A short overview of modern interplays between Functional Analysis, Geometry, Probability, and Statistics

Jordan Serres *

August 21, 2025

Contents

1 Geometry and Statistics **1**
1.1 Spaces with curvature bounded below 1
1.2 Statistical estimation in a geometric context 3

2 Stein's method and stability of functional inequalities **6**
2.1 Stein's method 6
2.2 Stability of functional inequalities 8
2.3 Beyond stability: Bubbling 11

3 Renormalization **12**
3.1 Renormalization and the Polchinski flow 12
3.2 Dynamical Γ-calculus 14

1 Geometry and Statistics

1.1 Spaces with curvature bounded below

A Riemannian structure is the specification of a means of measuring lengths. From a practical point of view, this can be considered as the choice of a system of units of measurement, but the mathematical abstraction goes a little further and allows this system of measurements to depend on where the measurement is actually taken. Of all those possible systems, the one we are familiar with, and in which the Pythagorean theorem we learned at school is valid, is called 'Euclidean', named after the Greek mathematician Euclid.

Given a smooth space with a Riemannian structure, the Riemann curvature tensor is a mathematical object that measures how far the local geometry deviates from Euclidean geometry. It is null if, and only if, the space is Euclidean, and moreover, it completely characterises the geometry, in the sense that, given the Riemann curvature tensor, we can completely recover the Riemannian metric structure. From a physical point of view, the Riemann curvature tensor at a point in space describes the tidal forces that would affect a physical body located at that point.

*Sorbonne Université, Laboratoire de Probabilités, Statistiques et Modélisation (LPSM), Paris, France, serres@lpsm.paris

By considering only the average effect of Riemann curvature, which mathematically consists of contracting the tensor, we obtain the Ricci curvature tensor, which has become a central concept since the formulation of general relativity, and in particular Einstein's revolutionary statement that the Ricci tensor of spacetime is equal to the energy-momentum tensor $\mathbf{T}$ compensated by half of its trace T times the metric g, up to Newton's universal constant G, hence asserting that spacetime is deformed by the matter it contains:

$$\mathrm{Ric} = 8\pi G \left(\mathbf{T} - \frac{1}{2} T g \right).$$

The property for a space to have a Ricci tensor bounded below by some $K \in \mathbb{R}$ therefore means in some sense that this space contains at least an amount K of energy, and has many implications, including the Bishop-Gromov inequality for the growth of the volume of its geodesic balls, the local Poincaré inequality, the local doubling condition, the Cheeger-Gromoll splitting theorem, Myers's theorem for its diameter, or the Lichnerowicz inequality for its fundamental frequency.

Moreover, given some $K \in \mathbb{R}$ and some $N \in \mathbb{N}$, the class $R_{K,N}$ of all Riemannian manifolds of dimension $n \leq N$ having a Ricci tensor bounded below by K, form a totally bounded space for the Gromov-Hausdorff topology. The theory of so-called Ricci-limit spaces, appearing from limit of sequence of manifolds in $R_{K,N}$, has been developped in the nineties by Cheeger and Colding, see e.g. [CC97].

Since the Gromov-Hausdorff topology does not assume the type of smoothness of the manifold structure, the search for the closure of $R_{K,N}$ is in some way equivalent to the search for a synthetic notion of curvature bounded below, that is a notion that does not involve any differential manifold structure.

To date, we do not yet have a synthetic notion of lower bounded Ricci curvature that corresponds exactly to the closure of $R_{K,N}$ with respect to the Gromov-Hausdorff topology. However, there are several very interesting synthetic notions, all of which correspond to a class strictly larger than the closure of $R_{K,N}$.

The class of spaces $CD(K,N)$ is certainly the most important historically. It was defined using optimal transport by Sturm [Stu06a, Stu06b] and Lott and Villani [LV09], as spaces with a geodesically convex entropy functional on the Wasserstein space W_2, which is the space of probability distributions with finite second moments, equipped with the L^2-optimal transport distance. Note that the letters CD stand for 'curvature-dimension'.

A more refined version is the class of $RCD(K,N)$ spaces, which were defined by Ambrosio-Gigli-Savaré [AGS15] as $CD(K,N)$ spaces having a quadratic Cheeger energy, a property known as infinitesimal Hilbertianity and ensuring that tangent spaces are Euclidean by asking the Sobolev space $W^{1,2}$ to be a Hilbert space. Note that here the letter R stands for Riemann.

Today, the class $RCD(K,N)$ seems to be the best way to look at the closure of $R_{K,N}$, even if both are not equal, see [BKMR21], and this is due to the large amount of research successfully done to show that the good properties of manifolds in $R_{K,N}$ are still true for RCD spaces. In particular, the Bishop-Gromov inequality and all theorems listed above remain true for $RCD(K,N)$ spaces. My contribution [BOS24] with Brunel and Ohta is part of this objective, generalizing Grünbaum's inequality, which is a classical result in the Minkowski theory of convex sets, to the setting of $RCD(0,N)$ spaces.

Several studies in this field have yielded results that were even new in the smooth case. For example, Lichnerowicz's quantitative inequality in the case of essentially non-branching $CD(K,N)$ spaces by Cavalletti, Mondino and Semola [CMS23], and its strengthening in the case of RCD spaces in my collaboration with Fathi and Gentil [FGS24].

For a very nice exposition on the topic of spaces with curvature bounded below, we refer the reader to Tewodrose's PhD thesis [Tew18]. Let us mention some other very interesting synthetic notion for Ricci curvature bounded below, including the Measure-Contraction property introduced by Ohta [Oht07], defined in terms of the contraction of

a measure on a set to a point; the Bishop-Gromov synthetic Ricci curvature introduced by Besson and Gallot [BG21], defined in terms of the growth of the volume of balls; or the barycenter curvature-dimension condition introduced by Han, Liu and Zhu in [HLZ25], in terms of Wasserstein barycenters.

From the point of view of RCD spaces, the notions of space and volume are formalized in the axiomatic of metric measure spaces. It is possible to understand those notions in a rather different manner by mean of Brownian motions. For more intuition about it, see my thesis manuscript [Ser22, Section 2.3]. The idea is not to understand a space in itself, but to understand a space through a stochastic motion within that space, usually named Brownian motion for historical reasons. There are two main rigorous formalizations of this intuition that are possible: one using Bakry-Émery theory, the other using the concept of Dirichlet forms.

The formalization by means of the Bakry-Émery theory is in fact equivalent to the description of RCD spaces given above, up to the technical assumption known as the Sobolev-to-Lipschitz property, see [Hon18]. For a brief introduction to the Bakry-Émery theory, see Section 3.2.

Formalization using Dirichlet spaces consists of starting from a notion of energy, called Dirichlet energy, and considering stochastic motion as a particle following the (random) trajectory that minimises this energy, see [FOT11]. This formalization makes it possible, in particular, to study singular spaces such as fractals, see for example [ARBC$^+$21]. Dirichlet and RCD spaces are closely related. This is actually proved by Suzuki's deep result [Suz19] stating that under the RCD condition, the pointed measure Gromov convergence is equivalent to the weak convergence of Brownian motions.

General relativity was the main starting point for the development of Ricci curvature theory, so it is natural to want to adapt the synthetic notion of curvature bounded below to the case of Lorentzian spaces, that is spacetime geometry. We refer the reader to Cavalletti and Mondino [CM22] for a review of this theory.

1.2 Statistical estimation in a geometric context

Although traditionally seen as rather distinct areas of mathematics, statistics and geometry have, in recent years, developed a rapidly growing interface, with many insights to be gained on both sides. At a high level, geometry enters statistics through the *structure* of the data.

It is clear that certain types of data naturally possess a geometric structure. A typical example arises in shape analysis, whose aim is to automatically detect similarities between the shapes of objects in a database, with applications such as face recognition. However, the striking feature of recent decades is that geometric data structures appear almost universally, as they are linked to the high dimensionality of data.

