Isoperimetric inequalities in high-dimensional convex sets

Lecture notes by Bo’az Klartag*

Institut Henri Poincaré (IHP), Paris, May 21–24, 2024

Contents

1 The Poincaré inequality

Even if we were not hosted by an institution that honors Poincaré, a good starting point for these lectures would be the mathematical inequality that carries his name. It was published by Poincaré in 1892–1896 in the case where the dimension is 2 or 3, and the measure μ is the uniform probability measure on the convex body K.

Recall that an absolutely-continuous probability measure μ in $\mathbb{R}^n$ is *log-concave* if its density ρ satisfies

$$\rho(\lambda x + (1-\lambda)y) \geq \rho(x)^{\lambda}\rho(y)^{1-\lambda} \qquad (x, y \in \mathbb{R}^n, 0 < \lambda < 1). \tag{1}$$

A probability measure μ in $\mathbb{R}^n$ is log-concave if it is supported in an affine subspace and has a log-concave density in this subspace. The uniform probability measure on a convex body is log-concave, as well as all Gaussian measures.

Theorem 1 (“The Poincaré inequality”). *Let $K \subseteq \mathbb{R}^n$ be a convex body, let μ be a log-concave probability measure on K. Then for any C^1-smooth function $f : K \to \mathbb{R}$ with $\int_K f d\mu = 0$,*

$$\int_K f^2 d\mu \leq C_P(\mu) \cdot \int_K |\nabla f|^2 d\mu \tag{2}$$

where $C_P(\mu) \leq C_n \cdot Diam^2(K)$, and $C_n > 0$ depends only on the dimension n.

*draft, version of May 27, 2024

Here $Diam(K) = \sup_{x,y\in K} |x-y|$ is the diameter of K and $|\cdot|$ is the standard Euclidean norm in $\mathbb{R}^n$. Intuitively, the inequality says that if f does not vary too wildly locally, i.e. controlled gradient, then it does not vary too much globally, i.e. bounded variance.

For a historical account of the Poincaré inequality, see Allaire [1]. The Poincaré constant $C_P(\mu)$ of the probability measure μ is defined as the smallest number for which (2) is valid for all C^1-smooth functions f with $\int f d\mu = 0$.

The quantity $1/C_P(\mu)$ is often referred to as the *spectral gap* of μ, for reasons to be explained. In 1960, Payne and Weinberger [42] found that for any n, the best possible value of the supposedly-dimensional constant C_n is in fact

$$C_n = \frac{1}{\pi^2},$$

which does not depend on the dimension. We proceed with an adaptation of the original proof by Poincaré, a proof which does not yield the optimal (in)dependence on the dimension, yet it suffices for some purposes.

Proof of Theorem 1. Passing to a subspace if necessary, we may assume that the probability measure μ is absolutely-continuous with a log-concave density $\rho : \mathbb{R}^n \to [0, \infty)$, which vanishes outside K. We express the variance as a double integral and use the fundamental theorem of calculus:

$$\begin{aligned}\int_K f^2 d\mu &= \frac{1}{2}\int_K\int_K |f(y)-f(x)|^2 d\mu(x)d\mu(y)\\ &= \frac{1}{2}\int_K\int_K \left|\int_0^1 \nabla f((1-t)x+ty)\cdot(y-x)dt\right|^2 d\mu(x)d\mu(y)\\ &\le \frac{Diam^2(K)}{2}\int_K\int_K\int_0^1 \left|\nabla f((1-t)x+ty)\right|^2 \rho(x)\rho(y)dtdxdy,\end{aligned}$$

where we used the inequality $|y-x| \le Diam(K)$. Let us show that for any $0 \le t \le 1$,

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \left|\nabla f((1-t)x+ty)\right|^2 \rho(x)\rho(y)dxdy \le C_{n,t}\int_K |\nabla f|^2 d\mu. \tag{3}$$

Our goal is to replace the product $\rho(x)\rho(y)$ in (3) by some expression involving $\rho((1-t)x+ty)$ and then apply a linear change of variables. Log-concavity will be handy here. We split the argument into two cases. If $t \approx 1/2$, then we will use the inequality

$$\min\{\rho(x), \rho(y)\} \le \rho((1-t)x+ty)$$

that follows from the definition (1) of log-concavity. It implies that

$$\rho(x)\rho(y) \le \rho((1-t)x+ty)\cdot\max\{\rho(x),\rho(y)\} \le \rho((1-t)x+ty)\cdot[\rho(x)+\rho(y)].$$

Thus the integral in (3) is at most

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\nabla f((1-t)x+ty)|^2 \rho((1-t)x+ty)\cdot[\rho(x)+\rho(y)]dxdy \qquad \text{“}u=(1-t)x+ty\text{''}$$
$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\nabla f(u)|^2\rho(u)\rho(x)\frac{du}{t^n}dx + \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\nabla f(u)|^2\rho(u)\rho(y)\frac{du}{(1-t)^n}dy$$
$$= \left[\frac{1}{t^n}+\frac{1}{(1-t)^n}\right]\int |\nabla f|^2 d\mu.$$

In the case where t is not too close to $1/2$ we will use the inequality

$$\rho(x)\rho(y) \le \rho((1-t)x+ty)\rho(tx+(1-t)y)$$

and change variables linearly via

$$u=(1-t)x+ty, \qquad v=tx+(1-t)y.$$

Since $du_j\wedge dv_j = [(1-t)^2-t^2]dx_j\wedge dy_j$ for $j=1,\dots,n$, the integral in (3) is bounded by

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\nabla f((1-t)x+ty)|^2 \rho((1-t)x+ty)\rho(tx+(1-t)y)dxdy$$
$$= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\nabla f(u)|^2\rho(u)\rho(v)\frac{dudv}{|t^n-(1-t)^n|} = \frac{1}{|t^n-(1-t)^n|}\int_{\mathbb{R}^n}|\nabla f|^2 d\mu.$$

Thus the Poincaré inequality follows with

$$C_n \le \frac{1}{2}\int_0^1 \min\left\{\frac{1}{t^n}+\frac{1}{(1-t)^n}, \frac{1}{|t^n-(1-t)^n|}\right\}dt \le C\cdot\frac{2^n}{n},$$

for some universal constant $C>0$, where we separately consider the contribution of the interval $[1/2-1/n, 1/2+1/n]$ to the integral. □

Throughout these lectures, we write $C, c, \tilde{C}, \tilde{c}, \bar{C}$ etc. to denote various positive universal constants whose value may change from one line to the next. Consider the case where μ is the uniform probability measure on a domain $K\subseteq\mathbb{R}^n$. Its Poincaré constant, sometimes denoted also by $C_P(K)$, measures the *conductance* of K. It is large when K has a bottleneck.

(some picture here)

Intuitively, it seems that convexity assumptions rule out many types of bottlenecks, possibly in high dimensions as well. Can we describe the Poincaré constant in terms of simple geometric characteristics of $K\subseteq\mathbb{R}^n$, under convexity assumptions?

Conjecture 2 (Kannan-Lovász-Simonovits [24])**.** *For any log-concave probability measure μ on $\mathbb{R}^n$,*

$$\|Cov(\mu)\|_{op} \leq C_P(\mu) \leq C \cdot \|Cov(\mu)\|_{op} \tag{4}$$

where $C > 0$ is a universal constant.

Here $\|A\|_{op}$ is the operator norm of the symmetric matrix $A \in \mathbb{R}^{n \times n}$, i.e., its maximal eigenvalue in absolute value, and $Cov(\mu) \in \mathbb{R}^{n \times n}$ is the inertia matrix or the *covariance matrix* of μ. The i, j entry of the matrix $Cov(\mu)$ is

$$\int_{\mathbb{R}^n} x_i x_j d\mu(x) - \int_{\mathbb{R}^n} x_i d\mu \int_{\mathbb{R}^n} x_j d\mu(x).$$

The covariance matrix is a symmetric, positive semi-definite matrix. If X is a random vector with law μ and density ρ, we write $C_P(X) = C_P(\mu) = C_P(\rho)$ and $Cov(X) = Cov(\mu) = Cov(\rho)$. With this notation, the Poincaré inequality states that for any C^1-smooth function f,

$$Var f(X) \leq C_P(X) \cdot \mathbb{E}|\nabla f(X)|^2.$$

We note that the left-hand side inequality in (4) is a trivial fact: for any linear functional $f_\theta(x) = x \cdot \theta$ with $\theta \in S^{n-1} = \{x \in \mathbb{R}^n \, ; \, |x| = 1\}$,

$$Cov(X)\theta \cdot \theta = Var(f_\theta(X)) \leq C_P(X) \cdot \mathbb{E}|\nabla f_\theta(X)|^2 = C_P(X),$$

and (4) follows by taking the supremum over all $\theta \in S^{n-1}$. Thus the KLS conjecture suggests that in the log-concave case, the Poincaré inequality is saturated by linear functions, up to a universal constant. Some examples:

1. Consider the one-dimensional case, where X is a random variable that is distributed uniformly in some interval of length L. Then,

   $$Var(X) = \frac{L^2}{12} \qquad \text{and} \qquad C_P(X) = \frac{L^2}{\pi^2},$$

   with the extremal function for the Poincaré inequality on $[0, \pi]$ being $f(x) = \cos x$.

2. Consider the case where X is distributed uniformly in $K = [0, 1]^n$. In this case,

   $$Diam(K) = \sqrt{n}$$

   while by the tensorization property of the Poincaré constant,

   $$C_P(X) = \frac{1}{\pi^2}$$

   and

   $$Cov(X) = \frac{1}{12} \cdot Id.$$

   We thus see that the diameter bound for the Poincaré constant is rather weak in high dimensions, even with the optimal, dimension-independent constant.

3. Suppose that X is distributed uniformly in a Euclidean ball. The Euclidean unit ball $B^n = \{x \in \mathbb{R}^n \, ; \, |x| \leq 1\}$ has volume

$$\frac{\pi^{n/2}}{\Gamma(1+n/2)} = \left(\frac{\sqrt{2\pi e} + o(1)}{\sqrt{n}} \right)^n,$$

which is a rather small number in high dimensions. In order to normalize the volume (or the covariance, or the Poincaré constant), we had better look at the random vector X that is distributed uniformly in a Euclidean ball $K = \sqrt{n} \cdot B^n$. In this case,

$$diam(K) = 2\sqrt{n}, \qquad Cov(X) = \frac{n}{n+2} \cdot \mathrm{Id}.$$

The Poincaré constant of X may be described using Bessel functions, and it has the order of magnitude of a universal constant, in accordance with the KLS conjecture. The Szegö-Weinberger inequality [45, 46] states that among all uniform distributions on domains in $\mathbb{R}^n$ of fixed volume, the Poincaré constant is minimized for a Euclidean ball.

4. Next we discuss the case where X is a standard Gaussian random vector in $\mathbb{R}^n$. Here,

$$Cov(X) = \mathrm{Id} \qquad \text{and} \qquad C_P(X) = 1.$$

Thus the Poincaré inequality in the Gaussian case is precisely saturated by linear functions.

Furthermore, by considering Hermite polynomials one can show the following: In the Gaussian case, a function nearly saturates the Poincare inequality if and only if it is nearly a low-degree polynomial. Indeed, in one direction, if f is a polynomial of degree at most d in n real variables then we can reverse the Poincaré inequality as follows:

$$\mathbb{E}|\nabla f(X)|^2 \leq d \cdot Var(f(X)).$$

In the other direction, if f is a smooth function with

$$\mathbb{E}|\nabla f(X)|^2 \leq R \cdot Var(f(X))$$

then the function f may be approximated by a polynomial of bounded degree: For any $d \geq 0$ there exists a polynomial P of degree at most d such that

$$\mathbb{E}|(f-P)(X)|^2 \leq \frac{R}{d+1} \cdot Var(f(X)).$$

5. Let us work in $\mathbb{C}^n$ and consider the probability measure μ on $\mathbb{C}^n$ with density

$$\prod_{j=1}^n \frac{e^{-|z_j|}}{2\pi}.$$

The measure μ is a log-concave probability measure on $\mathbb{C}^n$. Its covariance matrix is

$$Cov(\mu) = 3 \cdot \mathrm{Id}$$

and its Poincaré constant has the order of magnitude of a universal constant, in accordance with the KLS conjecture.

The density of μ decays expoentially at infinity. Exponentially, but not faster; any log-concave probability density decays exponentially at infinity, yet the Gaussian density decay even faster. This reflects on spectral properties. In the exponential case there are functions that nearly saturate the Poincaré inequality, and they do not necessarily resemble low-degree polynomials. For instance:

Claim: For any holomorphic function $f : \mathbb{C}^n \to \mathbb{C}$ with $f \in L^2(\mu)$ and $\int f d\mu = 0$ (or equivalently, with $f(0) = 0$), the Rayleigh quotient satisfies

$$\frac{1}{3} \leq \frac{\int_{\mathbb{C}^n} |\nabla f|^2 d\mu}{\int_{\mathbb{C}^n} |f|^2 d\mu} \leq \frac{1}{2}. \tag{5}$$

Here is a proof for $n = 1$, which can be easily generalized for any dimension. It suffices to check the validity of (5) for monomials z^k, because of orthogonality relations. If $f(z) = z^k$ with $k \geq 1$ then,

$$\|f\|^2_{L^2(\mu)} = (2k+1)!$$

while

$$\|f'\|^2_{L^2(\mu)} = k^2(2k-1)!$$

The ratio between the two is always between 4 and 6. We remark that by considering the real part of f, we see that (5) holds true for any pluri-harmonic function f, and in particular, when $n = 1$ the relation (5) holds true for any harmonic function $f : \mathbb{R}^2 \to \mathbb{R}$ (thanks to A. Eskenazis for suggesting to add this remark).

