{L}f\big]$$

for f, g in a suitable algebra $\mathcal{A}$ (of smooth functions), and they are both linked with the invariant measure μ by the integration by parts formula

$$\int_E f(-\mathrm{L}g)d\mu \,=\, \int_E \Gamma(f,g)d\mu.$$

The state space E may be endowed with an intrinsic distance for which Lipschitz functions f are such that $\Gamma(f)$ is bounded (μ-almost everywhere), which coincides with the Euclidean and Riemannian distance for Laplacians on Riemannian manifolds.

The Γ_2 operator is then defined by analogy with the carré du champ Γ as

$$\Gamma_2(f,g) \,=\, \frac{1}{2}\big[\mathrm{L}\big(\Gamma(f,g)\big) - \Gamma(f, \mathrm{L}g) - \Gamma(g, \mathrm{L}f)\big]$$

for all $f, g \in \mathcal{A}$. We write $\Gamma(f)$ for $\Gamma(f,f)$ and similarly for Γ_2. A curvature condition $CD(\rho, \infty)$ for some $\rho \in \mathbb{R}$ amounts to the condition

$$\Gamma_2(f) \,\geq\, \rho\,\Gamma(f), \quad f \in \mathcal{A}. \tag{6}$$

The Γ_2 operator and the associated curvature condition (6) stem from the Bochner identity for Laplacians on Riemannian manifolds, in which the $CD(\rho, \infty)$ condition expresses equivalently that the Ricci curvature of the manifold is bounded from below by ρ (in the sense of symmetric matrices). (The "∞" in the $CD(\rho, \infty)$ condition refers to an infinite dimension in the class of curvature-dimension condition $CD(\rho, n)$ not discussed here – see [1].)

The setting covers weighted Riemannian manifolds, and one prototypical example in this regard is the Gaussian model $E = \mathbb{R}^n$ equipped with the Euclidean gradient $\Gamma(f) = |\nabla f|^2$ for smooth functions $f : \mathbb{R}^n \to \mathbb{R}$ and the standard Gaussian measure $d\nu(x) = \mathrm{e}^{-|x|^2/2}\frac{dx}{(2\pi)^{n/2}}$. Here, and throughout the note, $|\cdot|$ denotes the Euclidean norm (and the associated metric). The underlying (Ornstein-Uhlenbeck) diffusion operator $\mathrm{L} = \Delta - x \cdot \nabla$ for which ν is the invariant measure gives rise to

$$\Gamma_2(f) \,=\, |\nabla^2 f|^2 + |\nabla f|^2 \,\geq\, |\nabla f|^2 \,=\, \Gamma(f)$$

for smooth $f : \mathbb{R}^n \to \mathbb{R}$. Thus, the Gaussian model is of curvature $CD(1, \infty)$. More generally, if $d\mu = \mathrm{e}^{-V}dx$ is a centered probability measure on the Borel sets with smooth potential

$V : \mathbb{R}^n \to \mathbb{R}$ such that $V(x) - \rho \frac{|x|^2}{2}$, $x \in \mathbb{R}^n$, is convex for some $\rho > 0$, the associated Markov Triple $E = \mathbb{R}^n$, $d\mu = \mathrm{e}^{-V} dx$, $\Gamma(f) = |\nabla f|^2$, has curvature $CD(\rho, \infty)$.

One significant application of the curvature $CD(\rho, \infty)$ condition is the logarithmic Sobolev inequality. Let (E, μ, Γ) be a Markov Triple with μ a probability measure satisfying the curvature condition $CD(\rho, \infty)$ for some $\rho > 0$. Then, μ satisfies the logarithmic Sobolev inequality

$$\int_E f \log f \, d\mu \,\leq\, \frac{1}{2\rho} \int_E \frac{\Gamma(f)}{f} \, d\mu \tag{7}$$

for every positive density f (in $\mathcal{A}$) with respect to μ. In particular, this logarithmic Sobolev inequality holds with constant $\rho = 1$ for the standard Gaussian measure ν on $\mathbb{R}^n$.

