# Optimal Transport and Convex Localization

**Bo'az Klartag**
Department of Mathematics
Weizmann Institute of Science

**Version: May 31, 2023**

Scribe: Arianna Piana

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

# Contents


**1 Poincaré inequality** **5**

1.1 Framework . . . 5
1.2 Poincaré inequality . . . 6
1.3 Part of Poincaré's motivation . . . 7
1.4 Conductance and isoperimetric inequality . . . 10
1.4.1 Co-aerea inequality . . . 11
1.4.2 From Co-aerea inequality to Cheeger's inequality . . . 13
1.5 Isoperimetric inequality . . . 16
1.6 Brunn-Minkowski inequality . . . 17
1.7 *log*-concave random vectors . . . 19

**2 Convex localization** **21**

2.1 Hyperplane bisection . . . 21
2.2 Needle decomposition . . . 23
2.3 Reverse Hölder inequality . . . 27

**Index** **33**

# Chapter 1

# Poincaré inequality

The Poincaré inequality is a powerful tool in mathematical analysis that is useful for a wide range of applications. One of its main uses is in the study of partial differential equations, particularly in the analysis of solutions to the Laplace equation. The inequality can be used to derive estimates for the solutions to the Laplace equation, which can be helpful in understanding the behavior of physical systems or geometric structures.

Another important application of the Poincaré inequality is in the field of Markov Chain Monte Carlo (MCMC) methods. MCMC methods are used to simulate complex systems by generating a sequence of samples that converge to the desired distribution. However, these methods can be computationally expensive, especially when the target distribution has a complicated geometry. The Poincaré inequality can be used to derive bounds on the mixing time of the Markov chain, which is the time it takes for the chain to converge to the target distribution. These bounds can be used to optimize the MCMC algorithm and improve its efficiency.

## 1.1 Framework

We consider **metric-measure space** $(X, d, \mu)$ where

- $X$ is a complete and separable space (e.g. metric space, $\mathbb{R}^n$, $\mathbb{S}^n$, Riemann manifold, $\{0,1\}^n$, ...),
- $d$ is a distance function,
- $\mu$ is a Borel measure.

If $\mu$ is a probability measure (i.e. $\mu(X) = 1$), it is called **metric probability space**.

**Example 1.1.1.** $(K, d_E, \lambda)$, where $K \subset \mathbb{R}^n$ is convex, $d_E$ is the Euclidean distance and $\lambda$ is the Lebesgue measure restricted on $K$.

**Example 1.1.2.** $\mathbb{S}^n$ equipped with the geodesic distance and the uniform measure.

**Example 1.1.3.** $(\mathbb{R}^n, d_E, \gamma_n)$ where $\gamma_n$ is the standard Gaussian measure, i.e.

$$\frac{\mathrm{d}\gamma_n}{\mathrm{d}x} = (2\pi)^{-\frac{n}{2}} e^{-\frac{||x||^2}{2}}$$

**Example 1.1.4.** $G = (V, E)$ weighted graph with weight function $w : G \to \mathbb{R}_+$ equipped with distance

$$d(u, v) = \min_{\gamma(u,v)} \ell(\gamma),$$

where $\gamma(u, v)$ denotes a path from $u, v$ and $\ell(\gamma) = \sum_{e \in \gamma} w(e)$.

## 1.2 Poincaré inequality

The original proof of the Poincaré inequality can be traced back to the work of Henri Poincaré himself, who introduced the inequality as part of his study of differential equations in the late 19th century. Poincaré's work on the inequality was published in several papers, including "*Sur les équations aux dérivées partielles de la physique mathématique*" (On the partial differential equations of mathematical physics) [12]. For the Poincaré inequality on convex sets and its history consult [1].