6. **Exercise:** ("subadditivity of the Poincaré constant") For two independent random vectors X and Y in $\mathbb{R}^n$,

$$C_P(X+Y) \leq C_P(X) + C_P(Y).$$

1.1 Applications

Poincaré's original motivation for his inequality was related to analysis of partial differential equations such as the *heat equation*. The motivation of Kannan, Lovász and Simonovits in the 1990s came from algorithms based on Markov chains (MCMC) for sampling and for estimating the volume of a high-dimensional convex body. Such tasks appear in linear programming. Another motivation for this research direction, that was put forth by Ball in the early 2000s and later jointly with Nguyen [4], was the relation to Bourgain's slicing problem discussed below. There are models in probability and statistical physics for which log-concavity and Poincaré inequalities are relevant. Let us describe here anther application, related to the *Central Limit Theorem for Convex Sets* [27] from 2006.

A random vector X in $\mathbb{R}^n$ is *isotropic* or *normalized* if $\mathbb{E}X = 0$ and

$$Cov(X) = \text{Id}.$$

Any random vector with finite second moments can be made isotropic by applying an affine-linear transformation. The relation between Gaussian approximation and the Poincaré constant stems from the following:

(i) The Poincaré inequality with $f(x) = |x|$ yields $Var(|X|) \leq C_P(X)$. Thus most of the mass of an isotropic random vector X is contained in spherical shell

$$\left\{ x \in \mathbb{R}^n \, ; \, \sqrt{n} - 3\sqrt{C_P(X)} \leq |x| \leq \sqrt{n} + 3\sqrt{C_P(X)} \right\},$$

whose width has the order of magnitude of the square root of the Poincaré constant.

(ii) Gaussian approximation principle (Sudakov [44], Diaconis-Freedman [16]): When most of the mass of the isotropic random vector X is contained in a thin spherical shell, we have *approximately Gaussian marginals*.

(some picture here)

The following theorem is the current state of the art on Gaussian approximation under Poincaré inequality. We write σ_{n-1} for the uniform probability measure on the unit sphere S^{n-1}.

Theorem 3 (Bobkov, Chistyakov, Götze [6, Proposition 17.5.1])**.** *Let X be an isotropic random vector in $\mathbb{R}^n$. Then there exists a subset $\Theta \subseteq S^{n-1}$ with $\sigma_{n-1}(\Theta) \geq 9/10$ such that any $\theta \in \Theta$,*

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}(X \cdot \theta \leq t) \; - \; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-s^2/2} ds \right| \leq \frac{C \log n}{n} \cdot C_P(X)^2,$$

where $C > 0$ is a universal constant.

It is currently known that $C_P(X) \leq C \cdot \log n$ for an isotropic, log-concave random vector X in $\mathbb{R}^n$, see [33]. Consequently Theorem 3 yields good error estimates in the Central Limit Theorem for Convex sets.

If all we know about the Poincaré constant is the diameter bound, then even in the case of the cube we would be off by a factor of n, and we would not obtain any non-trivial bound for the Central Limit Theorem for Convex sets. Thus in high dimensions it is necessary to refine the diameter bound, as suggested in the KLS conjecture.

What techniques can we use to this end, techniques that go beyond change of variables, Fubini theorem, and the Cauchy-Schwartz inequality used above? High-dimensional convex geometry is a playground for various geometric and analytic ideas that transcend the field of convexity. Any list of approaches that have proven useful to convexity must include convex localization, optimal transport, curvature and the Bochner formula, semigroup tools, geometric measure theory, stochastic localization and complex analysis. In these lectures we explore only some of these directions.

1.2 1D log-concave distributions

Before going on to study methods for high dimensions, let us briefly discuss the one-dimensional case. What do log-concave densities look like in one dimension? They all look like this:

(some picture here)

Proposition 4 ("How to think on 1D log-concave random variables"). *Let $X \in \mathbb{R}$ be a log-concave random variable with density ρ which is isotropic. Then for any $x \in \mathbb{R}$,*

$$c' 1_{|x| \le c''} \le \rho(x) \le C e^{-c|x|}$$

where $c', c'', c, C > 0$ are universal constants.

Can you prove this proposition by yourself? How would you use log-concavity? A hint for the upper bound, is that if $\rho(b) < \rho(a)/2$ for some $a < b$, then ρ decays exponentially and in fact $\rho(x) \le \rho(b) 2^{-x/(b-a)}$ for all $x > b$. As for the lower bound, it's enough to show that $\rho(x) > c'$ for some $x > c''$ and for some $x < -c''$.

Corollary 5 ("reverse Hölder inequalities"). *For any isotropic, log-concave, real-valued random variable X and any $p > -1$,*

$$c \cdot \min\{p+1, 1\} \le \|X\|_p = (\mathbb{E}|X|^p)^{1/p} \le C(|p|+1), \tag{6}$$

where $c, C > 0$ are universal constants.

The case $p = 0$ in (6) is interpreted by continuity, i.e.,

$$\|X\|_0 = \exp(\mathbb{E} \log |X|).$$

This is not a norm, yet a nice feature is its multiplicativity: for any random variables X and Y, possibly dependent,

$$\|XY\|_0 = \|X\|_0 \|Y\|_0.$$

Proof of Corollary 5. Begin with the inequality on the right-hand side. By the monotonicity of $p \mapsto \|X\|_p$, it is enough to look at $p > 0$. In this case,

$$\|X\|_p^p = \int_{-\infty}^{\infty} |t|^p \rho(t)dt \leq C \int_{-\infty}^{\infty} |t|^p e^{-c|t|} dt = \frac{2C}{c^{p+1}} \Gamma(p+1) \leq (\tilde{C}p)^p.$$

where we used the fact that for integer p, we have $\Gamma(p+1) = p! \leq p^p$. For the lower bound, by monotonicity it suffices to look at $p < 0$. Setting $q = -p \in (0,1)$ we have

$$\mathbb{E}\frac{1}{|X|^q} \leq C \int_{-\infty}^{\infty} \frac{1}{|t|^q} e^{-c|t|} dt \leq \frac{C'}{1-q}$$

and hence

$$\|X\|_p = \left(\mathbb{E}\frac{1}{|X|^q}\right)^{-1/q} \geq (C'(1-q))^{1/q} \geq \tilde{C}(1-q).$$

□

We proceed to discuss the isoperimetric profile of a log-concave distribution in one dimension. Bobkov [5] shows that for a probability density ρ on the real line,

$$\rho \text{ is log-concave} \qquad \Longleftrightarrow \qquad \rho \circ \Phi^{-1} : [0,1] \to (0, \infty) \text{ is concave} \tag{7}$$

where $\Phi(x) = \int_{-\infty}^{x} \rho(t)dt$ and $\Phi^{-1}(y) = \inf\{x \in \mathbb{R}\,;\, \Phi(x) \geq y\}$. Once stated, (7) is not difficult to prove. It follows from (7) that the function

$$I(x) = \min\left\{\rho \circ \Phi^{-1}, \rho \circ (1-\Phi)^{-1}\right\}$$

is concave. Write μ for the measure whose density is ρ, and note that

$$I(x) = \min\{\rho(\partial H)\,;\, H \text{ is a ray with } \mu(H) = x\}$$

Since the boundary ∂H is a singleton as H is a ray, in this case we abbreviate $\rho(\partial H) = \rho(a)$ if $\partial H = \{a\}$. The following Proposition by Bobkov implies that the concave function I is the *isoperimetric profile* of the probability density ρ.

We prefer to discuss isoperimetry through ε-neighborhoods. For $\varepsilon > 0$ and a subset $A \subseteq \mathbb{R}$ we write $A_\varepsilon = \{x \in \mathbb{R}\,;\, \inf_{y \in A} |x - y| < \varepsilon\}$ for its ε-neighborhood. We remark that analogously to (7), the log-concavity of ρ implies that the function $x \mapsto \Phi(\Phi^{-1}(x) + \varepsilon)$. This shows that the function

$$I_\varepsilon(x) = \min\{\mu(H_\varepsilon)\,;\, H \text{ is a ray with } \mu(H) = x\}$$

is a concave function of $x \in [0,1]$.

Proposition 6 (Bobkov [5])**.** *Let μ be a log-concave probability measure on $\mathbb{R}$ with density ρ. Fix $0 < p < 1, \varepsilon > 0$. Then among all Borel subsets $A \subseteq \mathbb{R}$ with $\mu(A) = p$, the infimum of $\mu(A_\varepsilon)$ is attained for a half line.*

Sketch of Proof. It suffices to show that half lines are better than finite unions of intervals. How can we deal with a subset A that is a finite union of intervals? Using the following claim. For $a \in \mathbb{R}$ with $\mu([a,\infty)) > p$ consider the unique interval $J(a) = (a,b)$ such that $\mu(J(a)) = p$. The claim is that the function

$$a \mapsto \mu(J(a)_\varepsilon)$$

is unimodal, thanks to log-concavity (i.e., the function is increasing and then decreasing). Again, once stated this is not too difficult to prove. Given this claim, one may fix all intervals in A but one, and then move the remaining one around and expand and shrink it so as to preserve the total μ-measure. It follows that gluing this interval to one of the sides cannot increase the μ-measure of the ε-neighborhood. $\square$

Corollary 7. *Let μ be an isotropic, log-concave probability measure on $\mathbb{R}$ and let $\varepsilon, p \in (0,1)$. Then for any Borel set $S \subseteq \mathbb{R}$ with $\mu(S) = p$,*

$$\mu(S_\varepsilon \setminus S) \geq c \cdot \varepsilon \cdot \min\{p, 1-p\}$$

where $c > 0$ is a universal constant.

Exercise: Fill in the details in the proofs of Proposition 6 and Corollary 7.

2 Optimal Transport theory with the Monge cost

Let μ_1 and μ_2 be two measures in $\mathbb{R}^n$, say compactly-supported and absolutely continuous, with the same total mass, i.e., $\mu_1(\mathbb{R}^n) = \mu_2(\mathbb{R}^n)$. We would like to push-forward the measure μ_1 to the measure μ_2 in the most efficient way, that minimizes the average distance that points have to travel. That is, we look at the optimization problem

$$\inf_{S_*(\mu)=\nu} \int_{\mathbb{R}^n} |Sx - x| d\mu_1(x).$$

This is the problem of Optimal Transport with the Monge cost or the L^1 cost, considered by Monge in 1781. See Cayley's review of Monge's work [12] from 1882. Here is a heuristics from Monge's paper that explains why this problem induces a partition into segments.

Monge heuristics: For the optimal transport map T, the segments $(x, T(x))$ $(x \in Supp(\mu_1))$ do not intersect, unless they overlap.

Explanation. Suppose that the segments (x, Tx) and (y, Ty) intersect at a point z, and apply the Triangle Inequality. $\square$

(some picture here)

This is related to the following elementary riddle: given 50 red points and 50 blue points in the plane, in general position, find a matching so that the corresponding segments do not intersect.

Since the above argument relies only on the triangle inequality, you would expect that the optimal transport problem would induce a partition into geodesics also for Riemannian manifolds, or Finslerian manifolds, or measure metric spaces of some type – basically wherever the triangle inequality holds true (under some regularity assumptions).

2.1 Linear programming relaxation and the dual problem

In Monge's problem we minimize over all maps S that push-forward μ_1 to μ_2. There is a relaxation of this problem, that looks at all possible *couplings*, or transport plans, of the two distributions. That is, instead of mapping a point x to a single point Tx, we are allowed to spread the mass across a region. Thus we look at all measures γ on $\mathbb{R}^n \times \mathbb{R}^n$ with

$$(\pi_1)_* \gamma = \mu_1 \qquad \text{and} \qquad (\pi_2)_* \gamma = \mu_2.$$

where $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$. Such a measure is called a *coupling* of μ and ν. In other words, we now look at *transport plans* rather than *transport maps*. The advantage is that the space of all couplings is a convex set. The relaxed optimal transport problem involves minimizing the average distance that points travel, namely we look at

$$\inf_{(\pi_1)_*\gamma = \mu, (\pi_2)_*\gamma = \nu} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y| d\gamma(x, y).$$

Hence we minimize a linear function on a convex set, this is Linear Programming or Functional Analysis (see e.g. Kantorovich and Akilov [25, Section VIII.4]).