4 A link

Let (E, μ, Γ) be a Markov Triple with μ a probability measure satisfying the curvature condition $CD(\rho, \infty)$ for some $\rho > 0$.

From the logarithmic Sobolev inequality (7), a standard consequence known as the Herbst argument (cf. [1]) expresses that Lipschitz functions on (E, μ, Γ) have Gaussian tails. That is, if $F : E \to \mathbb{R}$ is Lipschitz with respect to Γ with Lipschitz coefficient $\|F\|_{\mathrm{Lip}} = \|\Gamma(f)\|_\infty$, then F is integrable and for any $u \geq 0$,

$$\mu\Big(\big| F - \textstyle\int_E F d\mu \big| \geq u \Big) \,\leq\, 2 \, \mathrm{e}^{-\rho \, u^2 / 2 \|F\|^2_{\mathrm{Lip}}}. \tag{8}$$

Given then a class $\mathcal{F}$ of Lipschitz functions on E, consider the (centered) process (on the probability space $(\Omega, \Sigma, \mathbb{P}) = (E, \mathcal{B}, \mu)$)

$$X_F \,=\, F - \textstyle\int_E F d\mu, \quad F \in \mathcal{F}.$$

By (8), for any $F, G \in \mathcal{F}$,

$$\mu\big(\big| X_F - X_G \big| \geq u \big) \,\leq\, 2 \, \mathrm{e}^{-\rho \, u^2 / 2 \|F - G\|^2_{\mathrm{Lip}}}, \quad u \geq 0.$$

The process $(X_F)_{F \in \mathcal{F}}$ is thus sub-Gaussian with respect to the Lipschitz metric

$$d_{\mathrm{Lip}}(F, G) \,=\, \|F - G\|_{\mathrm{Lip}}$$

on $\mathcal{F}$, and therefore by the upper-bound (3),

$$M_{\mathcal{F}} \,=\, \mathbb{E}\big(\sup_{F \in \mathcal{F}} X_F \big) \,\leq\, \frac{K}{\sqrt{\rho}} \, \gamma_2(\mathcal{F}, d_{\mathrm{Lip}})$$

where $K > 0$ is numerical.

As a consequence, it follows that the lower bound $\rho > 0$ in the Γ_2 curvature criterion $CD(\rho, \infty)$ is actually upper-bounded by the γ_2 functional of families of Lipschitz functions.

Theorem. *Let (E, μ, Γ) be a Markov Triple with μ a probability measure. Then,*

$$\sup \big\{ \sqrt{\rho}\, ; (E, \mu, \Gamma) \text{ satisfies } CD(\rho, \infty) \big\} \, \leq \, K \inf \frac{\gamma_2(\mathcal{F}, d_{\mathrm{Lip}})}{M_{\mathcal{F}}}$$

where the infimum runs over all classes $\mathcal{F}$ of Lipschitz functions on (E, μ, Γ).

5 An example

As emphasized in the introduction, a first application (which inspired this note) has been developed by E. Meckes and M. Meckes in the study of bounds on Kantorovich distances for spectral measures of random matrices [6]. We outline here a further, simpler illustration.

Let E be $\mathbb{R}^n$ with its Euclidean structure and $d\mu = \mathrm{e}^{-V} dx$ be centered probability measure such that $V(x) - \rho \frac{|x|^2}{2}$, $x \in \mathbb{R}^n$, is convex for some $\rho > 0$. The associated Markov Triple is thus of curvature $CD(\rho, \infty)$.