Statistical estimation has long been confronted with the challenge of high dimensionality. When data points $X_1, \dots, X_N$ lie in a space $\mathbb{R}^D$ with $D \gg 1$, a situation that frequently occurs in practice today, for instance when the X_i are audiovisual files with ambient dimension on the order of several millions, statistical estimation rapidly becomes computationally demanding. This difficulty is commonly known as the *curse of dimensionality*, a term introduced by Bellman in his 1957 book [Bel57], where he explicitly emphasized how the explosive growth of volume in high-dimensional spaces renders data increasingly sparse and isolated.

To illustrate how our low-dimensional intuitions (in dimensions 2 or 3) can fail in higher dimensions, let us mention Stein's celebrated paper [Ste56], in which he showed that in dimensions greater than three, the usual estimator of the mean is no longer optimal

in terms of quadratic risk. This phenomenon is now known as *Stein's paradox*.

When working with high-dimensional data, it is often unrealistic to assume that the data are truly generated by a process with as many degrees of freedom as the ambient dimension. Consider, for example, photographs of hands. Such images are naturally embedded in a Euclidean space whose dimension equals the number of pixels. However, it would be unreasonable to believe that a hand can vary freely in so many directions. In reality, the effective degrees of freedom are much closer to the number of joints in the hand, at most a few dozens, which is dramatically smaller than the number of pixels (on the order of 5 million).

Therefore, when dealing with high-dimensional data, the most natural way to model randomness is not through a *uniform measure*, as in lower-dimensional settings, but rather through a probability distribution supported on a lower-dimensional manifold. In this framework, the data $(X_1, \dots, X_n)$ are assumed to lie on a submanifold $\mathcal{M} \subset \mathbb{R}^D$. This perspective on high-dimensional data is now widely known as the *manifold hypothesis*, and it continues to be an active area of research in statistics. To mention just a few contributions: Belkin and Niyogi's 2009 conference [NB09], the work of Fefferman et al. [FMN16], which develops statistical tests to assess the validity of the manifold hypothesis; and, on a more playful note, the paper [CIDSZ08] showing that 3×3 image patches naturally form a Klein bottle.

Of course, real data are inevitably affected by noise, so the model suggested by the manifold hypothesis is more accurately described as "manifold + noise". From this viewpoint, several fundamental questions naturally arise, among which the three most prominent are:

- *Metric learning*, which consists in identifying the manifold $\mathcal{M}$ itself,
- *Clustering*, which focuses on detecting the connected (or nearly connected) components of $\mathcal{M}$,
- *Dimensionality reduction*, which seeks to embed the manifold $\mathcal{M}$ into a lower-dimensional space of dimension $d \ll D$, $\mathcal{M} \subset \mathcal{N}^d \subset \mathbb{R}^D$.

Let us expand a little further on the notion of dimensionality reduction. Linear techniques such as Principal Component Analysis (PCA) are well-established and widely used, but they suffer from a crucial drawback: they fail to capture the intrinsic structure of the data. This limitation becomes particularly problematic when working with highly structured data. In response, recent years have witnessed the development of *nonlinear* dimensionality-reduction methods, which essentially project the data onto a lower dimensional submanifold $\mathcal{N}^d \subset \mathbb{R}^D$. From this perspective, the problem of metric learning can be viewed as the most refined form of dimensionality reduction, since it amounts to recovering the manifold $\mathcal{M}$ itself. However, $\mathcal{M}$ cannot always be accurately approximated, and in practice it is often sufficient to identify another manifold $\mathcal{N}^d \subset \mathbb{R}^D$, as long as $d \ll D$. A wide variety of techniques have been proposed in this direction, notably Laplacian eigenmaps [BN03], ISOMAP [TSL00], and diffusion maps [CL06], to mention just a few. Let us also highlight the role of *autoencoders* [HS06], a class of neural networks specifically designed to learn a representation of data in a *latent space*, which can be interpreted as a compressed encoding of the original data while preserving its most relevant features.

In the area of deep learning, a particularly exciting connection between statistics and geometry arises through the *attention mechanism* of *transformers*. Roughly speaking, the procedure begins by dividing a text into small blocks, called *tokens*, which are then represented in a high-dimensional space so that tokens with similar usage in the text are mapped to nearby points. This representation is known as *word embedding*, and can

be adapted to data types beyond text. The attention mechanism, or more precisely the self-attention mechanism [VSP+17], then analyzes the importance of each token relative to all the others in the embedding space. In this way, the meaning assigned to each token is determined by its context, and is therefore inherently contextual.

Put in more intuitive terms, large language models demonstrate that the meaning of words can effectively be reduced to a *geometry of words*, and the self-attention mechanism provides a means of computing this geometry.

Another way to view the interplay between geometry and statistics is to observe that when certain information on a manifold is *learnable*, *i.e.*, can be efficiently approximated by statistical estimators, this reflects less a property of the estimators themselves than a statistical property of the underlying geometry. From this perspective, the most universal results seems to be Gromov's reconstruction theorem, which states that a metric measure space is uniquely determined by the law of the distance matrix between samples of arbitrary size [Gro07, Section $3\frac{1}{2}$]. See also [Ver04] for a proof based on the law of large numbers. More specific geometric properties also give rise to statistical consequences. In particular, non-negative lower bounds on the Ricci curvature translate into notions of *statistical depth*, while upper bounds on sectional curvatures lead to *finite-sample bounds*.

Statistical depth provides a way to quantify the centrality of a point with respect to a cloud of data. One of the most widely used notions is Tukey's depth [Tuk75], which is based on half-spaces and measures how difficult it is to separate a point from the rest of the data cloud. Intuitively, points with high depth in this sense play the role of *medians*, and this notion naturally leads to the construction of *robust* estimators [DG92]. In collaboration with Brunel and Ohta, we have shown that a non-negative lower bound on the Ricci curvature implies that the barycenters of convex sets are *deep* in a generalized sense of Tukey's depth, where half-spaces are replaced by horoballs, see [BOS24] for more details.

Let us expand a little on upper bounds for sectional curvature. Just as lower bounds on Ricci curvature can be characterized in a synthetic way, that is, without any reference to differentiability (see Section 1.1), upper (and likewise lower) bounds on sectional curvature can also be described synthetically by means of triangle comparison. This leads to what is known as Alexandrov geometry [AKP24]. The key idea is based on Toponogov's theorem, which states that if a Riemannian manifold has sectional curvature everywhere bounded above by some $\kappa \in \mathbb{R}$, then its geodesic triangles are *thinner* than the corresponding triangles in the model space of constant curvature κ: hyperbolic space if $\kappa < 0$, Euclidean space if $\kappa = 0$, and the sphere if $\kappa > 0$. Since the notion of a triangle makes sense even in the non-smooth setting of a geodesic metric space, *Alexandrov curvature* provides a way to define curvature bounds directly through triangle comparison. In particular, an upper bound on sectional curvature implies a local strong geodesic convexity of the squared distance to any point. In the case of nonpositive sectional curvature, this geodesic convexity is even global. Such convexity properties provide a natural framework to define and study the notion of a *mean* or *barycenter*. The most celebrated, and perhaps the most intuitive, statistical estimator in Euclidean space is the empirical mean, given by

$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

for an i.i.d data set $(X_1, \dots, X_n) \in (\mathbb{R}^d)^n$. One of the simplest properties of S_n is its linearity in the data points X_i. However, when the data lie in a non-linear space, $(X_1, \dots, X_n) \in \mathcal{M}^n$, the usual linear definition of the mean no longer applies. A natural way to generalize S_n to a non-linear setting is to observe that it is the least squares

estimator, that is,

$$S_n \in \operatorname*{argmin}_{p \in \mathcal{M}} \frac{1}{n} \sum_{i=1}^{n} d(p, X_i)^2 \tag{1}$$

Minimizing this quantity makes sense in any metric space, and the resulting points are known as *Fréchet means* [Fré48]. The Fréchet mean enjoys properties of uniqueness, consistency and asymptotic normality, which are closely related to upper bounds on the sectional curvature of the space through the geodesic convexity properties mentioned above, see [BP03, BP05]. An iterative procedure to approximate the minimizer of Problem (1) can be performed via a stochastic gradient-type descent, as introduced by Sturm [Stu03]. In the case of nonpositive sectional curvature, the resulting *inductive mean* b_n converges to the population barycenter b^*, and in particular,

$$\mathbb{E}\left[d(b^*, b_n)^2\right] \leq \frac{\sigma^2}{n},$$

where $\sigma^2 = \mathbb{E}\left[d(b^*, X_1)^2\right]$ is the variance. Moreover, this convergence can be strengthened into a Bernstein-type concentration bound: for all $\delta \in (0, 1)$, it holds with probability at least $1 - \delta$ that

$$d(b^*, b_n) \leq \frac{\sigma}{\sqrt{n}} + K\sqrt{\frac{\log(1/\delta)}{n}},$$

when X_1 is K^2-sub-Gaussian. See [BS24] for the case of nonpositive sectional curvature ($\kappa \leq 0$) and [BS25] for the case of a positive sectional curvature upper bound ($\kappa > 0$), where the analysis is more involved and the bounds are not yet sharp. Both results are obtained in collaboration with Brunel.