Theorem 8. *(The dual problem) Let μ_1, μ_2 be two absolutely-continuous measures in $\mathbb{R}^n$ with the same total mass. Assume that*

$$\int_{\mathbb{R}^n} |x| d\mu_1(x) < \infty \qquad \textit{and} \qquad \int_{\mathbb{R}^n} |x| d\mu_2(x) < \infty.$$

Denote $\mu = \mu_2 - \mu_1$. Then the following quantities are equal:

1. *The minimum over all couplings γ of μ_1 and μ_2 of the integral*

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y| d\gamma(x, y).$$

2. *The maximum over all 1-Lipschitz functions $u : \mathbb{R}^n \to \mathbb{R}$ of*

$$\int_{\mathbb{R}^n} u d\mu$$

3. *The minimum over all maps T with $T_*\mu_1 = \mu_2$ of*

$$\int_{\mathbb{R}^n} |x - Tx| d\mu_1(x).$$

Proof sketch. We refer to Ambrosio [2] for full details. For the easy direction of the linear programming duality, pick a 1-Lipschitz map u and a coupling γ. For any points $x, y \in \mathbb{R}^n$,

$$u(y) - u(x) \leq |x - y|.$$

Integrating with respect to γ, we get

$$\int_{\mathbb{R}^n} u d\mu = \int_{\mathbb{R}^n \times \mathbb{R}^n} [u(y) - u(x)] d\gamma(x,y) \leq \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y| d\gamma(x,y). \tag{8}$$

Hence we need to find u and γ so that equality is attained in (8). The argument goes roughly as follows. A compactness argument shows that the infimum over all couplings is attained. Indeed, by Alaoglu's theorem, the collection of all couplings is compact in the w^*-topology (integration against continuous functions on $\mathbb{R}^n$ whose limit at infinity exists). The functional $\gamma \mapsto \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y| d\gamma(x,y)$ is lower semi-continuous in w^*-topology, hence its minimum is attained.

Similarly to the Monge heuristics, the optimality implies that the support of γ must be cyclically monotone: If $(x_i, y_i) \in Supp(\gamma) \subseteq \mathbb{R}^n \times \mathbb{R}^n$ for $i = 1, \ldots, N$ then for any permutation $\sigma \in S_N$,

$$\sum_{i=1}^{N} |x_i - y_i| \leq \sum_{i=1}^{N} |x_i - y_{\sigma(i)}|. \tag{9}$$

Indeed, otherwise one may pick small balls around x_i and y_i and rearrange them to contradict optimality. Similarly to Rockafellar's theorem from convex geometry, condition (9) implies that there exists a 1-Lipschitz function $u : \mathbb{R}^n \to \mathbb{R}$ with

$$(x, y) \in Supp(\gamma) \qquad \Longrightarrow \qquad u(y) - u(x) = |y - x|. \tag{10}$$

Indeed, fix $(x_0, y_0) \in Supp(\gamma)$ and define $u(x)$ as the supremum over all lower bounds with $u(x_0) = 0$,

$$u(x) = \sup_{N, (x_1, y_1), \ldots, (x_N, y_N) \in Supp(\gamma)} \{|x_0 - y_0| - |y_0 - x_1| + |x_1 - y_1| - |y_1 - x_2| + \ldots - |y_N - x|\}$$

(some picture here)

It follows from (9) that $u(x_0) = 0$. The function u is a 1-Lipschitz function as a supremum of 1-Lipschitz functions. It follows from the definition of u that (10) holds true. Hence we found u and γ so that equality is attained in (8). The proof that γ can also be replaced by a transport map is due to Evans and Gangbo [19]. This relies on analysis of the structure of u that will be described next. $\square$

Remark 9. The minimizers γ or T are not at all unique. It is actually the 1-Lipschitz function u which is essentially determined. More precisely, the gradient ∇u is determined μ-almost everywhere.

We move on to discuss the structure of 1-Lipschitz functions. Observe that when a 1-Lipschitz function u satisfies $|u(x) - u(y)| = |x - y|$, for some points $x, y \in \mathbb{R}^n$, it necessarily grows in speed one along the segment from x to y. A maximal open segment I on which u grows with speed one, i.e., $|u(x) - u(y)| = |x - y|$ for all $x, y \in I$, is called a *transport ray*. Theorem 8 tells us that optimal transport only happens only along transport rays, we only rearrange mass along transport rays.

It is illuminating to draw the transport rays of the function $u(x) = x_1$ in connection with Fubini's theorem

$$\int_{\mathbb{R}^2} \varphi = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \varphi(x_1, x_2) dx_1 \right) dx_2,$$

and of the function $u(x) = |x|$ on $\mathbb{R}^2 \cong \mathbb{C}$ in connection with integration in polar coordinates:

$$\int_{\mathbb{R}^2} \varphi = \int_0^{2\pi} \left(\int_0^{\infty} \varphi(re^{i\theta}) r dr \right) d\theta.$$

Note that the Jacobian factor on the needle is log-concave in both examples.

(some picture here)

The next step is to understand the *disintegration of measure* or *conditional probabilities* induced by the partition into transport rays. Let u be a maximizer as above, with

$$\mu = \mu_2 - \mu_1,$$

and with the two measures satisfying the requirements of Theorem 8. As it turns out, it is guaranteed that transport rays of positive length form a partition of the entire support of the measure μ, up to a set of measure zero. Write

$$f = \frac{d\mu}{d\lambda}$$

where λ is the Lebesgue measure on $\mathbb{R}^n$, or better: We may work with any log-concave reference measure λ in $\mathbb{R}^n$, not just the Lebesgue measure. The assumption that $\mu_1(\mathbb{R}^n) = \mu_2(\mathbb{R}^n)$ is equivalent to the requirement that

$$\int_{\mathbb{R}^n} f d\lambda = 0.$$

The following theorem requires careful regularity analysis, and in addition to Evans and Gangbo [19] it builds upon works by Caffarelli, Feldman and McCann [11] as well as [29]. It is analogous to integration in polar coordinates, yet with respect to a general 1-Lipschitz guiding function, rather than just $u(x) = |x|$. In the following theorem a line segment could also mean a singleton, a ray or a line.

Theorem 10. *Under the above assumptions, there is a collection Ω of line segments that form a partition of $\mathbb{R}^n$, a family of measures $\{\lambda_{\mathcal{I}}\}_{\mathcal{I}\in\Omega}$, and a measure ν on the space of segments Ω, such that*

1. *For any $\mathcal{I} \in \Omega$ the measure $\lambda_{\mathcal{I}}$ is supported on the line segment $\mathcal{I}$. If $\mathcal{I}$ is of non-zero length, then it is a transport ray of the 1-Lipschitz function u.*

2. *Disintegration of measure*
$$\lambda = \int_{\Omega} \lambda_{\mathcal{I}} d\nu(\mathcal{I}).$$

3. *Mass balance condition: for ν-almost any $\mathcal{I} \in \Omega$,*
$$\int_{\mathcal{I}} f d\lambda_{\mathcal{I}} = 0.$$

4. *For ν-almost any $\mathcal{I} \in \Omega$, the measure $\lambda_{\mathcal{I}}$ has a C^∞-smooth, positive density ρ on the segment $\mathcal{I}$ which is log-concave.*
(In fact, in the case where λ is the Lebesgue measure, it is a polynomial of degree $n-1$ with real roots, that does not vanish in the support of $\lambda_{\mathcal{I}}$).

Remark 11. This theorem may be generalized to any Riemannian manifold with non-negative Ricci curvature. We replace the line segment $\mathcal{I}$ by a unit-speed geodesic $\gamma = \gamma_{\mathcal{I}}$, and set $\kappa(t) = Ricci(\dot\gamma(t), \dot\gamma(t)), n = \dim(M)$. Denote by $\rho = \rho_{\mathcal{I}}$ the density of $\mu_{\mathcal{I}}$ with respect to arclength on the geodesic $\gamma = \gamma_{\mathcal{I}}$. Then,

$$\left(\rho^{\frac{1}{n-1}}\right)'' + \frac{\kappa}{n-1} \cdot \rho^{\frac{1}{n-1}} \leq 0.$$

The Riemannian version may be used to prove isoperimetric inequalities under lower bounds on the Ricci curvature, as well as Poincaré inequalities, log-Sobolev inequalities, Brunn-Minkowski inequalities etc.

Some ideas from the proof of Theorem 10. The proof of Theorem 10 does not use sophisticated results from Geometric Measure Theory, but it consists of several steps. Essentially,

- Show that a 1-Lipschitz u is always differentiable in the relative interior of a transport ray.
- The next step is to show that ∇u is a locally-Lipschitz function on a set which is only slightly smaller than the union of all transport rays, and that the restriction of u to this set may be extended to a $C^{1,1}$-function on $\mathbb{R}^n$.
- This is just enough regularity in order to allow change of variables in an integral, which yields the disintegration.
- By differentiating the Jacobian one sees that the logarithmic derivative of the needle density is the mean curvature of the level set of u, and the inverse principal curvatures grow linearly along the needle. This yields log-concavity along each needle.
- The mass balance condition follows from the fact that γ is a coupling between μ_1 and μ_2, and that transport happens only along transport rays (thanks to S. Szarek for this remark). Alternatively, one can use a perturbative argument based on the maximality of the integral $\int u f d\lambda$.

$\square$

As an application of this theorem, let us prove the reverse Cheeger inequality of Buser [10] and Ledoux [36], and in fact a refinement due to E. Milman [41]. In Joseph's lectures you can see another proof, using semi-group methods, of the following:

Proposition 12. *Let μ be a log-concave probability measure on $\mathbb{R}^n$ and $R > 0$. Assume that for any 1-Lipschitz function $u : \mathbb{R}^n \to \mathbb{R}$ there exists $\alpha \in \mathbb{R}$ with*

$$\int_{\mathbb{R}^n} |u(x) - \alpha| d\mu(x) \leq R. \tag{11}$$

(this is a weaker condition than requiring $C_P(\mu) \leq R^2$). Then for any measurable set $S \subseteq \mathbb{R}^n$ and $0 < \varepsilon < R$,

$$\mu(S_\varepsilon \setminus S) \geq c \cdot \frac{\varepsilon}{R} \cdot \mu(S) \cdot (1 - \mu(S)), \tag{12}$$

where $c > 0$ is a universal constant, and where S_ε is the ε-neighborhood of S.

Proof. Denote $t = \mu(S) \in [0, 1]$ and set $f(x) = 1_S(x) - t$ for $x \in \mathbb{R}^n$. Then $\int f d\mu = 0$. Let u be a 1-Lipschitz function maximizing

$$\int_{\mathbb{R}^n} u f d\mu.$$

After adding a constant to u, we may assume that

$$\int_{\mathbb{R}^n} |u| d\mu \leq R.$$

By Theorem 10, we obtain a needle decomposition: measures $\{\mu_{\mathcal{I}}\}_{\mathcal{I} \in \Omega}$ on $\mathbb{R}^n$, and a measure ν on the space Ω of transport rays which yield a disintegration of measure. We may normalize and assume that all of these measures are probability measures (i.e., replace $\mu_{\mathcal{I}}$ and ν by $\tilde{\mu}_{\mathcal{I}}$ and $\tilde{\nu}$ where $\tilde{\mu}_{\mathcal{I}} = \mu_{\mathcal{I}}/\mu_{\mathcal{I}}(\mathbb{R}^n)$ and $d\tilde{\nu}/\nu(\mathcal{I}) = \mu_{\mathcal{I}}(\mathbb{R}^n)$). Hence,

$$\int_\Omega \left(\int_{\mathcal{I}} |u| d\mu_{\mathcal{I}} \right) d\nu(\mathcal{I}) = \int_{\mathbb{R}^n} |u| d\mu \leq R.$$

Denote

$$B = \left\{ \mathcal{I} \in \Omega \, ; \, \int_{\mathcal{I}} |u| d\mu_{\mathcal{I}} \leq 2R \right\}.$$

By the Markov-Chebyshev inequality,

$$\nu(B) \geq 1/2. \tag{13}$$

For ν-almost all intervals $\mathcal{I} \in \Omega$ we know that $\int_{\mathcal{I}} f d\mu_{\mathcal{I}} = 0$, hence

$$\mu_{\mathcal{I}}(S) = t \cdot \mu_{\mathcal{I}}(\mathbb{R}^n) = t.$$

We would like to prove that for any $\mathcal{I} \in B$ and any $0 < \varepsilon < R$,

$$\mu_{\mathcal{I}}(S_\varepsilon \setminus S) \geq c \cdot \frac{\varepsilon}{R} \cdot t(1-t), \tag{14}$$

for a universal constant $c > 0$. Once (14) is proven, the bound (12) follows by integrating (14) with respect to ν and using (13), since

$$\mu(S_\varepsilon \setminus S) \geq \int_B \mu_{\mathcal{I}}(S_\varepsilon \setminus S) d\nu(\mathcal{I}) \geq \nu(B) \cdot c \cdot \frac{\varepsilon}{R} \cdot t(1-t) \geq \frac{c}{2} \cdot \frac{\varepsilon}{R} \cdot t(1-t).$$

What remains to be proven is a one-dimensional statement about log-concave measures: If $\eta = \mu_{\mathcal{I}}$ is a log-concave probability measure on $\mathbb{R}$ with $\int_{\mathbb{R}} |t| d\eta(t) \leq R$, then (14) holds true. This follows from Corollary 7 and a scaling argument. □

The same proof applies for any complete Riemannian manifold with non-negative Riemannian curvature. In fact, completeness in unneeded, the weaker geodesic-convexity assumption suffices here. There are quite a few other applications for this theorem, which helps reduce the task of proving an n-dimensional inequality to the task of proving a 1-dimensional inequality ("localization"). In a simply-connected space of constant sectional curvature, most of these applications – like reverse Hölder inequalities for polynomials – may also be proven using a localization method based on hyperplane bisections that go back to Payne and Weinberger [42], Gromov and Milman [22] and Kannan, Lovász and Simonovotis [24]. Proposition 12 seems to be an exception, our proof requires the 1-Lipschitz guiding function.