Let T be a subset of $\mathbb{R}^n$ and let $\mathcal{F}$ be the class of functions $F_t(x) = \langle t, x \rangle$, $x \in \mathbb{R}^n$, $t \in T$. Clearly, for $s, t \in T$,

$$d_{\mathrm{Lip}}(F_s, F_t) \, = \, \|F_s - F_t\|_{\mathrm{Lip}} \, = \, |s - t|.$$

Therefore, as a consequence of the theorem,

$$\int_{\mathbb{R}^n} \sup_{t \in T} \langle t, x \rangle d\mu(x) \, \leq \, \frac{K}{\sqrt{\rho}} \, \gamma_2\big(T, | \cdot |\big).$$

Now, if ν is the standard Gaussian measure on $\mathbb{R}^n$, it is also true from (4) that

$$\int_{\mathbb{R}^n} \sup_{t \in T} \langle t, x \rangle d\nu(x) \, \geq \, \frac{1}{K} \, \gamma_2\big(T, | \cdot |\big).$$

As a consequence (cf. (5)),

$$\int_{\mathbb{R}^n} \sup_{t \in T} \langle t, x \rangle d\mu(x) \, \leq \, \frac{K^2}{\sqrt{\rho}} \int_{\mathbb{R}^n} \sup_{t \in T} \langle t, x \rangle d\nu(x). \tag{9}$$

This result applies in particular to T the unit ball of the dual norm of a norm $\| \cdot \|$ on $\mathbb{R}^n$ so to yield

$$\int_{\mathbb{R}^n} \|x\| d\mu(x) \, \leq \, \frac{K^2}{\sqrt{\rho}} \int_{\mathbb{R}^n} \|x\| d\nu(x).$$

The inequality (9) is of course not new, and may be easily deduced, for example, from Caffarelli's contraction principle from optimal transport theory (cf. [9]). Namely, there is a Lipschitz map $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ with $\|\mathcal{T}\|_{\mathrm{Lip}} \leq \frac{1}{\sqrt{\rho}}$ pushing forward the Gaussian measure ν onto μ. Then, following [7], whenever $x, y \in \mathbb{R}^n$ and $x_\theta = x \sin\theta + y \cos\theta$,

$$\mathcal{T}(x) - \mathcal{T}(y) \, = \, \int_0^{\pi/2} \frac{d}{d\theta} \mathcal{T}(x_\theta) d\theta \, = \, \int_0^{\pi/2} \mathcal{T}'(x_\theta) \cdot x'_\theta \, d\theta.$$

Therefore (with T finite to start with),

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \sup_{t\in T} \langle t, \mathcal{T}(x) - \mathcal{T}(y)\rangle d\nu(x)d\nu(y) \;\leq\; \int_0^{\pi/2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \sup_{t\in T} \big\langle t, \mathcal{T}'(x_\theta)\cdot x'_\theta\big\rangle d\nu(x)d\nu(y)d\theta.$$

Since the couple (x_θ, x'_θ) has the same distribution as (x,y) under $d\nu(x)d\nu(y)$, it follows that

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \sup_{t\in T} \langle t, \mathcal{T}(x) - \mathcal{T}(y)\rangle d\nu(x)d\nu(y) \;\leq\; \frac{\pi}{2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \sup_{t\in T} \big\langle t, \mathcal{T}'(x)\cdot y\big\rangle d\nu(x)d\nu(y).$$

Now, for (almost) each x fixed, for every $t\in\mathbb{R}^n$,

$$\int_{\mathbb{R}^n} \big\langle t, \mathcal{T}'(x)\cdot y\big\rangle^2 d\nu(y) \;\leq\; \frac{1}{\rho}\int_{\mathbb{R}^n} \langle t, y\rangle^2 d\nu(y)$$

by the contraction property. Standard Gaussian comparison properties (cf. [3]) then ensure that

$$\int_{\mathbb{R}^n} \sup_{t\in T} \langle t, \mathcal{T}'(x)\cdot y\rangle d\nu(y) \;\leq\; \frac{1}{\sqrt{\rho}}\int_{\mathbb{R}^n} \sup_{t\in T} \langle t, y\rangle d\nu(y).$$