2 Stein's method and stability of functional inequalities

2.1 Stein's method

Stein's method is a set of techniques extensively developed from the seminal paper [Ste72] of Charles Stein in 1972. The aim of these techniques is to give quantitative bound for the distance between two probability measures. This method was introduced to quantify the asymptotic normality of certain statistical estimators, and proved extremely fruitful in mathematical statistics at a time when non-asymptotic bounds were establishing themselves as major theoretical guarantees for practical applications. We refer the reader to the survey by Ross [Ros11] which remains the most cited survey on Stein's method.

From a statistical perspective, the concept of distance between probability distributions serves as a bridge between asymptotic theory and finite sample theory. Asymptotic theory asserts that with a sufficiently large amount of data, one can accurately estimate the quantity of interest, commonly referred to as the estimand. However, "sufficient" can range from a manageable sample size to one that is practically unattainable. In contrast, finite sample theory focuses on determining the exact number of observations required to estimate the estimand with a predefined level of precision. Let us illustrate this with the example of the Central Limit Theorem (CLT). The CLT states that under mild assumptions, the empirical mean (the estimator) is asymptotically a Gaussian perturbation of the population mean (the estimand)

$$\frac{1}{n} \sum_{i=1}^{n} X_i \underset{n \to \infty}{\sim} \mathbb{E}[X] + \frac{1}{\sqrt{n}} \mathcal{N}(0, 1).$$

In this setting, distances between probability distributions are a natural concept to mathematically formalize what a level of precision is, and therefore enable us to replace the

asymptotic guarantee $\underset{n\to\infty}{\sim}$ by a quantitative bound of the form

$$\operatorname{dist}(\mu_n, \mathcal{N}(0,1)) \leq \frac{1}{\sqrt{n}},$$

where μ_n stands for the distribution of $Z_n = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i - \mathbb{E}[X]\right)$. In particular, if such a bound is available, we know that to achieve an error of at most 10^{-2}, it suffices to take a sample of size n such that $\sqrt{n} = 10^2$, i.e., $n = 10{,}000$.

From a probabilistic standpoint, the concept of distance between probability distributions has many applications, one of which is the analysis of mixing times in Markov chains. These chains represent stochastic processes where the immediate future depends only on the present state, not on the past. A classic example is shuffling a deck of cards using the riffle shuffle technique: after several repetitions, the deck becomes randomized. While continued shuffling still changes the card order at each step, the overall distribution remains uniform, indicating that the system has reached equilibrium. By analogy with this example, the mixing time refers to how long a Markov chain takes to reach such equilibrium. Most Markov chains do, in fact, admit an equilibrium, even if not all. From a more metaphysical perspective, this suggests that when time reaches infinity, only the past remains and there is no more "present" to influence the future. In that sense, the Markov property, which states that the future depends only on the present, implies that the process becomes constant: with no present left, change ceases, and the system settles permanently into its equilibrium. This makes it clear why the concept of distance between probability distributions is so valuable in this context: it provides a way to quantify how far the distribution of a Markov process at a given time is from its equilibrium distribution. In doing so, it offers a rigorous foundation for the definition of mixing time.

Stein's method takes a reversed perspective by using Markov chains to define a new notion of distance between probability distributions. The underlying heuristic is that if a Markov chain converges reliably to its equilibrium, then the time it takes to do so serves as a meaningful measure of the distance between its initial distribution and the equilibrium distribution. After all, if all roads lead to Rome, the time it takes to get there should give a good sense of how far we are, even if we are traveling at random. Let us illustrate it with the example of the Ornstein-Ulhenbeck process (OU). The OU process admits the Gaussian as unique equilibrium probability distribution, and has the advantage to allow almost all computations to be explicit. Let $(X_t)_{t\geq 0}$ be the OU process, μ be the initial distribution (that is the one of X_0), and γ be the Gaussian equilibrium (that is the one of X_∞). Consider a (real) test function h and look at the following function

$$f_h(x) = -\int_0^\infty \left(\mathbb{E}[h(X_t)|X_0 = x] - \mathbb{E}[h(X_\infty)]\right) dt.$$

If the process $(X_t)_t$ converges well, the integral above is finite, and the function f_h is well defined. This is indeed the case for the OU process. We can now see how this function f_h reflects the earlier heuristic: the integrand compares the distribution of X_t with the equilibrium distribution X_∞ at time t. Integrating this comparison over all times $t \in (0,\infty)$ amounts to tracing the entire trajectory of the process as it converges toward its equilibrium. Now, the crucial thing is that the function f_h satisfies the following ODE for all $x \in \mathbb{R}$,

$$f_h''(x) - x\, f_h'(x) = f_h(x) - \mathbb{E}[h(X_\infty)],$$

and moreover its second derivative is bounded with respect to h in the following manner

$$||f''||_\infty \leq 2\,||h'||_\infty.$$

Those two facts allows to derive inequalities of the following type: For all probability distributions μ,

$$W_1(\mu,\gamma) \leq \sup_{||f''||_\infty \leq 2} \mathbb{E}\left[f''(X_0) - X\, f'(X_0)\right] \tag{2}$$

where we denote by W_1 the Wasserstein L^1-transport distance, and by γ the standard normal distribution. The proof is very short, starting from Kantorovich dual formulation of the L^1 optimal transport problem, we write

$$\begin{aligned} W_1(\mu,\gamma) &= \sup_{||h'||_\infty \leq 1} \mathbb{E}[h(X_0) - h(X_\infty)] \\ &= \sup_{||h'||_\infty \leq 1} \mathbb{E}\left[f_h''(X_0) - X\, f_h'(X_0)\right] \\ &\leq \sup_{||f''||_\infty \leq 2} \mathbb{E}\left[f''(X_0) - X\, f'(X_0)\right] \end{aligned}$$

where we used the ODE solved by f_h at the second inequality, and the second derivative bound for the inequality at last line.

The supremum quantity appearing in Inequality (2) ultimately defines the new notion of distance we were seeking between any initial distribution μ and the Gaussian equilibrium γ. The philosophy underlying Stein's method can thus be reinterpreted as the idea that a differential operator can characterize a probability distribution, specifically, the operator $f \mapsto f'' - xf'$ in the case of the Gaussian. This approach is known as the *Barbour generator approach* [Bar90] and has led to the extension of the method to a wide range of probability distributions through the use of differential operators that generate Markov processes. In particular, to mention my own contributions, I have developed Stein's method for certain families of exponential-type probability distributions [Ser23b], and for Beta distributions in collaboration with Fathi and Gentil [FGS24]. I also introduced a version of Stein's method that applies to shapes, see [Ser24b]. This approach considers uniform probability distributions over domains $\Omega \subset \mathbb{R}^n$ rather than fully supported measures, thus embracing a more geometric perspective. In this setting, the notion of distance provided by Stein's method connects naturally with geometric notions of distance between shapes, such as Fraenkel asymmetry.