Exercise: ("reverse Hölder inequalities for polynomials") Let X be a log-concave random vector in $\mathbb{R}^n$, and let $f : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree at most d. Then for any $0 < p \leq q$,

$$\|f(X)\|_q \leq C_{q,d} \cdot \|f(X)\|_p,$$

for some constant $C_{q,d}$ depending only on q and d.

The solution uses needle decomposition. As a hint, let us prove it in one dimension, following Bobkov [7]. We may assume that f is a monic polynomial in one real variable, hence

$$f(X) = \prod_{i=1}^{d} (X - z_i)$$

for some $z_1, \ldots, z_d \in \mathbb{C}$. Consequently, by Hölder inequality and by Corollary 5,

$$\|f(X)\|_q = \left\| \prod_{i=1}^{d} (X - z_i) \right\|_{dq} \leq \prod_{i=1}^{d} \|X - z_i\|_{dq} \leq \prod_{i=1}^{d} Cd(q+1)\|X - z_i\|_0 = (Cd(q+1))^d \|f(X)\|_0.$$

2.2 Isoperimetry and the Poincaré inequality

The Cheeger inequality [13] states that for any absolutely-continuous probability measure on $\mathbb{R}^n$ satisfying sone mild regularity assumptions,

$$C_P(\mu) \leq 4\psi_\mu^2 \tag{15}$$

where ψ_μ is the *isoperimetric constant* of the probability measure μ, defined via

$$\frac{1}{\psi_\mu} = \inf_{A \subseteq \mathbb{R}^n} \left\{ \frac{\int_{\partial A} \rho}{\min\{\mu(A), 1 - \mu(A)\}} \right\}$$

where ρ is the density of μ and where the infimum runs over all open sets $A \subseteq \mathbb{R}^n$ with smooth boundary satisfying $0 < \mu(A) < 1$. Inequality (15) is proven by the co-area formula and the Cauchy-Schwartz inequality, see Joseph's lectures.

Proposition 12 thus implies that the Poincaré inequality and the isoperimetric inequality are equivalent in the log-concave case, up to a universal constant. De Ponti and Mondino [15] used the technique from Ledoux [36] in order to find the optimal value of the universal constant, and showed that

$$\frac{1}{\pi}\psi_\mu^2 \leq C_P(\mu) \leq 4\psi_\mu^2.$$

Proposition 12 moreover implies that in the log-concave case, there exists a 1-Lipschitz function f such that

$$\psi_\mu^2 \leq C \cdot Var_\mu(f).$$

The Cheeger inequality thus leads to the following corollary of Proposition 12:

Corollary 13 (E. Milman [41]). *Let μ be a log-concave probability measure on $\mathbb{R}^n$. Then there exists a 1-Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ such that*

$$c \cdot C_P(\mu) \leq Var_\mu(f) \leq C_P(\mu) \tag{16}$$

where $c > 0$ is a universal constant.

3 Bochner identities and curvature

In this lecture we discuss a technique that originated in Riemannian Geometry and connects the Poincaré inequality and Curvature. It started with the works of Bochner in the 1940s and also Lichnerowicz in the 1950s. The approach fit well with convex bodies and log-concave measures in high dimension. In a nutshell, the idea is to make local computations involving something like curvature, as well as integrations by parts, and then dualize and obtain Poincaré-type inequalities. This may sound pretty vague, let us explain what we mean.

Suppose that μ is an absolutely-continuous log-concave probability measure in $\mathbb{R}^n$. Then μ is supported in an open, convex set $K \subseteq \mathbb{R}^n$ and it has a positive, log-concave density $\rho = e^{-\psi}$ in K. We will measure distances using the Euclidean distances in $\mathbb{R}^n$, but we will measure volumes using the measure μ. We thus look at the *weighted Riemannian manifold* or the *metric-measure space*

$$(K, |\cdot|, \mu).$$

Thus the Dirichlet energy of a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ is

$$\|f\|^2_{\dot{H}^1(\mu)} = \int_K |\nabla f|^2 d\mu.$$

Indeed, we measure the length of the gradient with respect to the Euclidean metric, while we integrate with respect to the measure μ. As was already defined in Joseph's lectures, the Laplace-type operator associated with this measure-metric space is defined, initially for $u \in C_c^\infty(K)$, via

$$Lu = L_\mu u = \Delta u - \nabla\psi \cdot \nabla u = e^{\psi} div(e^{-\psi}\nabla u).$$

This reason for this definition is that for any smooth functions $u, v : \mathbb{R}^n \to \mathbb{R}$, with one of them compactly-supported in K,

$$\int_{\mathbb{R}^n} (Lu) v d\mu = -\int_{\mathbb{R}^n} [\nabla u \cdot \nabla v] e^{-\psi}.$$

and in particular

$$\langle -Lu, u \rangle_{L^2(\mu)} = \int_{\mathbb{R}^n} |\nabla u|^2 d\mu.$$

Thus L is a symmetric operator in $L^2(\mu)$, defined initially for $u \in C_c^\infty(K)$. It can have more than one self-adjoint extension, for example corresponding to the Dirichlet or Neumann boundary conditions when K is bounded. When discussing the Bochner technique, it is customary and possible to find ways to circumvent spectral theory of the operator L. Still, spectral theory helps us understand and form intuition, and we will at least quote the relevant spectral theory.

It will be convenient to make an (inessential) regularity assumption on μ, so as to avoid all boundary terms in all integrations by parts. We say that μ is a regular, log-concave measure in $\mathbb{R}^n$ if its density, denoted by $e^{-\psi}$, is smooth and positive in $\mathbb{R}^n$ and the following two requirements hold:

(i) Log-concavity amounts to ψ being convex, so $\nabla^2\psi \geq 0$ everywhere in $\mathbb{R}^n$. We require a bit more, that there exists $\varepsilon > 0$ such that for all $x \in \mathbb{R}^n$,

$$\varepsilon \cdot \mathrm{Id} \leq \nabla^2\psi(x) \leq \frac{1}{\varepsilon} \cdot \mathrm{Id}. \tag{17}$$

(ii) The function ψ, as well as each of its partial derivatives, grows at most polynomially at infinity.

Exercise (regularization process): Begin with an arbitrary log-concave measure μ on $\mathbb{R}^n$, convolve it by a tiny Gaussian, and then multiply its density by $\exp(-\varepsilon|x|^2)$ for small $\varepsilon > 0$. Show that the resulting measure is regular, log-concave, with approximately the same covariance matrix, and that the Poincaré constant cannot jump down by much under this regularization process.

From now on, we assume that our probability measure μ is regular, log-concave measure. It turns out that in this case, the operator L, initially defined on $C_c^\infty(\mathbb{R}^n)$, is essentially self-adjoint, positive semi-definite operator in $L^2(\mu)$ with a discrete spectrum. Its eigenfunctions $1 \equiv \varphi_0, \varphi_1, \ldots$ constitute an orthonormal basis, and the eigenvalues of $-L$ are

$$0 = \lambda_0(L) < \lambda_1(L) = \frac{1}{C_P(\mu)} \leq \lambda_2(L) \leq \ldots$$

with the eigenfunction corresponding to the trivial eigenvalue 0 being the constant function. The eigenfunctions are smooth functions in $\mathbb{R}^n$ that do not grow too fast at infinity: the function

$$\varphi_j e^{-\psi/2}$$

decays exponentially at infinity. Also $(\partial^\alpha \varphi_j)e^{-\psi/2}$ decays exponentially at infinity for any partial derivative α. This follows from known results on exponential decay of eigenfunctions of Schrödinger operators. The eigenvalues are given by min-max of the Rayleigh quotients,

$$\lambda_k(L) = \inf_{f \perp \varphi_0, \ldots, \varphi_{k-1}} \frac{\int_{\mathbb{R}^n} |\nabla f|^2 d\mu}{\int_{\mathbb{R}^n} f^2 d\mu}$$

where the infimum runs over all (say) locally-Lipschitz functions $f \in L^2(\mu)$. Since $\varphi_0 \equiv Const$, we indeed see that the first eigenfunction φ_1 saturates the Poincaré inequality for μ. For proofs of these spectral theoretic facts, see references in [33].

Let us return to Geometry. In Riemannian geometry, the Ricci curvature appears when we commute the Laplacian and the gradient. Analogously, here we have the easily-verified commutation relation

$$\nabla(Lu) = L(\nabla u) - (\nabla^2\psi)(\nabla u),$$

where $L(\nabla u) = (L(\partial^1 u), \ldots, L(\partial^n u))$. Hence the matrix $\nabla^2\psi$ corresponds to a curvature term, analogous to the Ricci curvature.

Proposition 14 (Integrated Bochner's formula)**.** *For any* $u \in C_c^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} (Lu)^2 \, d\mu = \int_{\mathbb{R}^n} \left(\nabla^2 \psi\right) \nabla u \cdot \nabla u \, d\mu + \int_{\mathbb{R}^n} \|\nabla^2 u\|_{HS}^2 d\mu,$$

where $\|\nabla^2 u\|_{HS}^2 = \sum_{i=1}^n |\nabla \partial_i u|^2$.

Proof. Integration by parts gives

$$\begin{aligned}\int_{\mathbb{R}^n} (Lu)^2 \, d\mu = -\int_{\mathbb{R}^n} \nabla(Lu) \cdot \nabla u \, d\mu &= -\int_{\mathbb{R}^n} L(\nabla u) \cdot \nabla u \, d\mu + \int_{\mathbb{R}^n} \left[(\nabla^2 \psi) \nabla u \cdot \nabla u\right] \, d\mu \\ &= \sum_{i=1}^n \int_{\mathbb{R}^n} |\nabla \partial_i u|^2 \, d\mu + \int_{\mathbb{R}^n} \left(\nabla^2 \psi\right) \nabla u \cdot \nabla u \, d\mu.\end{aligned}$$

□

The assumption that u is compactly-supported was used in order to discard the boundary terms when integrating by parts. In fact, it suffices to know that u is μ-tempered. We say that u is μ-tempered if it is a smooth function, and $(\partial^\alpha u) e^{-\psi/2}$ decays exponentially at infinity for any partial derivative $\partial^\alpha u$. Any eigenfunction of L is μ-tempered. if f is μ-tempered, then so is Lf.

The following inequality from [33] is analogous to some investigations of Lichnerowicz [39]. It is concerned with distributions that are more log-concave than a Gaussian distribution, in the sense that their logarithmic Hessian is uniformly bounded by that of the Gaussian.

Theorem 15 (improved log-concave Lichnerowicz inequality)**.** *Let* $t > 0$ *and assume that* $\nabla^2 \psi(x) \geq t$ *for all* $x \in \mathbb{R}^n$*. Then,*

$$C_P(\mu) \leq \sqrt{\|Cov(\mu)\|_{op} \cdot \frac{1}{t}}.$$

Equality in Theorem 15 is attained when μ is a Gaussian measure. Write γ_s for the law of distribution of a Gaussian random vector of mean zero and variance $s \cdot \mathrm{Id}$ in $\mathbb{R}^n$. Then γ_s satisfies the assumptions of Theorem 15 for $t = 1/s$ while $C_P(\gamma_s) = \|Cov(\gamma_s)\|_{op} = s$.