**Theorem 1.2.1.** *Let $K \subset \mathbb{R}^n$ be a convex, bounded and open set. Let $f : K \to \mathbb{R}$ be a $C^1$-smooth function with $\int_K f \mathrm{d}\lambda = 0$. Then*

$$\int_K f^2 \mathrm{d}\lambda \leq C_P(K) \int_K |\nabla f|^2 \mathrm{d}\lambda, \tag{1.2.1}$$

*where $C_P(K) = 2^{n-1} diam^2(K)$.*

**Remark 1.2.2.** $C_P(K)$ is called the Poincaré constant, but it's not optimal.

**Remark 1.2.3.** The assumption $\int_K f \mathrm{d}\lambda = 0$ can be removed and equation (1.2.1) can be restated as

$$\mathrm{Var}_K(f) \leq C_P(K) \int_K |\nabla f|^2 \mathrm{d}\lambda \,. \tag{1.2.2}$$

The convexity condition ensures that the distance between any two points in the set is well-defined and bounded above by the diameter of the set. This property is important because it allows us to use geometric arguments to derive the Poincaré inequality.

Now, regarding the example of a body with a small bottleneck, let us consider a two-

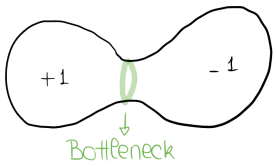



Figure 1.1: Non-convex example of a 2-dimensional body with bottleneck.

dimensional set with the shape of an hourglass, as in Figure 1.1. This set is not convex, and in fact, it has a bottleneck that connects two larger regions. If we try to apply the Poincaré inequality to a function defined on this set, we may run into problems. The bottleneck region will cause the gradient of the function to become large, while the function itself may not vary much, leading to a large ratio between the integrals of the square of the function and the square of its gradient. As a result, the constant $C_P(K)$ may become very large, which can limit the usefulness of the inequality in this case.

*Proof.* Wlog assume $\mathrm{Vol}_n(K) = 1$ (otherwise scale). Note that using the assumption $\int_K f \mathrm{d}\lambda = 0$ we can write

$$\int_K f^2(x) \mathrm{d}\lambda(x) = \frac{1}{2} \int_K \int_K |f(x) - f(y)|^2 \mathrm{d}x \mathrm{d}y \,.$$

By fundamental theorem of calculus we have

$$\begin{aligned} f(x) - f(y) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} f\big((1-t)x + ty\big) \mathrm{d}t \\ &= \int_0^1 \nabla f\big((1-t)x + ty\big)(y - x) \mathrm{d}t \,. \end{aligned}$$

Thus,

$$|f(x) - f(y)| \leq \operatorname{diam}(K) \int_0^1 |\nabla f\big((1-t)x + ty\big)| \mathrm{d}t \,.$$

Taking the squares and applying Cauchy-Schwarz inequality we get

$$|f(x) - f(y)|^2 \leq \operatorname{diam}^2(K) \int_0^1 \big|\nabla f\big((1-t)x + ty\big)\big|^2 \mathrm{d}t \,.$$

We integrate over $K$

$$\begin{aligned}\int_K f^2(x) \mathrm{d}\lambda(x) &= \frac{1}{2}\operatorname{diam}^2(K) \int_K \int_K \int_0^1 \big|\nabla f\big((1-t)x + ty\big)\big|^2 \mathrm{d}x\mathrm{d}y\mathrm{d}t \\ &= \operatorname{diam}^2(K) \int_{\frac{1}{2}}^1 \int_K \int_K \big|\nabla f\big((1-t)x + ty\big)\big|^2 \mathrm{d}x\mathrm{d}y\mathrm{d}t \,,\end{aligned}$$

where in the second line we exploit the symmetry in $x$ and $y$ of the integrals. Fix $x$, we change variable $z(y) = (1-t)x + ty$ for $t \geq \frac{1}{2}$ (thus $\mathrm{d}z = t^n \mathrm{d}y$). Then

$$\begin{aligned}\int_K \big|\nabla f\big((1-t)x + ty\big)\big|^2 \mathrm{d}y &= \int_{(1-t)x + tK \subset K} |\nabla f(z)| \frac{\mathrm{d}z}{t^n} \\ &\leq 2^n \int_K |\nabla f(z)|^2 \mathrm{d}z.\end{aligned}$$