Let us note that, in practice, deriving inequalities of the type of (2) is only the first step in Stein's method. The second, and typically far more challenging step, is to bound the supremum appearing in the inequality. From a purely accounting viewpoint, the success of Stein's method can nonetheless be largely attributed to the fact that bounding such supremum terms is often easier than directly bounding distances like the Wasserstein distance. This is because the operator $f'' - xf'$ resembles the beginning of a Taylor expansion, thereby allowing one to leverage the powerful and well-established tools of classical calculus.

Stein's inequality (2) can also be seen as a transport cost inequality, and can therefore be paralleled with the Bobkov-Götze L^1-transport-cost inequality or Talagrand Inequality. Noticing this, Ledoux, Nourdin and Peccati [LNP15] interpreted the supremum in (2) as an entropy-like term they called the Stein discrepancy, and proved the HSI inequality improving Otto and Villani's famous HWI inequality [OV00] which connects the entropy, the Wasserstein-2 distance and the Fisher information.

2.2 Stability of functional inequalities

In general, almost philosophical terms, the question of stability can be formulated as follows: *If we have almost solved a problem, does that mean we are necessarily close to a real solution?* Intuitively, one might be inclined to answer affirmatively: for example, if one has nearly solved the equation $2x = 6$, this means that one has found a number x_0 such that $2x_0 \approx 6$, and consequently x_0 must be close to the exact solution $x = 3$. However, this notion of *stability* does not hold universally. There exist problems for which this property fails. Such problems may be regarded as 'ill-behaved' in the sense that the

lack of stability implies that approximating their solutions is inherently difficult, since being almost correct does not guarantee proximity to the actual solution. At least from a heuristic point of view, instability can arise from two reasons: a high degree of non-linearity, and/or a lack of compactness. A toy example illustrating the lack of compactness phenomenon could be the problem of minimizing the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x^2, & x \leq 1 \\ 1/x, & x > 1 \end{cases}$$

There is a unique solution to this problem, as the function attains its minimum only at 0. However, the problem lacks stability, due to the fact that $f(x) \to 0$ as $x \to +\infty$. As a consequence, any large number $\omega \gg 1$ constitutes an approximate solution, since $f(\omega) = 1/\omega$ is arbitrarily close to zero, yet ω is far from the true solution, which is 0.

It is important to emphasise that the problem of stability should not be confused with the problem of continuity, which poses the converse question, namely whether candidates close to a solution actually provide an approximate solution to the problem. It should also be noted that the question of stability is in fact a reverse problem: it seeks to reconstruct the cause (something close to a true solution) from the observation of its effects (an almost resolution of the problem).

The question of stability is really interesting for functional inequalities, which are problems that are infinite-dimensional and often have a geometric aspect. The most famous example is perhaps that of the isoperimetric inequality: among all shapes of fixed volume $v > 0$, the goal is to find the shapes with the smallest perimeter. It can be written

$$\inf_{|\Omega|=v} \mathcal{P}(\Omega)$$

where $\mathcal{P}$ is the perimeter functional. Up to translation, the extremizers are the balls of of volume v, and the isoperimetric inequality formalizes that fact:

$$\mathcal{P}(\Omega) \geq n \left|B_1\right|^{1/n} |\Omega|^{(n-1)/n}$$

where B_1 stands for the ball of radius 1. The stability problem can be approached in two ways: *qualitative* stability or *quantitative* stability. Qualitative stability consists of showing that if a sequence of domains Ω_n satisfies $\mathcal{P}(\Omega_n) \to \mathcal{P}(B)$, then (up to a subsequence), $\Omega_n \to B$ in an appropriate topology. This is nothing other than the Palais-Smale compactness condition. On the other hand, quantitative stability consists of being able to control exactly the distance between Ω and B given the deficit $\mathcal{P}(\Omega) - \mathcal{P}(B)$. Note that the distinction between qualitative and quantitative stability is exactly the same as the distinction between asymptotic theory and finite sample estimation theory that we mentioned in statistics, see section 2.1. Over the past twenty years or so, there is been growing interest in understanding the *quantitative* stability of functional inequalities. A good example of this attention is Alessio Figalli's talk at the Sixth European Congress of Mathematics [Fig13].

In the case of the isoperimetric inequality, the question boils down to whether the isoperimetric deficit controls a certain distance from the ball. In this case, the answer is yes, and it is now well known: It holds that

$$\mathcal{P}(\Omega) - n \left|B_1\right|^{1/n} |\Omega|^{(n-1)/n} \geq C_n \, |\Omega|^{(n-1)/n} \, \mathcal{A}(\Omega)^2$$

where $\mathcal{A}$ stands for the Fraenkel asymmetry, which measures the distance between Ω and the balls which are solutions, and C_n is a constant that depends only on the dimension n, see for example, Fusco survey [Fus15].

Many functional inequalities show quantitative stability, such as the Brunn-Minkowski inequality [FVHT23], the Faber-Krahn inequality [BPV15] or the Sobolev inequality [DEF+25], to name but a few, and even in the context of manifold, see for example the overview [Nob24].

Among many diverse proof techniques to study the stability of functional inequalities, let us mention:

- Entropy methods: A functional inequality often quantifies the rate at which a certain flow converges to equilibrium in terms of a distance, commonly referred to as "entropy" by analogy with physics. Therefore, if a function f is nearly optimal for the inequality, we gain insight into the convergence speed of the flow starting from f toward a true optimizer. By integrating this entropy over time, we can then expect to obtain a bound on the distance between f and the optimizer, see [Dol21].

- Symmetrization methods: Extremizers often exhibit a high degree of symmetry, for instance, balls in the case of the isoperimetric inequality. The underlying idea is that near-extremizers should inherit, at least approximately, these same symmetries, see [FMP08].

- Transportation methods: When it is possible to transport an almost extremizer onto a true minimizer, analyzing the fine properties of the transport map provides insight into how close the two are. This approach is particularly elegant in the case of the isoperimetric problem, where it refines Gromov's proof of the inequality by considering the optimal transport map between the uniform measures on the near-optimal set Ω and the ball B. This method involves highly nonlinear PDE analysis, notably the Monge–Ampère equation, which governs optimal transport maps, see [FMP10].

- The selection principle: This general principle asserts that a minimizing sequence for an optimization problem tends to "select" a dominant structure. In the context of the isoperimetric problem, this selected structure is typically close to a ball, see [CL12].

- The ABP method: Alexandroff-Bakelman-Pucci estimates (ABP) are L^∞ bounds for solutions of Poisson equations associated to linear second order elliptic operators. The method is similar to the transportation one, but the gradient of some solution of a linear Neumann equation is used instead of an usual transport map. ABP estimates then allow to get fine properties for that gradient, and deduce stability, see [CROS16].

Following Courtade and Fathi's ideas [CF20], Stein's method can also be used to obtain stability results for functional inequalities. The starting point is to say that extremizers are critical point, and so they satisfy the Euler-Lagrange equation. The idea is then to view the differential operator appearing in the Euler–Lagrange equation as a characterizing operator for the extremizers, in the spirit of Stein's method, see Section 2.1. If all goes well, a stability result follows. I used this type of idea during my PhD in the articles [Ser23b], which establish stability results for the Poincaré constant of the reversible distribution of a diffusion process; [Ser23a], which prove stability results for eigenvalues of any order of a one-dimensional diffusion; and [FGS24], where, in collaboration with Fathi and Gentil, we state a quantitative Lichnerowicz inequality in the framework of RCD spaces, for more details about this type of spaces, see Section 1.1. Note that, through the spectral interpretation of Poincaré's inequalities, the stability problem in this case can be seen as a relaxation of Kac's famous question: "Can one hear the shape of a drum?". In

particular, unlike Kac's question, this relaxed version admits a positive answer, as shown in the above-mentioned articles. Recently, I also provided a new proof of the stability of the Brock–Weinstock inequality (involving the first nonzero Steklov eigenvalue) by introducing a version of Stein's method for shapes; see [Ser24b]. For the original proof of the stability of the Brock–Weinstock inequality, based on a calibration technique, see [BDPR12].