Proof of Theorem 15. Denote $f = \varphi_1$, the first eigenfunction, normalized so that $\|f\|_{L^2(\mu)} = 1$. Set $\lambda = 1/C_P(\mu)$. By the Bochner formula and the Poincaré inequality for $\partial^i f$ $(i = 1, \ldots, n)$,

$$\begin{aligned}\lambda^2 = \int_{\mathbb{R}^n} (Lf)^2 d\mu &= \int_{\mathbb{R}^n} [(\nabla^2 \psi) \nabla f \cdot \nabla f] d\mu + \int_{\mathbb{R}^n} \|\nabla^2 f\|_{HS}^2 d\mu \\ &\geq t \int_{\mathbb{R}^n} |\nabla f|^2 d\mu + \lambda \left[\int_{\mathbb{R}^n} |\nabla f|^2 d\mu - \left|\int_{\mathbb{R}^n} \nabla f d\mu\right|^2\right] \\ &= (t + \lambda) \cdot \lambda - \lambda \left|\int_{\mathbb{R}^n} \nabla f d\mu\right|^2. \end{aligned} \tag{18}$$

Therefore the first eigenfunction has a "preferred direction", i.e.,

$$\left| \int_{\mathbb{R}^n} \nabla f d\mu \right|^2 \geq t. \tag{19}$$

We remark that in the general case, under log-concavity assumptions it is known that $\int_{\mathbb{R}^n} \nabla f d\mu \neq 0$, see [28], and this leads to a bound on the dimension of the first eigenspace. The lower bound (19) is a quantitative version, relying on the assumption of a uniform lower bound on the log-concavity. Using that the i^{th} coordinate of ∇f is $\nabla f \cdot \nabla x_i$ and integrating by parts we have

$$\int_{\mathbb{R}^n} \nabla f d\mu = -\int_{\mathbb{R}^n} (Lf) x d\mu = \lambda \int_{\mathbb{R}^n} f x d\mu$$

Since $\int f d\mu = 0$, by Cauchy-Schwartz, for some $\theta \in S^{n-1}$,

$$\left| \int_{\mathbb{R}^n} \nabla f d\mu \right| = \int_{\mathbb{R}^n} \langle \nabla f, \theta \rangle d\mu = \lambda \int_{\mathbb{R}^n} f(x) \langle x, \theta \rangle d\mu(x) \leq \lambda \|f\|_{L^2(\mu)} \cdot \sqrt{Cov(\mu)\theta \cdot \theta} \leq \lambda \|Cov(\mu)\|_{op}.$$

This expression is at least t, and the theorem follows. □

Since $\|Cov(\mu)\|_{op} \leq C_P(\mu)$, we deduce from Theorem 15 that

$$C_P(\mu) \leq \frac{1}{t}. \tag{20}$$

Inequality (20) is sometimes referred to as the log-concave Lichnerowicz inequality. Therefore the bound in Theorem 15 is a geometric average of the Lichnerowicz bound and the conjectural KLS bound.

The Bochner identity has quite a few additional applications in the study of log-concave measures, beyond the improved log-concave Lichnerowicz inequality. Especially if one introduces the semigroup $(e^{tL})_{t \geq 0}$ associated with the operator L (see e.g. Ledoux [35]), as we see in Joseph's lectures. Yet even simple integrations by parts and duality arguments based on the Bochner identity lead to non-trivial conclusions. One example is the Brascamp-Lieb inequality [9] from the 1970s:

Theorem 16 (Brascamp-Lieb). *For any C^1-smooth $f \in L^2(\mu)$,*

$$Var_\mu(f) \leq \int_{\mathbb{R}^n} \left(\nabla^2 \psi \right)^{-1} \nabla f \cdot \nabla f \, d\mu(x),$$

where $Var_\mu(f) = \int_{\mathbb{R}^n} (f - E)^2 \, d\mu(x)$, and $E = \int_{\mathbb{R}^n} f d\mu$.

Proof. We will only prove this inequality for regular, log-concave measures, though it holds true under weaker regularity assumptions. The space of all μ-tempered functions is denoted by $\mathcal{F}_\mu$. It is clearly a dense subspace of $L^2(\mu)$ and in fact its image under L is dense in

$$\varphi_0^\perp = \left\{ g \in L^2(\mu) \, ; \, \int_{\mathbb{R}^n} g d\mu = 0 \right\}.$$

Indeed, the image contains all finite linear combinations of all eigenfunctions $\varphi_1, \varphi_2, \dots$ (without φ_0) which is dense in H. Assume $\int f \, d\mu = 0, \varepsilon > 0$ and pick $u \in \mathcal{F}_\mu$ such that

$$\|Lu - f\|_{L^2(\mu)} < \varepsilon.$$

Then,

$$\begin{aligned} Var_\mu(f) = \|f\|^2_{L^2(\mu)} &= \|Lu - f\|^2_{L^2(\mu)} + 2\int f Lu \, d\mu - \int (Lu)^2 \, d\mu \\ &\leq \varepsilon^2 - 2\int \nabla f \cdot \nabla u \, d\mu - \int (\nabla^2 \psi) \nabla u \cdot \nabla u \, d\mu \\ &\leq \varepsilon^2 + \int (\nabla^2 \psi)^{-1} \nabla f \cdot \nabla f \, d\mu, \end{aligned}$$

where we have used the fact that

$$\int (Lu)^2 \, d\mu \geq \int (\nabla^2 \psi) \nabla u \cdot \nabla u \, d\mu,$$

which follows from Bochner's formula and

$$-2x \cdot y - Ax \cdot x \leq A^{-1} y \cdot y \iff |\sqrt{A}x + \sqrt{A^{-1}}y|^2 \geq 0.$$

The desired inequality follows by letting ε tend to zero. $\square$

Remark. The Brascamp-Lieb inequality is an infinitesimal version of the Prékopa-Leindler inequality. Suppose that $f_0, f_1 : \mathbb{R}^n \to [0, \infty)$ are integrable, log-concave functions and

$$f_t(x) = \sup_{x=(1-t)y+yz} f_0(y)^{1-t} f_1(z)^t.$$

The Prékopa-Leindler inequality implies that $\log \int_{\mathbb{R}^n} f_t$ is concave in t. The second derivative in t is non-negative, and this actually amounts to the Brascamp-Lieb inequality. Thus the Brascamp-Lieb inequality is yet another incarnation of the Brunn-Minkowski theory.

We say that a function ψ on the orthant $\mathbb{R}^n_+$ is p-convex if $\psi(x_1^{1/p}, \dots, x_n^{1/p})$ is a convex function of $(x_1, \dots, x_n) \in \mathbb{R}^n_+$.

Corollary 17. *Let μ be a probability measure in the orthant $\mathbb{R}^n_+$, set $e^{-\psi} = d\mu/dx$ and assume that ψ is p-convex for $p = 1/2$. Then for any C^1-smooth function $f \in L^2(\mu)$,*

$$Var_\mu(f) \leq 4 \int_{\mathbb{R}^n} \sum_{i=1}^n x_i^2 |\partial_i f|^2 \, d\mu(x).$$

For general $p > 1$, replace the coefficient 4 by $p^2/(p-1)$.

Proof. Change variables and use the Brascamp-Lieb inequality. Denote $\frac{d\mu}{dx} = e^{-\psi}$. Then for

$$\pi(x_1, \cdots, x_n) = (x_1^2, \cdots, x_n^2),$$

the function $\psi(\pi(x))$ is convex. Set

$$\varphi(x) = \psi(\pi(x)) - \sum_{i=1}^{n} \log(2x_i).$$

Then π^{-1} pushes-forward μ to the measure with density $e^{-\varphi}$. Moreover,

$$\nabla^2\varphi(x) \geq \nabla^2 \left(-\sum_{i=1}^{n} \log(2x_i) \right) = \begin{pmatrix} \frac{1}{x_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{x_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{x_n^2} \end{pmatrix} > 0,$$

and therefore

$$\left(\nabla^2\varphi(x)\right)^{-1} \leq \begin{pmatrix} x_1^2 & 0 & \cdots & 0 \\ 0 & x_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n^2 \end{pmatrix}.$$

Set $g(x) = f(\pi(x))$. By the Brascamp-Lieb inequality,

$$Var_{e^{-\varphi}}(g) \leq \int_{\mathbb{R}_+^n} \left[\left(\nabla^2\varphi\right)^{-1} \nabla g \cdot \nabla g \right] e^{-\varphi(x)}\, dx \leq \int_{\mathbb{R}_+^n} \sum_{i=1}^{n} x_i^2 |\partial_i g(x)|^2 e^{-\varphi(x)}\, dx.$$

The corollary follows since

$$Var_{e^{-\varphi}}(g) = Var_{e^{-\psi}}(f).$$

and since when $y = \pi(x) = (x_1^2, \cdots, x_n^2)$ we have

$$x_i \partial_i g(x) = 2 y_i \partial_i f(y).$$

□

Exercise. if $\psi : \mathbb{R}_+^n \to \mathbb{R}$ is convex and increasing in all of the coordinate directions, then ψ is p-convex for $p = 1/2$, i.e., $\psi(x_1^2, \ldots, x_n^2)$ is convex in the orthant.

A function $\psi : \mathbb{R}^n \to \mathbb{R}$ is invariant under coordinate reflections (a.k.a unconditional) if

$$\psi(x_1, \ldots, x_n) = \psi(|x_1|, \ldots, |x_n|) \qquad \text{for all } x \in \mathbb{R}^n.$$

If ψ is moreover convex, then $\psi|_{\mathbb{R}_+^n}$ is increasing in all coordinate directions. The following thin shell bound from [28] is optimal.

Corollary 18. *Suppose that X is a random vector that is log-concave, isotropic and unconditional in $\mathbb{R}^n$. Then,*

$$Var(|X|) \leq C.$$

Proof.

$$\begin{aligned} Var(|X|) &\leq \mathbb{E}(|X| - \sqrt{n})^2 \leq \frac{1}{n}\mathbb{E}(|X|^2 - n)^2 = \frac{1}{n}Var(|X|^2) \\ &\leq \frac{4}{n}\sum_{i=1}^n \mathbb{E}X_i^2(2X_i)^2 = \frac{16}{n}\sum_{i=1}^n \mathbb{E}X_i^4 \leq \frac{C}{n}\sum_{i=1}^n (\mathbb{E}X_i^2)^2 \leq C \end{aligned}$$

where we used reverse Hölder inequalities in the last passage. ☐

We remark that as of May 2024, the state of affairs is that the KLS conjecture is still open already in the particular case of unconditional convex bodies. A logarithmic bound for the Poincaré constant in this case is known for years, see [28], and it is subsumed by recent bounds for the general case.

4 Gaussian Localization

Yesterday we discussed localization of a log-concave measure into *needles*, one-dimensional segments. We proceed by discussing Gaussian localization, decomposing the given measure into a mixture of measures, each of which involves multiplying the given measure by a Gaussian. The Gaussians bring with them a wealth of connections and elegant formulae, as we see below. The method was invented by Ronen Eldan [17] and it is coined *Eldan's Stochastic Localization*. We present a rather degenerate case of Eldan's method, which does not require stochastic processes.

Let Z be a standard Gaussian random vector in $\mathbb{R}^n$, of mean zero and covariance Id. Recall that for $s > 0$ write γ_s for the density of $\sqrt{s} \cdot Z$. Let X be a log-concave random vector in $\mathbb{R}^n$ independent of Z, with density ρ. For $s \geq 0$ consider the random vector

$$Y_s = X + \sqrt{s}Z$$

whose density is $\rho * \gamma_s$.

One could think of (Y_s) as a process parameterized by s, perhaps as a Browniam motion starting at the initial distribution of X. This point of view, with the time reversal $t = 1/s$, is emphasized in Joseph's lectures. In the present lecture do not consider a stochastic process parameterized by s, and view $s > 0$ as a parameter whose value will be fixed later on. One of the simplest examples of Gaussian localization of the probability density ρ is given by the following:

Proposition 19. *For $s > 0, y \in \mathbb{R}^n$ consider the probability density*

$$\rho_{s,y}(x) = \frac{\rho(x)\gamma_s(x-y)}{\rho * \gamma_s(y)},$$

which we view as a localized "Gaussian needle" or "Gaussian piece" relative to ρ*. Then the original density* ρ *is a certain average of these Gaussian needles:*

$$\rho = \mathbb{E}\rho_{s,Y_s}.$$

One says that this is a disintegration of ρ *into the localized Gaussian pieces* $(\rho_{s,y})_{s\in\mathbb{R}^n}$.

Proof. The joint density of (X, Y_s) in $\mathbb{R}^n \times \mathbb{R}^n$ is

$$(x, y) \mapsto \rho(x)\gamma_s(y - x).$$

The family of densities $\rho_{s,y}$ give us the conditional distribution of X with respect to Y_s. That is, for any test function $f(x, y)$,

$$\int_{\mathbb{R}^n} f(x, y)\rho(x)\gamma_s(y - x)dxdy = \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x, y)\rho_{s,y}(x)dx\right] \rho * \gamma_s(y)dy$$

In particular, if the function $f(x, y)$ depends only on x, we get

$$\int_{\mathbb{R}^n} f\rho = \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f\rho_{s,y}\right] \rho * \gamma_s(y)dy = \mathbb{E}\int_{\mathbb{R}^n} f\rho_{s,Y_s}.$$

□

From the proof of Proposition 19 we see that the densities $\rho_{s,y}$ give us the conditional distribution of X with respect to Y_s. The conditional expectation operator is denoted by

$$Q_s f(y) = \int_{\mathbb{R}^n} f\rho_{s,y},$$

whenever the integral converges. Thus

$$Q_s f(Y_s) = \mathbb{E}\left[f(X)|Y_s\right].$$

Assume that the original density ρ is log-concave. Then each of the elements $\rho_{s,y}$ in the decomposition is more log-concave than the Gaussian γ_s. We have thus expressed our log-concave density as a mixture of measures that are uniformly log-concave. This decomposition is determined by the choice of the parameter $s > 0$.