Finally, since μ is the push-forward of ν and is centered,

$$\begin{aligned}\int_{\mathbb{R}^n} \sup_{t\in T} \langle t, \mathcal{T}(x) - \mathcal{T}(y)\rangle d\nu(x)d\nu(y) \;&=\; \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \sup_{t\in T} \langle t, x-y\rangle d\mu(x)d\mu(y)\\ &\geq\; \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \sup_{t\in T} \langle t, x\rangle d\mu(x).\end{aligned}$$

The preceding argument may be applied similarly to spectral norms of matrices. For example, the (normalized) Haar measure on the special orthogonal group $\mathbb{SO}(N)$ has a lower bound on the Ricci curvature of order N and thus satisfy a logarithmic Sobolev inequality. Similarly, as discussed in [6], the Haar measure μ on $\mathbb{O}(N)$ satisfies by conditioning a logarithmic Sobolev inequality with constant of order $\frac{1}{N}$ (when equipped with the Hilbert-Schmidt metric). Let therefore $\mathcal{F}$ be the class of functions on $\mathbb{O}(N)$ given by

$$F_t(X) \;=\; \langle Xt, t\rangle, \quad t\in T,$$

where $T\subset\mathbb{R}^n$, for which, for the Hilbert-Schmidt norm on $\mathbb{O}(N)$,

$$d_{\mathrm{Lip}}(F_s, F_t) \;=\; \|F_s - F_t\|_{\mathrm{Lip}} \;=\; \bigg(\sum_{i,j=1}^N |s_is_j - t_it_j|^2\bigg)^{1/2}, \quad s,t\in T.$$

The theorem then shows that for the standard Gaussian measure ν on $\mathbb{R}^{n^2}$

$$\int_{\mathbb{O}^N} \sup_{t\in T} \langle Xt, t\rangle d\mu(X) \;\leq\; \frac{L}{\sqrt{N}}\int_{\mathbb{R}^{n^2}} \sup_{t\in T} \langle xt, t\rangle d\nu(x)$$

for some numerical constant $L>0$.

Similar developments take place for (symmetric) random matrices with log-concave densities leading to spectral comparison properties.

References

[1] D. Bakry, I. Gentil, M. Ledoux. *Analysis and geometry of Markov diffusion operators.* Grundlehren der mathematischen Wissenschaften 348. Springer (2014).

[2] X. Fernique. *Régularité des trajectoires des fonctions aléatoires gaussiennes.* École d'été de Probabilités de Saint-Flour, IV-1974, Lecture Notes in Math. 480, 1–96. Springer (1975).

[3] M. Ledoux, M. Talagrand. *Probability in Banach spaces. Isoperimetry and processes.* Ergebnisse der Mathematik und ihrer Grenzgebiete 23. Springer (1991).

[4] E. Meckes. Approximation of projections of random vectors. *J. Theoret. Probab.* 25, 333–352 (2012).

[5] E. Meckes. Projections of probability distributions: A measure-theoretic Dvoretzky theorem. *Geometric Aspects of Functional Analysis, Israel Seminar (GAFA) 2011-2013*, Springer Lecture Notes in Math. 2050, 385–394 (2012).

[6] E. Meckes, M. Meckes. Concentration and convergence rates for spectral measures of random matrices. *Probab. Theory Related Fields* 156, 145–164 (2013).

[7] G. Pisier. *Probabilistic methods in the geometry of Banach spaces.* Probability and Analysis, Varenna (Italy) 1985, Lecture Notes in Math. 1206, 167–241. Springer (1986).

[8] M. Talagrand. *Upper and lower bounds for stochastic processes: modern methods and classical problems.* Ergebnisse der Mathematik und ihrer Grenzgebiete 60. Springer (2014).

[9] C. Villani. *Optimal transport. Old and new.* Grundlehren der mathematischen Wissenschaften 338. Springer (2009).

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