2.3 Beyond stability: Bubbling

What happens when stability fails? In some cases, a finer phenomenon known as bubbling emerges. Bubbling occurs when a sum of "weakly interacting" solutions is almost a solution, without being an actual one. This phenomenon can only arise in nonlinear problems, due to the lack of superposition. Bubbling breaks stability, as such sums can be arbitrarily far from any true solution. However, it raises a deeper, and often technically challenging question: *Are almost solutions necessarily close to a sum of weakly interacting true solutions?*

Let us consider the example of the Sobolev inequality: For all $u \in H^1(\mathbb{R}^n)$,

$$||u||_{2^*} \leq S_n \, ||\nabla u||_2$$

where $2^* = \frac{2n}{n-2}$ is the sharp exponent, and S_n is the Sobolev constant. All extremal functions are given by the so-called Aubin-Talenti bubbles

$$f_{\sigma,b,x_0}(x) = \left(\sigma^2 + b|x - x_0|^2\right)^{-\frac{n-2}{2}}, \quad \sigma > 0,\, b > 0,\, x_0 \in \mathbb{R}^n.$$

As said before, this inequality satisfies stability, as [BE91] first showed that for all $u \in H^1(\mathbb{R}^n)$, it holds that

$$\inf_{\sigma,b>0,x_0\in\mathbb{R}^n} ||\nabla \left(u - f_{\sigma,b,x_0}\right)||_2^2 \leq C_n \left(S_n^2 \, ||\nabla u||_2^2 - ||u||_{2^*}^2\right)$$

where C_n is a dimenional constant. However, one may wish to go further and ask whether the Euler–Lagrange equation associated with the Sobolev inequality also exhibits stability. This equation is given by

$$\Delta u + u \, |u|^{2^*-2} = 0. \tag{3}$$

The Aubin–Talenti bubbles are the unique nonnegative solutions, however, sign-changing solutions also exist. The question is then whether stability holds when restricting to non-negative functions. The answer is negative. Indeed, since a bubble f_{σ,b,x_0} is concentrated around x_0, one can construct a sum of such functions, each centered at well-separated points x_i, and obtain a function that is close to solving the Euler-Lagrange equation. Such a sum is referred to as a *sum of weakly interacting bubbles*, and although it approximates a solution, it remains far from any true nonnegative one. More precisely, due to the concentration property of the Aubin–Talenti bubbles, the function is locally close to a single bubble f_{σ_i,b_i,x_i} near each point x_i, and thus locally close to achieving equality in the Euler-Lagrange equation. Away from all x_i, the sum is close to zero, and again nearly satisfies the equation. However, such a function cannot be close to a single bubble unless all the parameters σ_i, b_i and x_i are nearly identical. Therefore, the stability property is broken in this setting.

There is, nevertheless, a *bubbling phenomenon* that occurs: A nonnegative almost solution of (3) is close to a sum of weakly interacting Aubin-Talenti bubbles. As in the case of stability, there exists both a *qualitative* and a *quantitative* version of this result. The qualitative version dates back at least to [Str84]. The quantitative version is more recent, see [DSW25, FG20], and has the surprising feature that the rate at which the deficit

controls the distance to a sum of weakly interacting bubbles depends on the dimension.

The bubbling phenomenon appears extensively in the literature on nonlinear PDEs, see for example, [dP17]. Among many instances, one can mention the case of harmonic maps, where stability holds when restricted to maps of degree $k \geq 1$, but bubbling occurs for maps of degree $k \geq 2$, see [SU83, BC85, Top97, Top23, Rup23].

3 Renormalization

3.1 Renormalization and the Polchinski flow

Renormalization first appeared in quantum field theory in the 1950s as a set of techniques for circumventing the appearance of infinite quantities in continuum models. Let's present it briefly through the ϕ^4 model.

In classical mechanics, the *least action principle* states that the motion of a body can be described as the function of time that minimizes the action of a suitable Lagrangian, depending on the problem under consideration. The same idea applies in Lagrangian field theory: the field, for instance a scalar field ϕ of mass m moving in a potential $V(\phi) \geq 0$, minimizes the action

$$\int_{\mathbb{R}^d} dx \, \left(\frac{1}{2} |\nabla \phi(x)|^2 + m\phi(x)^2 - V(\phi(x)) \right).$$

Since any minimizer must in particular be a critical point of the action, it satisfies the associated Euler-Lagrange equation, which in this context is precisely Newton's second law of motion.

Now, in the context of *quantum uncertainty*, instead of determining *the* exact minimum of the action functional, one seeks it only up to random fluctuations. From a mathematical perspective, this is modeled by a probability distribution concentrated around the exact minimizers, with formal density given by

$$\mathcal{D}\phi \, \exp\left\{ - \int_{\mathbb{R}^d} dx \left[\frac{1}{2} |\nabla \phi(x)|^2 + m\phi(x)^2 + V(\phi(x)) \right] \right\} \tag{4}$$

where $\mathcal{D}\phi$ denotes the Feynman measure, formally corresponding to the infinite-dimensional Lebesgue measure, which is not mathematically well defined. Nevertheless, the formal density

$$\mathcal{D}\phi \, \exp\left\{ -\frac{1}{2} \int_{\mathbb{R}^d} dx \left(|\nabla \phi(x)|^2 + m\phi(x)^2 \right) \right\}$$

corresponds to the massive Gaussian free field, which is well defined. Hence, a natural approach to make (4) rigorous is to try to define it as the measure with density $e^{-V(\phi)}$, up to a normalizing constant, with respect to the Gaussian free field. In particular, the choice of the double-well potential $V(\phi) := g(\phi^2 - 1)^2$ for some coupling constant $g > 0$, is known as the ϕ^4 model.

From a mathematical perspective, the problem of giving a rigorous meaning to (4) is both deep and extensively studied. The difficulties stem from two main issues: first, the absence of an analogue of the Lebesgue measure on function spaces; second, the fact that the random fields one can define in this framework possess very low regularity. In fact, they are not pointwise-defined functions, but rather distributions, defined only on regions of space and not at individual points. As a consequence, algebraic expressions such as ϕ^2 will, a priori, not have any meaning.

One approach is to define (4) as the limiting law of a suitable stochastic process. This technique, known as *stochastic quantization*, was introduced by Nelson [Nel66] and by

Parisi and Wu [PY80]. Indeed, at least formaly, (4) is the equilibrium distribution of the dissipative dynamics governed by the following stochastic partial differential equation:

$$\partial_t \phi_t = (\Delta + m)\phi_t - \nabla V(\phi_t) + \xi \tag{5}$$

where ξ denotes a spacetime white noise. The presence of noise forces ϕ to be a generalized function, so the term $\nabla V(\phi_t) \sim \phi_t^3$ is not defined, which makes this SPDE singular. The *renormalization* procedure consists in subtracting divergent counterterms in order to isolate the non-singular part of the Lagrangian, which is the quantity with a physical meaning. Formally, one writes (4) as

$$\mathcal{D}\phi \exp\left\{ - \int_{\mathbb{R}^d} dx \left[\frac{1}{2} |\nabla \phi|^2 + m\phi^2 + g_{\text{eff}} (\phi^2 - 1)^2 \right] - \text{counterterms} \right\}$$

where the part $- \int_{\mathbb{R}^d} dx \left[\frac{1}{2} |\nabla \phi|^2 + g_{\text{eff}} (\phi^2 - 1)^2 \right]$ is the effective Lagrangian, with the effective coupling constant g_{eff} which is an experimentally observable quantity.

To date, the most advanced mathematical frameworks for determining when such counterterms exist are Hairer's theory of regularity structures [Hai14], which has a more algebraic flavour, and the theory of paracontrolled distributions developed by Gubinelli, Imkeller, and Perkowski [GIP15], which adopts a more analytical perspective.