The critical value of s turns out to be $s \sim C_P(X)$. Roughly speaking, for much smaller values of s, we decompose into highly localized measures, maybe even resembling δ measures. For much larger values of s the decomposition is trivial for another reason: the localized pieces resemble the original measure. Abbreviate

$$\rho_s = \rho_{s,Y_s},$$

a random probability density. Recall that $\mathbb{E}\rho_s = \rho$ by Proposition 19. As usual, for a function f on $\mathbb{R}^n$ we write

$$Var_{\rho_s}(f) = \int_{\mathbb{R}^n} f^2 \rho_s - \left(\int_{\mathbb{R}^n} f \rho_s \right)^2,$$

provided that the integrals converge. Similarly, we also write $Var_\rho(f) = Var f(X)$. Then by the law of total variance,

$$Var f(X) = \mathbb{E}Var(f(X)|Y_s) + Var(\mathbb{E}(f(X)|Y_s)) = \mathbb{E}Var_{\rho_s}(f) + Var(Q_s f(Y_s))$$

When $s \gtrsim C_P(X)$, it is the first summand that is dominant:

Lemma 20. *For any $s > 0$ and a function f on $\mathbb{R}^n$ with $\mathbb{E}f^2(X) < \infty$,*

$$\mathbb{E}Var_{\rho_s}(f) \leq Var_\rho(f) \leq \left(2 + \frac{C_P(X)}{s} \right) \mathbb{E}Var_{\rho_s}(f).$$

Proof. We need to show that $Var Q_s f(Y_s)$ is not much larger than $\mathbb{E}Var_{\rho_s}(f)$. To this end, we will use the Poincaré inequality for the random vector Y_s. By the subadditivity property of the Poincaré constant,

$$C_P(Y_s) = C_P(X + \sqrt{s}Z) \leq C_P(X) + C_P(\sqrt{s}Z) = C_P(X) + s.$$

Hence

$$Var Q_s f(Y_s) \leq (C_P(X) + s) \cdot \mathbb{E}|\nabla Q_s f(Y_s)|^2.$$

Recall that

$$Q_s f(y) = \int_{\mathbb{R}^n} \rho_{s,y}(x) f(x) dx = \int_{\mathbb{R}^n} \frac{\rho(x)\gamma_s(x-y)}{\rho * \gamma_s(y)} f(x) dx.$$

Differentiating a Gaussian is easy, we have $\nabla \gamma_s(x) = -\gamma(s) \cdot x/s$. It follows that

$$\nabla Q_s f(y) = \int_{\mathbb{R}^n} \frac{x - a_s}{s} \rho_{s,y}(x) f(x) dx,$$

where $a_s = a_{s,y} = \int_{\mathbb{R}^n} x \rho_{s,y}(x) dx$ is the barycenter of the local measure $\rho_{s,y}$. Write $A_s = A_{s,y} = Cov(\rho_{s,y})$. By the Cauchy-Schwartz inequality, for $\theta \in S^{n-1}$,

$$\begin{aligned} \nabla Q_s f(y) \cdot \theta = \int_{\mathbb{R}^n} \frac{(x-a_s)\cdot\theta}{s} \rho_{s,y}(x) f(x) dx &\leq \frac{1}{s} \sqrt{\int_{\mathbb{R}^n} |(x-a_s)\cdot\theta|^2 p_{s,y}(x) dx} \sqrt{Var_{\rho_{s,y}}(f)} \\ &\leq \frac{1}{s} \sqrt{\|A_s\|_{op}} \cdot \sqrt{Var_{\rho_{s,y}}(f)}. \end{aligned}$$

Then by taking the supremum over $\theta \in S^{n-1}$,

$$Var Q_s f(Y_s) \leq \frac{C_P(X) + s}{s^2} \cdot \mathbb{E}\|A_s\|_{op} \cdot Var_{\rho_s}(f).$$

However, the random probability density ρ_s is always more log-concave than the Gaussian γ_s, and hence $A_s \leq s$. Consequently,

$$VarQ_s f(Y_s) \leq \frac{C_P(X) + s}{s} \cdot \mathbb{E}Var_{\rho_s}(f).$$

Since $Var_\rho(f)$ is the sum of the two expressions $VarQ_s f(Y_s)$ and $Var_{\rho_s}(f)$, the proposition is proven. □

To summarize, for $s \gtrsim C_P(\mu)$, the local measure ρ_s is typically close enough to the original measure, so the variance of any fixed function with respect to ρ is roughly the averaged variance with respect to ρ_s.

Remark. By differentiating with respect to s, one may improve upon Proposition 20 in two respects. First, it turns out that log-concavity is actually not needed in Proposition 20. It is proven in Klartag and Ordentlich [34] that for any random vector X and a function f with $\mathbb{E}f^2(X) < \infty$,

$$Var_\rho(f) \leq \left(1 + \frac{C_P(X)}{s}\right) \mathbb{E}Var_{\rho_s}(f). \tag{21}$$

This is a better bound than that of Lemma 20.

Corollary 21. *For any $s > 0$, setting $\alpha = s/C_P(X)$,*

$$C_P(X) \leq C\left(1 + \frac{1}{\alpha}\right) \cdot \mathbb{E}C_P(\rho_s),$$

where $C > 0$ is a universal constant.

Proof. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a 1-Lipschitz function with

$$Var_\mu(f) \geq c \cdot C_P(X),$$

whose existence in guaranteed by Corollary 13 due to E. Milman. By Proposition 20 and the Poincaré inequality,

$$Var_\mu(f) \leq \left(2 + \frac{1}{A}\right)\mathbb{E}Var_{\rho_s}(f) \leq \left(2 + \frac{1}{A}\right)\mathbb{E}C_P(\rho_s) \cdot \int_{\mathbb{R}^n} |\nabla f|^2 \rho_s \leq \left(2 + \frac{1}{A}\right)\mathbb{E}C_P(\rho_s).$$

□

Thus, in order to bound the Poincaré constant of X, we may apply Gaussian localization with $s \gtrsim C_P(\mu)$ and try to bound the Poincaré constant of ρ_s. An advantage of ρ_s over ρ is that ρ_s is more log-concave than the Gaussian γ_s. Hence, by the improved log-concave Lichnerowicz inequality, which is Theorem 15 above,

$$C_P(\rho_s) \leq \sqrt{s \cdot \|A_s\|_{op}}$$

where we recall that $A_s = Cov(\rho_s)$. Therefore, Corollary 21 leads to another corollary:

Corollary 22. *For any $s > 0$,*

$$C_P(X) \leq C \left(1 + \frac{C_P(X)}{s}\right) \cdot \sqrt{\mathbb{E}\|A_s\|_{op} \cdot s}.$$

What do we know about $\mathbb{E}\|A_s\|_{op}$? Assume from now on that X is log-concave and isotropic, so for large $s > 0$ we might expect A_s to be roughly $Cov(X) = \mathrm{Id}$. However, the operator norm involves a supremum, and this complicates matters. The evolution of the operator norm of the covariance matrix is analyzed in great detail in Joseph's lectures using stochastic processes and computations involving 3-tensors. He proves the following:

Theorem 23 (Eldan [17], Lee-Vempala [37], Chen [14], Lehec [38])**.** *Define*

$$s_0 = \min\{s > 0\,;\, \forall r > s,\ \mathbb{E}\|A_r\|_{op} \leq 5\}.$$

Then,

$$s_0 \leq C \log^2 n \tag{22}$$

where $C > 0$ is a universal constant. This bound utilizes the improved Lichnerowicz inequality, proven only recently. A slightly older bound that suffices here (e.g. [32, 38]) is

$$s_0 \leq C \log n \cdot \sup C_P(\mu)$$

where the supremum runs over all isotropic, log-concave probability measures μ on $\mathbb{R}^n$.

Moreover, $s_0 \geq c \log n$ in some examples, say when $1 + X_1, 1 + X_2, \ldots, 1 + X_n$ are independent, identically distributed standard Exponential random variables.

My guess is that stochastic processes and pathwise analysis of are not essential for the proof of Theorem 23, and that an analytic proof is possible to find. There are other applications of stochastic localization which seem to rely heavily on pathwise analysis (e.g., the complex waist inequalities in [30]). By using Theorem 23 and Corollary 22 with $s = C \log^2 n$ we thus arrive at

Corollary 24 ("best known bound for KLS")**.** *For any isotropic, log-concave random vector X in $\mathbb{R}^n$,*

$$C_P(X) \leq C \log n \tag{23}$$

where $C > 0$ is a universal constant.

Proof. We have

$$C_P(X) \leq C \left(1 + \frac{C_P(X)}{\log^2 n}\right) \cdot \sqrt{\log^2 n},$$

which implies (23). □

5 Bourgain's slicing problem

Consider a centrally-symmetric convex body $K \subseteq \mathbb{R}^n$ (i.e. $K = -K$). The maximal function operator associated with K, defined for $f : \mathbb{R}^n \to \mathbb{R}$ via

$$M_K f(x) = \sup_{r>0} \int_K f(x+ry) \frac{dy}{Vol_n(K)}.$$

Bourgain [8] proved that $\|M_K\|_{L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)} \leq C$ for a universal constant $C > 0$. This led him to study on another question, seemingly innocent:

Question 25. *Let $n \geq 2$ and suppose that $K \subseteq \mathbb{R}^n$ is a convex body of volume one. Does there exist a hyperplane $H \subseteq \mathbb{R}^n$ such that*

$$Vol_{n-1}(K \cap H) > c \tag{24}$$

for a universal constant $c > 0$?

This question is still not completely answered, and in the last four decades it emerged as an "engine" for the development of the research direction discussed in these lectures. It is shown in [33] that the bound (24) holds true if we replace the universal constant c on the right-hand side by $c/\sqrt{\log n}$. This is the currently best known result in the general case.

Theorem 26 (Hensley [23], Fradelizi [20])**.** *Let $K \subseteq \mathbb{R}^n$ be a convex body whose barycenter lies at the origin. Let X be a random vector distributed uniformly in K, and assume that $Cov(X)$ is a scalar matrix. Then for any $\theta_1, \theta_2 \in S^{n-1}$,*

$$Vol_{n-1}(K \cap \theta_1^{\perp}) \leq C \cdot Vol_{n-1}(K \cap \theta_2^{\perp})$$

where $C > 0$ is a universal constant. In fact, $C \leq \sqrt{6}$.

Proof. Let $\theta \in S^{n-1}$ and denote

$$\sigma = \sqrt{\mathbb{E}(X\cdot\theta)^2} = \sqrt{Cov(X)\theta\cdot\theta},$$

which is independent of θ. Write

$$\rho_\theta(t) = \frac{Vol_{n-1}(K \cap (t\theta + \theta^{\perp}))}{Vol_n(K)},$$

the density of the random variable $X \cdot \theta$. By the Brunn-Minkowski inequality, ρ_θ is a log-concave probability density. The log-concave random variable $X \cdot \theta/\sigma$ has mean zero and variance one, and its density is $x \mapsto \sigma\rho_\theta(x\sigma)$. According to Proposition 4 above, for any $x \in \mathbb{R}$,

$$c' 1_{\{|x| \leq c''\}} \leq \sigma\rho_\theta(x\sigma) \leq Ce^{-c|x|}$$

In particular, $c \leq \rho_\theta(0) \leq C$, for some universal constants $c, C > 0$. □

From this proof we may obtain a few more conclusions. First, that among all hyperplane sections parallel to a given hyperplane, the hyperplane section through the barycenter has the largest volume, up to a multiplicative universal constant. Second, that when $K \subseteq \mathbb{R}^n$ is a centered convex body of volume one, for any $\theta \in S^{n-1}$,

$$Vol_{n-1}(K \cap \theta^{\perp}) \cdot \sqrt{\mathbb{E}(X \cdot \theta)^2} \sim 1.$$

Here $\theta^{\perp} = \{x \in \mathbb{R}^n \,;\, x \cdot \theta = 0\}$ and we abbreviate $A \sim B$ if $c \cdot A \leq B \leq C \cdot A$ for universal constants $c, C > 0$. This leads to the following conclusion:

Corollary 27. *Let $K \subseteq \mathbb{R}^n$ be a convex body of volume one and let X be a random vector distributed uniformly on K. Then,*

$$\sup_H Vol_{n-1}(K \cap H) \sim \frac{1}{\sqrt{\|Cov(X)\|_{op}}},$$

where the supremum runs over all hyperplanes $H \subseteq \mathbb{R}^n$.