To name just a few important advances from the history of Physics, let us cite Kadanoff block spin Renormalization scheme [Kad66] in statistical field theory, which consists in integrating ϕ up to some energy scale, i.e. dividing $\mathbb{R}^d$ in little blocs and integrating the fluctuations of ϕ at the level of the size of the blocs. In Fourier space, this corresponds to integrating high fequencies. Later, K. Wilson [Wil71], then Polchinski [Pol84] reformulated this procedure as an infinite system of differential equations on the parameter space, thus presenting a semigroup structure, referred to as the renormalization group. Among the many technical difficulties is the fact that this procedure creates coupling parameters at all orders from the very first step. For example, whereas the ϕ^4 model only has coupling constants up to order 4, renormalization immediately produces further couplings at all orders.

Among the others approaches to defining (4), let us also mention the variational method of Barashkov and Gubinelli [BG20], which is based on a variational stochastic representation formula for the Gaussian Free Field (GFF), in the spirit of the Boué-Dupuis formula. As a variational method, it is perhaps the closest to the least action principle and the original ideas from Lagrangian mechanics.

It is possible to give meaning to (4) without resorting to the abstract infinite-dimensional machinery described above. This is the *constructive approach*, which consists in replacing the continuum $\mathbb{R}^d$ with a discrete space given by a lattice $L\mathbb{T}^d \cap \varepsilon\mathbb{Z}^d$, where $\mathbb{T}^d$ denotes the torus of width $L > 0$ and $\varepsilon > 0$ is the energy cut-off, corresponding to the block size in Kadanoff's scheme. The continuum model (4) is then replaced by a well defined probability distribution on $\mathbb{R}^K$ where K is the number of sites in the lattice. The idea is that all properties of the lattice model that are independent of ε, are ipso-facto true for the continuum model. Under this constructive formulation, it is possible to show that the continuum model (4) is Gaussian for all $d \geq 4$ (the case $d = 4$ having recently been solved by Aizenmann and Duminil-Copin [ADC21]), but is in fact non-Gaussian for $d = 2$ and $d = 3$. Note that in dimension $d = 2$, the continuum ϕ^4 measure is absolutely continuous with respect to the GFF in finite volume ($L < \infty$), but becomes singular in infinite volume ($L \to \infty$). In dimension $d = 3$, it is always singular with respect to the GFF. Let us also mention that the case $d = 1$ is non-Gaussian but essentially trivial: since the GFF is then a continuous function (namely, Brownian motion), the equation defining the model

reduces to an SDE and is therefore well understood, in contrast with (5).

For lattice models, up to a normalizing constant, the density (4) take the form

$$\nu_0(d\phi) = \gamma_{C_\infty}(d\phi)\exp(-V_0(\phi))$$

where the centered Gaussian γ_{C_∞} is the discrete GFF. Recently, Bauerschmidt and Bodineau [BB21] have performed a renormalization group procedure on such models by decomposing the covariance C_∞ at each order of fluctuations. The procedure then consists in disintegrating ν_0 into a renormalized part ν_t and a fluctuation (random) part μ_t^φ. This is the analogue of (3.1) where the renormalized measure corresponds to the effective part of the Lagrangian. The renormalized measure obeys a Hamilton-Jacobi equation that can be read directly as the exact equation of the Polchinki renormalization group. Bauerschmidt and Bodineau used it to decompose the entropy of the original measure ν_0 and state a generalized Bakry-Émery criterion, which enabled them to prove a logarithmic Sobolev inequality for the continuum two-dimensional sine-Gordon model. The same method also allows to prove a log-Sobolev inequality [BD22] for the ϕ^4 model in dimension 2 and 3.

Note that the fluctation measures μ_t^φ are deeply connected with Eldan's stochastic localisation, which is a powerfull method to study convex analysis problems and to derive mixing bound for Markov chains, see [CE22].

The flow of renormalized measures $(\nu_t)_{t\geq 0}$, known as the Polchinski flow, has the added advantage of generalizing the well-established Bakry-Émery theory. Building on these ideas, I suggested to introduce a dynamic version of the Γ-calculus from Bakry-Émery theory to study the Polchinski flow, and more generally non-homogeneous flows, in the article [Ser24a], inspired by and extending [KP23, Section 2].

On a somewhat different topic, but to highlight the recent importance of the renormalization group in mathematics, let us mention the work of Armstrong, Kuusi, and Mourrat [AKM19], where they implement a rigorous renormalization scheme to obtain quantitative bounds in the stochastic homogenization of certain PDEs with random coefficients.

3.2 Dynamical Γ-calculus

Bakry and Émery [BE85] introduced the Γ_2 criterion as a sufficient condition to ensure the hypercontractivity of a Markov semigroup. This celebrated Γ-calculus introduced very powerful tools for studying properties of a Markov semigroup, such as logarithmic Sobolev inequalities, concentration of measure, mixing time etc.

The main ideas are as follows. Since what a Markov process does at the present moment depends solely on what it did at the very last moment, it follows that its future can be predicted (let's say stochastically) by knowledge of the process at an infinitesimal variation of time. The way in which the present can be predicted at the next instant is determined by the so-called infinitesimal generator. For the deterministic motion of a point, the infinitesimal generator would correspond to the velocity of that point, and hence be a first order differential operator.

In the case of a diffusion Markov process, i.e. one that propagates in space in the same way as heat diffuses in the material, the infinitesimal generator can be written as a second-order differential operator, taking the form

$$\mathcal{L} = \sum_{i,j} a_{i,j}\,\partial_{i,j} + \sum_i b_i\,\partial_i.$$

A natural thing to do with this second-order differential formula is to measure the extent to which it is not first-order, *i.e.* the extent to which the process is non-deterministic. Since

it would be of first-order if, and only if, it would satisfies Leibniz rule of differentiation, one can measure it by the following quantity

$$\Gamma(f,g) = \frac{1}{2}\left[\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f\right],$$

which is called the *carré du champ* operator because it boils down to $\Gamma(f,f) = |\nabla f|^2$ when $a_{i,j} = \delta_{i,j}$, denoting δ the Kronecker symbol (in French, "carré du champ" means field squared, talking about the gradient field ∇f). The carré du champ $\Gamma(f,f)$ is a quadratic first-order operator which, in average with respect to the equilibrium distribution μ, is equal to $-\mathcal{L}$:

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

So one may now want to measure the extent to which the carré du champ Γ commutes with the generator $\mathcal{L}$ by defining

$$\Gamma_2(f,g) = \frac{1}{2}\left[\mathcal{L}(\Gamma(f,g)) - \Gamma(f,\mathcal{L}g) - \Gamma(g,\mathcal{L}f)\right],$$

which is simply called the operator Γ_2. The quantity $\Gamma_2(f,f)$ measures the evolution of $\Gamma(f,f)$ along the dynamics. The remarkable fact is that nonnegative lower bounds on Γ_2, which can be therefore interpreted as preventing the energy from dissipating too quickly, have incredibly deep implications. This would not seem so mysterious, however, looking at Bochner's formula, linking Γ_2 to the Ricci curvature tensor, and allowing the Γ_2 criterion to be read as a lower bound on Ricci curvature, as was done in the seminal article [AGS15]. Note also that the fact that the Γ_2 operator and the Ricci curvature tensor are related is not so surprising, since both are measures of a commutation defect: The first is the commutation defect between the carré du champ operator and the generator of a Markov process, the second is the commutation defect between the covariant derivative with respect to two vector fields. What is generally referred to as Γ-calculus is a set of techniques aimed at showing stochastic, geometric or functional properties of a Markov diffusion using computable properties of the three objects defined above: $\mathcal{L}$, Γ and Γ_2. We refer the reader to [BGL14] for a detailed presentation.

Among many other developments, these tools have been extended into integrated criteria to study the global properties of Markov processes [CG23], and they have also been extended to the study of hypocoercive diffusions [Bau17].

A natural extension is to adapt the Γ-calculus to the context of a flow of probability distribution, in order to obtain some control over the dynamics of a functional inequality along the flow. Klartag and Putterman's work on the Poincaré constant along the heat flow [KP23] constitutes pioneering work in this vein. We should also mention the work of Roberto and Zegarlinski on the hypercontractivity in Orlicz spaces for non homogeneous diffusions [RZ22].