We thus see that Bourgain's slicing problem can be formulated as a question on the relation between the covariance of a convex body and its volume. Note that the logarithm of the volume of a convex body is the differential entropy of a random vector X that is distributed uniformly over the convex body. In general, when the random vector X has density ρ in $\mathbb{R}^n$, its differential entropy is

$$Ent(X) = -\int_{\mathbb{R}^n} \rho \log \rho.$$

Definition 28. *For a convex body $K \subseteq \mathbb{R}^n$ we define its isotropic constant to be*

$$L_K = \left(\frac{\det Cov(K)}{Vol_n(K)^2} \right)^{\frac{1}{2n}}$$

where $Cov(K)$ is the covariance matrix of the uniform probability distribution on K. More generally, the isotropic constant of an absolutely-continuous, log-concave random vector X in $\mathbb{R}^n$ is

$$L_X = \left(\frac{\det Cov(X)}{e^{2Ent(X)}} \right)^{\frac{1}{2n}}. \tag{25}$$

The isotropic constant of a convex body $K \subseteq \mathbb{R}^n$ of volume one governs the volumes of its hyperplane sections. From Corollary 27 we see that when $Vol_n(K) = 1$, there always exists a hyperplane section $H \subseteq \mathbb{R}^n$ with

$$Vol_{n-1}(K \cap H) \geq c / L_K.$$

Moreover, if we additionally assume that $Cov(K)$ is a scalar matrix, then for any hyperplane $H \subseteq \mathbb{R}^n$ through the barycenter of K,

$$Vol_{n-1}(K \cap H) \sim \frac{1}{L_K}.$$

The slicing problem thus asks whether L_K is universally bounded from above.

Remark on the definition of the isotropic constant in the log-concave case. Some variants of this definition exist, sometimes one replaces $Ent(X)$ by $-\log \sup \rho$ or by $-\log \rho(\mathbb{E}X)$ or by $2 \log \mathbb{E}\rho^{-1/2}(X)$, where ρ is the density of X. These variants differ at most by a multiplicative universal constant, because of the following lemma:

Lemma 29. *Denoting $\psi = -\log \rho$, we have*

$$\psi(\mathbb{E}X) \leq Ent(X) \leq \inf \psi + n$$

and

$$\mathbb{E}e^{\frac{\psi(X)}{2}} \leq e^{\frac{\inf \psi}{2} + (\ln 2)n}.$$

Proof. We may assume that ρ is continuous in $\mathbb{R}^n$ in order to neglect boundary terms in the integration by parts below. Let $y \in \mathbb{R}^n$. Then by Jensen's inequality and by the fact that any convex function lies above its tangent at X,

$$\psi(\mathbb{E}X) \leq \mathbb{E}\psi(X) = Ent(X) = \mathbb{E}\psi(X) \leq \mathbb{E}\left[\psi(y) - \nabla\psi(X) \cdot (y - X)\right] = \psi(y) + n.$$

Additionally,

$$\mathbb{E}e^{\frac{\psi(X)}{2}} = e^{\frac{\psi(y)}{2}} \int_{\mathbb{R}^n} e^{-\frac{\psi(x)+\psi(y)}{2}} dx \leq e^{\frac{\psi(y)}{2}} \int_{\mathbb{R}^n} e^{-\psi\left(\frac{x+y}{2}\right)} dx = 2^n e^{\frac{\psi(y)}{2}} \int e^{-\psi} = 2^n e^{\frac{\psi(y)}{2}}.$$

The lemma follows by taking the infimum over all $y \in \mathbb{R}^n$ in these two inequalities $\square$

It what follows we work with the definition (25). While here we are interested only in the log-concave case, the definition makes sense for any absolutely-continuous random vector X with finite second moments in $\mathbb{R}^n$. The isotropic constant measures the difference between two ways to measure the "size" of a random vector: its entropy and its covariance. Here are some basic properties of the isotropic constant:

1. It is an affine invariant, $L_{T(X)} = L_X$ for any invertible linear-affine map $T : \mathbb{R}^n \to \mathbb{R}^n$.

2. If $X_1, X_2 \in \mathbb{R}^n$ are independent log-concave random vectors, then for $X = (X_1, X_2) \in \mathbb{R}^{2n} \cong \mathbb{R}^n \times \mathbb{R}^n$,
$$L_X = \sqrt{L_{X_1} L_{X_2}}.$$

3. For any dimension n and an absolutely-continuous random vector X with finite second moments in $\mathbb{R}^n$,
$$L_X \geq \frac{1}{\sqrt{2\pi e}},$$

with equality when X is Gaussian. Indeed, it is well-known that among all random vectors with a fixed covariance in $\mathbb{R}^n$, the differential entropy is maximal for the Gaussian distribution. The proof is short. Suppose that X is a random variable of mean zero, variance one and density ρ. Then for $\gamma(x) = (2\pi)^{-1/2}\exp(-x^2/2)$ the standard Gaussian density,

$$\int_{\mathbb{R}} \rho \log \frac{\gamma}{\rho} \leq \int_{\mathbb{R}} f\left(\frac{\gamma}{\rho} - 1\right) = 0$$

so

$$Ent(X) = -\int \rho \log \rho \leq -\int \rho \log \gamma = -\int_{-\infty}^{\infty} \rho(x) \log \frac{e^{-|x|^2/2}}{\sqrt{2\pi}} dx = \log \sqrt{2\pi e}.$$

Exercise. Explain why it is not a coincidence that this universal constant $\sqrt{2\pi e}$ is "the same number" from the asymptotics $Vol_n(\sqrt{n}B^n)^{1/n} \approx 1/\sqrt{2\pi e}$.

4. Some examples:

$$L_{[0,1]^n} = \frac{1}{\sqrt{12}}, \qquad\qquad L_{\Delta^n} = \frac{(n!)^{1/n}}{(n+1)^{(n+1)/(2n)}\sqrt{n+2}} \approx \frac{1}{e}.$$

where Δ^n is a regular simplex in $\mathbb{R}^n$.

There are quite a few equivalent formulations and conditional statements, relating the isotropic constant to classical conjectures and results:

- If the isotropic constant is maximized for the cube among all centrally-symmetric convex set, then the Minkowski lattice conjecture follows, see Magazinov [40] and references therein. The Minkowski lattice conjecture suggests that if $L \subseteq \mathbb{R}^n$ is a lattice of determinant one, then each of its translates intersects the set

$$\left\{ x \in \mathbb{R}^n \, ; \, \prod_{i=1}^n |x_i| \leq \frac{1}{2^n} \right\}.$$

This was proven in two dimensions by Minkowski in 1908.

- If the isotropic constant is maximized for the simplex among all convex bodies, then the Mahler conjecture follows in the non-symmetric case. This conjecture suggests that among all convex bodies $K \subseteq \mathbb{R}^n$, the volume product

$$Vol_n(K) \cdot Vol_n(K^\circ)$$

is minimized when K is a centered simplex [31]. This was proven in two dimensions by Mahler in 1908. Here

$$K^\circ = \{x \in \mathbb{R}^n \, ; \, \forall y \in K, \ x \cdot y \leq 1\}$$

is the dual body. Recall that $(K^\circ)^\circ = K$ when K is a closed, convex set containing the origin. The Bourgain-Milman inequality resolves this conjecture up to a factor that is only exponential in the dimension. It states that for any convex body $K \subseteq \mathbb{R}^n$ containing the origin,

$$Vol_n(K) \cdot Vol_n(K^\circ) \geq (c/n)^n,$$

for a universal constant $c > 0$.

- Suppose that $K \subseteq \mathbb{R}^n$ is a convex body. Is there an ellipsoid $\mathcal{E} \subseteq \mathbb{R}^n$ with $Vol_n(\mathcal{E}) = Von_n(K)$ such that

  $$Vol_n(K \cap C\mathcal{E}) \geq \frac{1}{2} \cdot Vol_n(K)$$

  where $C > 0$ is a universal constant? This is an equivalent formulation of the slicing problem.

  Exercise. Prove the equivalence using reverse Hölder inequalities for quadratic polynomials.

  For any convex body $K \subseteq \mathbb{R}^n$, Milman's ellipsoid theorem provides an ellipsoid $\mathcal{E} \subseteq \mathbb{R}^n$ with

  $$Vol_n(K \cap C\mathcal{E}) \geq c^n \cdot Vol_n(K).$$

  This suffices for developing the Milman ellipsoid theory, which contains the quotient of subspace theorem and reverse Brunn-Minkowski and the Bourgain-Milman inequality. See Pisier [43] and references therein. The slicing problem is a conjectural strengthening of Milman's ellipsoids.

We move on to discuss the $\sqrt{\log}$-bound for the isotropic constant, and the relation to the Poincaré constant and the thin shell constants. We define

$$\sigma_n = \sup_X \sqrt{Var(|X|^2)/n}$$

where the supremum ranges over all isotropic, log-concave random vectors X in $\mathbb{R}^n$. By reverse Hölder inequalities for polynomials we may show that $Var(|X|^2)/n \sim Var(|X|)$, and hence σ_n is roughly the maximal width of the thin spherical shell that captures most of the mass of an isotropic, log-concave random vector.

(some picture here)

From Corollary 24 we know that,

$$\sigma_n \leq \sup_X \sqrt{C_P(X) \cdot 4\mathbb{E}|X|^2/n} \leq \sup_X 2\sqrt{C_P(X)} \leq C\sqrt{\log n}.$$

Hence it remains to prove:

Theorem 30 (Eldan, K. '10)**.** *For any convex body $K \subseteq \mathbb{R}^n$,*

$$L_K \leq C\sigma_n.$$

(In fact, it is shown in [18] that $L_X \leq C\sigma_n$ for any log-concave random vector X in $\mathbb{R}^n$, but for simplicity we confine ourself here for the convex body case. The slicing problem for convex bodies and for log-concave measures are known to be equivalent, as shown by Ball [3, 26].

While yesterday we studied Gaussian convolution, the proof of Theorem 30 utilizes the closely related *Laplace transform*. Let us fix an isotropic, log-concave random vector X with density ρ in $\mathbb{R}^n$. Its logarithmic Laplace transform is

$$\Lambda(y) = \Lambda_X(y) = \log \mathbb{E} e^{X \cdot y}.$$

Since a log-concave random vector has exponential moments, the logarithmic Laplace transform is finite near the origin. In fact, it is smooth in the open convex set $\Omega = \{\Lambda < \infty\}$. For $y \in \Omega$ we write X_y for a random vector with density

$$\rho_y(x) = \frac{\rho(x) e^{x \cdot y}}{e^{\Lambda(y)}}.$$

It is again a log-concave random vector, not necessarily isotropic, and we think of it as a *tilted* version of the random vector X. We comment that it is possible to view tilts using projective transformations, this leads to conditional statement that the strong slicing conjecture implies the Mahler conjecture, see [31].

Lemma 31. *For any $y \in \Omega$,*

$$\nabla \Lambda(y) = \mathbb{E} X_y, \qquad \nabla^2 \Lambda(y) = Cov(X_y), \qquad \nabla^3 \Lambda(y) = \mathbb{E}(X_y - a_y)^{\otimes 3},$$

where $a_y = \mathbb{E} X_y$.

Lemma 31 is proven by direct computation; the logarithmic Laplace transform is the cumulant generating function. We see from Lemma 31 that Λ is convex, even strongly-convex as its Hessian is positive definite. In particular the gradient $\nabla \Lambda : \Omega \to \mathbb{R}^n$ is a one-to-one map. Consider the "tilted determinant" function

$$F(y) = \log \det \nabla^2 \Lambda(y) = \log \det \nabla^2 Cov(X_y).$$

It measures how the determinant of the covariance matrix changes when we tilt the given distribution. Occasionally we may view F as a function that is defined only up to an additive constant. Write $[F]$ for the equivalence class of F under the equivalence relation "F is equivalent to G if and only if $F - G$ is a constant function".

Lemma 32. *The following bound holds pointwise in all of Ω:*

$$(\nabla^2 \Lambda)^{-1} \nabla F \cdot \nabla F \leq n\sigma_n^2. \tag{26}$$

Proof. Let us prove this bound first for $y = 0$ using the isotropicity of X. Recalling how to differentiate a determinant, we see that for any unit vector $v \in S^{n-1}$,

$$\partial_v F(0) = Trace\left[(\nabla^2\Lambda)^{-1}(0) \cdot \partial_v \nabla^2 \Lambda(0)\right] = \mathbb{E}(X \cdot v)|X|^2 \leq \sqrt{\mathbb{E}(X \cdot v)^2 \cdot Var(|X|^2)} \leq \sqrt{n}\sigma_n.$$

By considering the supremum over all $v \in S^{n-1}$, we obtain the desired bound at $y = 0$.

In order to obtain the bound for any $y \in \Omega$ we may either make a computation, or alternatively, think invariantly without computing anything, as we now explain.

Define a Riemannian metric on Ω via the Hessian of the log-Laplace transform Λ. We look at the Hessian metric (Ω, g), where the scalar product of two tangent vectors $u, v \in T_x\mathbb{R}^n \cong \mathbb{R}^n$ is

$$g_x(u, v) = \nabla^2\Lambda(x)u \cdot v.$$

The main observation is that the expression on the left-hand side of (26) is the squared Riemannian length of the Riemannian gradient of the function $F : \Omega \to \mathbb{R}$. We say that

$$\mathcal{M}_X = (\Omega, g, [F])$$

is the "Riemannian package" associated with X. This means that (Ω, g) is a Riemannian manifold and that F is a function on Ω modulo an additive constant. An isomorphism between two Riemannian packages is a bijective map which is a Riemannian isometry and transforms correctly the function modulo the additive constant.