This is precisely the approach I have taken in [Ser24a], where I formulated and applied a dynamic Γ_2 criterion that reduces exactly to the original Bakry-Émery criterion in the case of a static flow. Furthermore, for the Polchinski flow of renormalized distributions, this dynamic criterion specializes to the multiscale Bakry-Émery criterion of Bauerschmidt and Bodineau [BB21].

References

[ADC21] Michael Aizenman and Hugo Duminil-Copin. Marginal triviality of the scaling limits of critical 4D Ising and ϕ^4_4 models. *Annals of Mathematics*, 194(1):163–235, 2021.

[AGS15] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Bakry–Émery curvature-dimension condition and Riemannian Ricci curvature bounds. *The Annals of Probability*, 43(1):339 – 404, 2015.

[AKM19] Scott Armstrong, Tuomo Kuusi, and Jean-Christophe Mourrat. *Quantitative stochastic homogenization and large-scale regularity*, volume 352. Springer, 2019.

[AKP24] Stephanie Alexander, Vitali Kapovitch, and Anton Petrunin. *Alexandrov geometry: foundations*, volume 236. American Mathematical Society, 2024.

[ARBC+21] Patricia Alonso-Ruiz, Fabrice Baudoin, Li Chen, Luke Rogers, Nageswari Shanmugalingam, and Alexander Teplyaev. Besov class via heat semigroup on Dirichlet spaces III: BV functions and sub-Gaussian heat kernel estimates. *Calculus of Variations and Partial Differential Equations*, 60(5):170, 2021.

[Bar90] A. D. Barbour. Stein's method for diffusion approximations. *Probab. Theory Related Fields*, 84(3):297–322, 1990.

[Bau17] Fabrice Baudoin. Bakry-Émery meet Villani. *Journal of functional analysis*, 273(7):2275–2291, 2017.

[BB21] Roland Bauerschmidt and Thierry Bodineau. Log-Sobolev inequality for the continuum sine-Gordon model. *Communications on Pure and Applied Mathematics*, 74(10):2064–2113, 2021.

[BC85] Haim Brezis and Jean-Michel Coron. Convergence of solutions of H-systems or how to blow bubbles. *Archive for rational mechanics and analysis*, 89(1):21–56, 1985.

[BD22] Roland Bauerschmidt and Benoit Dagallier. Log-Sobolev inequality for the φ_2^4, and φ_3^4 measures. *Communications on Pure and Applied Mathematics*, 2022.

[BDPR12] Lorenzo Brasco, Guido De Philippis, and Berardo Ruffini. Spectral optimization for the Stekloff–Laplacian: the stability issue. *Journal of Functional Analysis*, 262(11):4675–4710, 2012.

[BE85] Dominique Bakry and Michel Émery. Diffusions hypercontractives. *Séminaire de probabilités de Strasbourg*, 19:177–206, 1985.

[BE91] Gabriele Bianchi and Henrik Egnell. A note on the Sobolev inequality. *Journal of functional analysis*, 100(1):18–24, 1991.

[Bel57] Richard Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, 1957.

[BG20] N. Barashkov and M. Gubinelli. A variational method for Φ_3^4. *Duke Mathematical Journal*, 169(17):3339 – 3415, 2020.

[BG21] Gerard Besson and Sylvestre Gallot. On scalar and Ricci curvatures. *SIGMA. Symmetry, Integrability and Geometry: Methods and Applications*, 17:046, 2021.

[BGL14] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 103. Springer, 2014.

[BKMR21] Jérôme Bertrand, Christian Ketterer, Ilaria Mondello, and Thomas Richard. Stratified spaces and synthetic Ricci curvature bounds. In *Annales de l'Institut Fourier*, volume 71, pages 123–173, 2021.

[BN03] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

[BOS24] Victor-Emmanuel Brunel, Shin-ichi Ohta, and Jordan Serres. A generalization of Grünbaum's inequality in RCD(0,N)-spaces. *arXiv preprint arXiv:2408.15030*, 2024.

[BP03] Rabi Bhattacharya and Vic Patrangenaru. Large sample theory of intrinsic and extrinsic sample means on manifolds. *The Annals of Statistics*, 31(1):1–29, 2003.

[BP05] Rabi Bhattacharya and Vic Patrangenaru. Large sample theory of intrinsic and extrinsic sample means on manifolds—II. *The Annals of Statistics*, 33(3):1225 – 1259, 2005.

[BPV15] Lorenzo Brasco, Guido De Philippis, and Bozhidar Velichkov. Faber–Krahn inequalities in sharp quantitative form. *Duke Mathematical Journal*, 164(9):1777 – 1831, 2015.

[BS24] Victor-Emmanuel Brunel and Jordan Serres. Concentration of empirical barycenters in metric spaces. In *Proceedings of The 35th International Conference on Algorithmic Learning Theory*, volume 237 of *Proceedings of Machine Learning Research*, pages 337–361. PMLR, 25–28 Feb 2024.

[BS25] Victor-Emmanuel Brunel and Jordan Serres. Finite sample bounds for barycenter estimation in geodesic spaces. *arXiv preprint arXiv:2502.14069*, 2025.

[CC97] Jeff Cheeger and Tobias H Colding. On the structure of spaces with ricci curvature bounded below. i. *Journal of Differential Geometry*, 46(3):406–480, 1997.

[CE22] Yuansi Chen and Ronen Eldan. Localization schemes: A framework for proving mixing bounds for Markov chains. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 110–122. IEEE, 2022.

[CF20] Thomas A. Courtade and Max Fathi. Stability of the Bakry-Émery theorem on $\mathbb{R}^n$. *J. Funct. Anal.*, 279(2):108523, 28, 2020.

[CG23] Patrick Cattiaux and Arnaud Guillin. A journey with the integrated γ_2 criterion and its weak forms. In *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 2020-2022*, pages 167–208. Springer, 2023.

[CIDSZ08] Gunnar Carlsson, Tigran Ishkhanov, Vin De Silva, and Afra Zomorodian. On the local behavior of spaces of natural images. *International journal of computer vision*, 76(1):1–12, 2008.

[CL06] Ronald R Coifman and Stéphane Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1):5–30, 2006.

[CL12] Marco Cicalese and Gian Paolo Leonardi. A selection principle for the sharp quantitative isoperimetric inequality. *Archive for Rational Mechanics and Analysis*, 206(2):617–643, 2012.

[CM22] Fabio Cavalletti and Andrea Mondino. A review of Lorentzian synthetic theory of timelike Ricci curvature bounds. *General Relativity and Gravitation*, 54(11):137, 2022.

[CMS23] Fabio Cavalletti, Andrea Mondino, and Daniele Semola. Quantitative Obata's theorem. *Analysis & PDE*, 16(6):1389–1431, 2023.

[CROS16] Xavier Cabré, Xavier Ros-Oton, and Joaquim Serra. Sharp isoperimetric inequalities via the ABP method. *Journal of the European Mathematical Society (EMS Publishing)*, 18(12), 2016.

[DEF+25] Jean Dolbeault, Maria Esteban, Alessio Figalli, Rupert Frank, and Michael Loss. Sharp stability for Sobolev and log-Sobolev inequalities, with optimal dimensional dependence. *Cambridge Journal of Mathematics*, 13:359–430, 01 2025.

[DG92] David L. Donoho and Miriam Gasko. Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *The Annals of Statistics*, 20(4):1803–1827, 1992.

[Dol21] Jean Dolbeault. Functional inequalities: nonlinear flows and entropy methods as a tool for obtaining sharp and constructive results. *Milan Journal of Mathematics*, 89(2):355–386, 2021.

[dP17] Manuel del Pino. Bubbling blow-up in critical parabolic problems. In *Nonlocal and Nonlinear Diffusions and Interactions: New Methods and Directions: Cetraro, Italy 2016*, pages 73–116. Springer, 2017.

[DSW25] Bin Deng, Liming Sun, and Jun-cheng Wei. Sharp quantitative estimates of Struwe's decomposition. *Duke Mathematical Journal*, 174(1):159–228, 2025.