What happens to the Riemannian package associated with X when we do various operations?

- When we translate X, the Riemannian metric stays the same, as well as the function F. We get the same Riemannian package.
- Tilting X and switching to X_y yields an isomorphism of the two Riemannian packages by *translation* by y: We translate Ω, g and $[F]$ by the vector $y \in \Omega$. Any translation corresponds to a tilt and vice versa.
- Applying an invertible linear transformation to X induces an isomorphism of the Riemannian packages. We apply a linear transformation and push forward Ω, g and $[F]$. (See also the paragraph before the next lemma).

By the first and last items, we proved (26) at the point $y = 0$ for any log-concave random vector (not necessarily centered or isotropic). By the middle item, we proved (26) also at all other points of Ω. □

It makes sense to say that we think of X as a random vector defined on an abstract affine space, rather than on $\mathbb{R}^n$, and observe that the Riemannian manifold (Ω, g) is well-defined, as well as the function $F : \Omega \to \mathbb{R}$ modulo additive constants. What can we say about balls in this Riemannian manifold?

Lemma 33. *Assume that X is a centered, log-concave random vector in $\mathbb{R}^n$. Then for any $r > 0$,*

$$\frac{1}{2} \cdot \{\Lambda \leq r\} \subseteq B_g(0, \sqrt{r}).$$

Proof. Let $y \in \Omega$ satisfy $\Lambda(2y) \leq r$. We need to find a curve from 0 to y whose Riemannian length is at most r. Let us try a line segment:

$$\begin{aligned} Length_g([0,y]) &= \int_0^1 \sqrt{\nabla^2 \Lambda(ty) y \cdot y} dt = \int_0^1 \sqrt{\frac{d^2}{dt^2} \Lambda(ty)} dt \\ &\leq \sqrt{\int_0^2 (2-t) \frac{d^2}{dt^2} \Lambda(ty) dt \cdot \int_0^1 \frac{1}{2-t} dt} \\ &= \sqrt{\log 2} \cdot \sqrt{\Lambda(2y) - [\Lambda(0) + \nabla \Lambda(0) \cdot (2y)]} = \sqrt{\log 2} \cdot \sqrt{\Lambda(2y)} \leq \sqrt{r}. \end{aligned}$$

$\square$

Let X be an isotropic random vector in $\mathbb{R}^n$, distributed uniformly in a convex body $K \subseteq \mathbb{R}^n$. We need two estimates for the proof of Theorem 30:

(i) First, we need to show that for $r = n/\sigma_n^2$,

$$Vol_n(K) \geq e^{-n} \cdot Vol_n(B_g(0, \sqrt{r})),$$

the Euclidean volume of the Riemannian ball. This is related to mass transport in a simple case.

(ii) Second, we need to show that

$$Vol_n(\{\Lambda \leq r\})^{1/n} \geq c \frac{r}{n} L_K.$$

This is related to the Bourgain-Milman inequality.

Proof of Theorem 30. Since X is isotropic and log-concave, by (i), (ii) and Lemma 33,

$$L_K = Vol_n(K)^{-1/n} \leq c \cdot Vol_n(B_g(0, \sqrt{r}))^{-1/n} \leq c \cdot Vol_n(\{\Lambda \leq r\})^{-1/n} \leq C \frac{n}{r L_K} = C \frac{\sigma_n^2}{L_K},$$

and $L_K \leq C \sigma_n$. $\square$

Proof of estimate (i): The function F vanishes at the origin, and by Lemma 32 it is a Riemannian Lipschitz function with Lipschitz constant at most $\sqrt{n}\sigma_n$. Hence,

$$|F| \leq n \qquad \text{in} \qquad B_g(0, \sqrt{r}).$$

Consequently, for any $y \in B_g(0, \sqrt{r})$,

$$e^{-n} \leq \det \nabla^2 \Lambda(y) \leq e^n.$$

We will use the fact that $\nabla\Lambda(y) = \mathbb{E}X_y \in K$ and that $y \mapsto \nabla\Lambda(y)$ is one-to-one. Changing variables, we obtain

$$Vol_n(K) \geq \int_{\nabla\Lambda(B_g(0,\sqrt{r}))} 1 dx \overset{\text{“}x=\nabla\Lambda(y)\text{''}}{=} \int_{B_g(0,\sqrt{r})} \det \nabla^2\Lambda(y) \geq e^{-n} \cdot Vol_n(B_g(0, \sqrt{r})).$$

□

Proof of estimate (ii): For any $y \in rK^\circ$,

$$\Lambda(y) = \log \mathbb{E}e^{y \cdot X} \leq \log(e^r) = r.$$

Therefore,

$$\{\Lambda \leq r\} \supseteq rK^\circ.$$

By the Bourgain-Milman inequality,

$$Vol_n(\{\Lambda \leq r\})^{1/n} \geq Vol_n(rK^\circ)^{1/n} \geq c\frac{r}{n} Vol_n(K)^{-1/n} = c\frac{r}{n} L_K.$$

□

We remark that the Bourgain-Milman inequality has several proofs, and in particular it may be proven using more delicate analysis of the log-Laplace transform as shown by Giannopoulos, Paouris and Vrisiou [21].

References

[1] G. Allaire. A la recherche de l'inégalité perdue. `https://hal.science/hal-01111806/document`, 2012.

[2] L. Ambrosio. Lecture notes on optimal transport problems. In *Mathematical aspects of evolving interfaces (Funchal, 2000)*, volume 1812 of *Lecture Notes in Math.*, pages 1–52. Springer, Berlin, 2003.

[3] K. Ball. Logarithmically concave functions and sections of convex sets in $\mathbf{R}^n$. *Studia Math.*, 88(1):69–84, 1988.

[4] K. Ball and V. H. Nguyen. Entropy jumps for isotropic log-concave random vectors and spectral gap. *Studia Math.*, 213(1):81–96, 2012.

[5] S. Bobkov. Extremal properties of half-spaces for log-concave distributions. *Ann. Probab.*, 24(1):35–48, 1996.

[6] S. Bobkov, G. Chistyakov, and F. Götze. *Concentration and Gaussian approximation for randomized sums*, volume 104 of *Probability Theory and Stochastic Modelling*. Springer, 2023.

[7] S. G. Bobkov. Remarks on the growth of L^p-norms of polynomials. In *Geometric aspects of functional analysis*, volume 1745 of *Lecture Notes in Math.*, pages 27–35. Springer, Berlin, 2000.

[8] J. Bourgain. On high-dimensional maximal functions associated to convex bodies. *Amer. J. Math.*, 108(6):1467–1476, 1986.

[9] H. J. Brascamp and E. H. Lieb. On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Functional Analysis*, 22(4):366–389, 1976.

[10] P. Buser. A note on the isoperimetric constant. *Ann. Sci. École Norm. Sup. (4)*, 15(2):213–230, 1982.

[11] L. A. Caffarelli, M. Feldman, and R. J. McCann. Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs. *J. Amer. Math. Soc.*, 15(1):1–26, 2002.

[12] A. Cayley. On monge's "mémoire sur la théorie des déblais et des remblais.". *Proc. London Math. Soc.*, s1-14(1):139–143, 1882. Available at `http://dx.doi.org/10.1112/plms/s1-14.1.139`.

[13] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. *Problems in analysis, Princeton Univ. Press*, pages 195–199, 1970.

[14] Yuansi Chen. An almost constant lower bound of the isoperimetric coefficient in the KLS conjecture. *Geom. Funct. Anal.*, 31(1):34–61, 2021.

[15] N. De Ponti and A. Mondino. Sharp Cheeger-Buser type inequalities in RCD(K, ∞) spaces. *J. Geom. Anal.*, 31(3):2416–2438, 2021.

[16] P. Diaconis and D. Freedman. Asymptotics of graphical projection pursuit. *Ann. Statist.*, 12(3):793–815, 1984.

[17] R. Eldan. Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geom. Funct. Anal.*, 23(2):532–569, 2013.

[18] R. Eldan and B. Klartag. Approximately Gaussian marginals and the hyperplane conjecture. In *Concentration, functional inequalities and isoperimetry*, volume 545 of *Contemp. Math.*, pages 55–68. Amer. Math. Soc., Providence, RI, 2011.

[19] L. C. Evans and W. Gangbo. Differential equations methods for the Monge-Kantorovich mass transfer problem. *Mem. Amer. Math. Soc.*, 137(653):viii+66, 1999.

[20] Matthieu Fradelizi. Hyperplane sections of convex bodies in isotropic position. *Beiträge Algebra Geom.*, 40(1):163–183, 1999.

[21] A. Giannopoulos, G. Paouris, and B.-H. Vritsiou. The isotropic position and the reverse Santaló inequality. *Israel J. Math.*, 203(1):1–22, 2014.

[22] M. Gromov and V. D. Milman. A topological application of the isoperimetric inequality. *American Journal of Mathematics*, 105(4):843–854, 1983.

[23] D. Hensley. Slicing convex bodies—bounds for slice area in terms of the body's covariance. *Proc. Amer. Math. Soc.*, 79(4):619–625, 1980.

[24] R. Kannan, L. Lovász, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete & Computational Geometry*, 13:541–559, 1995.

[25] L. V. Kantorovich and G. P. Akilov. *Functional analysis*. Pergamon Press, Oxford-Elmsford, N.Y., second edition, 1982. Translated from the Russian by Howard L. Silcock.

[26] B. Klartag. On convex perturbations with a bounded isotropic constant. *Geom. Funct. Anal.*, 16(6):1274–1290, 2006.

[27] B. Klartag. A central limit theorem for convex sets. *Invent. Math.*, 168(1):91–131, 2007.

[28] B. Klartag. A Berry-Esseen type inequality for convex bodies with an unconditional basis. *Probab. Theory Related Fields*, 145(1-2):1–33, 2009.

[29] B. Klartag. Needle decompositions in Riemannian geometry. *Mem. Amer. Math. Soc.*, 249(1180):v+77, 2017.

[30] B. Klartag. Eldan's stochastic localization and tubular neighborhoods of complex-analytic sets. *J. Geom. Anal.*, 28(3):2008–2027, 2018.

[31] B. Klartag. Isotropic constants and Mahler volumes. *Adv. Math.*, 330:74–108, 2018.

[32] B. Klartag. On yuansi chen's work on the KLS conjecture. *Lecture notes prepared for a winter school at the Hausdorff Institute*, 2021. Available at `https://www.him.uni-bonn.de/fileadmin/him/Lecture_Notes/chen_lecture_not`

[33] B. Klartag. Logarithmic bounds for isoperimetry and slices of convex sets. *Ars Inveniendi Analytica*, Paper No. 4, 17pp, 2023.

[34] B. Klartag and O. Ordentlich. The strong data processing inequality under the heat flow. *In preparation*, 2024.

[35] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[36] M. Ledoux. Spectral gap, logarithmic Sobolev constant, and geometric bounds. In *Surveys in differential geometry. Vol. IX*, volume 9 of *Surv. Differ. Geom.*, pages 219–240. Int. Press, Somerville, MA, 2004.

[37] Y. T. Lee and S. Vempala. Eldan's stochastic localization and the KLS conjecture: Isoperimetry, concentration and mixing. *Ann. of Math. (2)*, 199(3):1043–1092, 2024.

[38] J. Lehec. Lecture notes on isoperimetric inequalities in high-dimensional convex sets. *Institut Henri Poincaré (IHP), Paris*, 2024.

[39] A. Lichnerowicz. *Géométrie des groupes de transformations*, volume III of *Travaux et Recherches Mathématiques*. Dunod, Paris, 1958.

[40] A. Magazinov. A proof of a conjecture by Haviv, Lyubashevsky and Regev on the second moment of a lattice Voronoi cell. *Adv. Geom.*, 20(1):117–120, 2020.

[41] E. Milman. On the role of convexity in isoperimetry, spectral gap and concentration. *Invent. Math.*, 177(1):1–43, 2009.

[42] L. E. Payne and H. F. Weinberger. An optimal Poincaré inequality for convex domains. *Arch. Rational Mech. Anal.*, 5:286–292, 1960.

[43] G. Pisier. *The volume of convex bodies and Banach space geometry*, volume 94 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1989.

[44] V. N. Sudakov. Typical distributions of linear functionals in finite-dimensional spaces of high dimension. *Dokl. Akad. Nauk SSSR*, 243(6):1402–1405, 1978. English translation: Soviet Math. Dokl. 19 (1978), no. 6, 1578–1582 (1979).

[45] G. Szegö. Inequalities for certain eigenvalues of a membrane of given area. *J. Rational Mech. Anal.*, 3:343–356, 1954.

[46] H.F. Weinberger. An isoperimetric inequality for the n-dimensional free membrane problem. *J. Rational Mech. Anal.*, 5:633–636, 1956.