[FG20] Alessio Figalli and Federico Glaudo. On the sharp stability of critical points of the Sobolev inequality. *Archive for rational mechanics and analysis*, 237(1):201–258, 2020.

[FGS24] Max Fathi, Ivan Gentil, and Jordan Serres. Stability estimates for the sharp spectral gap bound under a curvature-dimension condition. *Annales de l'Institut Fourier*, 74(6):2425–2459, 2024.

[Fig13] Alessio Figalli. Stability in geometric and functional inequalities. In *European congress of mathematics*, pages 585–599. Citeseer, 2013.

[FMN16] Charles Fefferman, Sanjoy Mitter, and Hariharan Narayanan. Testing the manifold hypothesis. *Journal of the American Mathematical Society*, 29(4):983–1049, 2016.

[FMP08] Nicola Fusco, Francesco Maggi, and Aldo Pratelli. The sharp quantitative isoperimetric inequality. *Annals of mathematics*, pages 941–980, 2008.

[FMP10] Alessio Figalli, Francesco Maggi, and Aldo Pratelli. A mass transportation approach to quantitative isoperimetric inequalities. *Inventiones mathematicae*, 182(1):167–211, 2010.

[FOT11] Masatoshi Fukushima, Yoichi Oshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19. Walter de Gruyter, 2011.

[Fré48] Maurice Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de l'Institut Henri Poincaré*, 10(4):215–310, 1948.

[Fus15] Nicola Fusco. The quantitative isoperimetric inequality and related topics. *Bulletin of Mathematical Sciences*, 5:517–607, 2015.

[FVHT23] Alessio Figalli, Peter Van-Hintum, and Marius Tiba. Sharp quantitative stability of the Brunn-Minkowski inequality. *arXiv preprint arXiv:2310.20643*, 2023.

[GIP15] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular pdes. In *Forum of Mathematics, Pi*, volume 3, page e6. Cambridge University Press, 2015.

[Gro07] Mikhail Gromov. *Metric structures for Riemannian and non-Riemannian spaces*. Springer, 2007.

[Hai14] Martin Hairer. A theory of regularity structures. *Inventiones mathematicae*, 198(2):269–504, 2014.

[HLZ25] Bang-Xian Han, Deng-Yu Liu, and Zhuo-Nan Zhu. Barycenter curvature-dimension condition for extended metric measure spaces. *Indagationes Mathematicae*, 2025.

[Hon18] Shouhei Honda. Bakry-Émery conditions on almost smooth metric measure spaces. *Analysis and Geometry in Metric Spaces*, 6(1):129–145, 2018.

[HS06] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.

[Kad66] Leo P Kadanoff. Scaling laws for ising models near t c. *Physics Physique Fizika*, 2(6):263, 1966.

[KP23] Bo'az Klartag and Eli Putterman. Spectral monotonicity under Gaussian convolution. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, Ser. 6, 32(5):939–967, 2023.

[LNP15] Michel Ledoux, Ivan Nourdin, and Giovanni Peccati. Stein's method, logarithmic Sobolev and transport inequalities. *Geom. Funct. Anal.*, 25(1):256–306, 2015.

[LV09] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Annals of Mathematics*, pages 903–991, 2009.

[NB09] Partha Niyogi and Mikhail Belkin. Geometric methods and manifold learning. Tutorial lecture at the Machine Learning Summer School (MLSS 2009), Chicago, June 2009. Event dates: June 1–11, 2009. Video available at VideoLectures.NET.

[Nel66] Edward Nelson. Derivation of the Schrödinger equation from Newtonian mechanics. *Physical review*, 150(4):1079, 1966.

[Nob24] Francesco Nobili. An overview of the stability of Sobolev inequalities on Riemannian manifolds with Ricci lower bounds , 2024. cvgmt preprint.

[Oht07] Shin-ichi Ohta. On the measure contraction property of metric measure spaces. *Comment. Math. Helv.*, 82(4):805–828, 2007.

[OV00] Felix Otto and Cédric Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *Journal of Functional Analysis*, 173(2):361–400, 2000.

[Pol84] Joseph Polchinski. Renormalization and effective Lagrangians. *Nuclear Physics B*, 231(2):269–295, 1984.

[PY80] G Parisi and Wu Yongshi. *Perturbation theory without gauge fixing*, volume 7. July, 1980.

[Ros11] Nathan Ross. Fundamentals of Stein's method. *Probab. Surv.*, 8:210–293, 2011.

[Rup23] Melanie Rupflin. Sharp quantitative rigidity results for maps from $\mathbb{S}^2$ to $\mathbb{S}^2$ of general degree. *arXiv preprint arXiv:2305.17045*, 2023.

[RZ22] C Roberto and B Zegarlinski. Hypercontractivity for markov semi-groups. *Journal of Functional Analysis*, 282(12):109439, 2022.

[Ser22] Jordan Serres. *Around stability for functional inequalities*. PhD thesis, Université Paul Sabatier-Toulouse III, 2022.

[Ser23a] Jordan Serres. Stability of higher order eigenvalues in dimension one. *Stochastic Processes and their Applications*, 155:459–484, 2023.

[Ser23b] Jordan Serres. Stability of the Poincaré constant. *Bernoulli*, 29(2):1297–1320, 2023.

[Ser24a] Jordan Serres. Behavior of the poincaré constant along the polchinski renormalization flow. *Communications in Contemporary Mathematics*, 26(07):2350035, 2024.

[Ser24b] Jordan Serres. Isoperimetric and geometric inequalities in quantitative form: Stein's method approach. *arXiv preprint arXiv:2410.20844*, 2024.

[Ste56] Charles Stein. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the third Berkeley symposium on mathematical statistics and probability, volume 1: Contributions to the theory of statistics*, volume 3, pages 197–207. University of California Press, 1956.

[Ste72] Charles Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability , Vol. II: Probability theory*, pages 583–602, 1972.

[Str84] Michael Struwe. A global compactness result for elliptic boundary value problems involving limiting nonlinearities. *Mathematische zeitschrift*, 187(4):511–517, 1984.

[Stu03] Karl-Theodor Sturm. Probability measures on metric spaces of nonpositive. *Heat kernels and analysis on manifolds, graphs, and metric spaces*, 338:357, 2003.

[Stu06a] Karl-Theodor Sturm. On the geometry of metric measure spaces. I. *Acta Mathematica*, 196(1):65 – 131, 2006.

[Stu06b] Karl-Theodor Sturm. On the geometry of metric measure spaces. II. *Acta Mathematica*, 196(1):133 – 177, 2006.

[SU83] Richard Schoen and Karen Uhlenbeck. Boundary regularity and the dirichlet problem for harmonic maps. *Journal of Differential Geometry*, 18(2):253–268, 1983.

[Suz19] Kohei Suzuki. Convergence of Brownian motions on metric measure spaces under Riemannian Curvature–Dimension conditions. *Electronic Journal of Probability*, 24(none):1 – 36, 2019.

[Tew18] David Tewodrose. *Some functional inequalities and spectral properties of metric measure spaces with curvature bounded below*. PhD thesis, Université Paris sciences et lettres; Scuola normale superiore (Pise, Italie), 2018.

[Top97] Peter Miles Topping. Rigidity in the harmonic map heat flow. *Journal of Differential Geometry*, 45(3):593–610, 1997.

[Top23] Peter M Topping. A rigidity estimate for maps from $\mathbb{S}^2$ to $\mathbb{S}^2$ via the harmonic map flow. *Bulletin of the London Mathematical Society*, 55(1):338–343, 2023.

[TSL00] Joshua B Tenenbaum, Vin de Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.

[Tuk75] John W Tukey. Mathematics and the picturing of data. In *Proceedings of the international congress of mathematicians*, volume 2, pages 523–531. Vancouver, 1975.

[Ver04] Anatoly M Vershik. Random and universal metric spaces. *Nonlinear Phenomena and complex systems*, 10:199–228, 2004.

[VSP+17] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[Wil71] Kenneth G Wilson. Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture. *Physical review B*, 4(9):3174, 1971.