Hence,

$$\begin{aligned}\int_K f^2(x) \mathrm{d}\lambda(x) &\leq \left( \int_{\frac{1}{2}}^1 \mathrm{d}t \int_K \mathrm{d}x \int_K |\nabla f(z)^2| \mathrm{d}z \right) 2^n \operatorname{diam}^2(K) \\ &= 2^{n-1} \operatorname{diam}^2(K) \int_K |\nabla f(z)^2| \mathrm{d}z \,.\end{aligned}$$

□

**Remark 1.2.4.** If $K = B^n = \{x \in \mathbb{R}^n : ||x|| \leq 1\}$, then $C_P(K) < \frac{C}{n}$ for some universal constant $C > 0$. In this case Poincaré inequality is not optimal. The best constant was found by Payne-Weinberg (gaps were fixed later) in [11].

**Theorem 1.2.5.** *$\forall K \subset \mathbb{R}^n$ convex, open and bounded we have*

$$C_P(K) \leq \frac{\operatorname{diam}^2(K)}{\pi^2} \,.$$

**Exercise 1.2.6.** Show that for $K = [0, \pi]$, Theorem 1.2.5 is tight.

## 1.3 Part of Poincaré's motivation

The **heat equation (or heat flow)** [5] (Chapter 4) is a classical partial differential equation that describes the evolution of the temperature in a physical system over time. The **heat equation with Neumann boundary condition** is a specific version of the heat equation that takes into account the flux of heat at the boundary of the system, instead of specifying the temperature at the boundary. This condition is often used when the boundary is insulated. The Neumann boundary condition states that the derivative of the temperature with respect to the normal direction at the boundary is equal to a given function.

Let $K \subset \mathbb{R}^n$ be a smooth, open, bounded set, not necessary convex. Let us write as $u_t(x)$, for $t \geq 0$ and $x \in K$ the temperature at the point $x \in K$ at time $t$. The evolution of temperature in time is given by

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} u_t(x) = \Delta u_t(x) & x \in K \\ \langle \nabla u_t(x), \nu(x) \rangle = 0 & x \in \partial K \end{cases} \tag{1.3.1}$$

where $\nu$ is the outer unit normal vector.

The **Fourier law** relates the flux of heat through a boundary to the temperature gradient at that boundary. It states that the heat flux is proportional to the negative gradient of the temperature. In other words, if we have a boundary $\partial\Omega$ with unit normal vector $\nu$ and temperature $u(x,t)$, then the flux of heat through this boundary is given by $-k\frac{\partial u}{\partial \nu}$ where $k$ is the thermal conductivity of the material. The minus sign indicates that heat flows from higher temperature to lower temperature.

A classical result from PDE class is the following

**Theorem 1.3.1.** *For all smooth $u_0 : K \to \mathbb{R}$, it exists an unique solution to (1.3.1) starting from $u_0$ and it is smooth.*

**Lemma 1.3.2.** *(Preservation of total heat) $\int_K u_t$ is constant in $t$.*

*Proof.*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_K u_t = \int_K \frac{\mathrm{d}}{\mathrm{d}t} u(t,x) = \int_K \Delta u_t \mathrm{d}x = \int_{\partial K} \langle \nabla u_t, \nu \rangle = 0\,,$$

by divergence theorem. □

**Proposition 1.3.3.** *Wlog suppose $\mathrm{Vol}_n(K) = 1$. For all $u_0 : K \to \mathbb{R}$ such that $u_0$ is $C^1$-smooth and $\int_K u_0(x)\mathrm{d}\lambda(x) = 1$ then it holds*

$$||u_t - 1||^2_{L^2(K)} \leq e^{-\frac{t}{C_P(K)}} ||u_0 - 1||^2_{L^2(K)}\,, \tag{1.3.2}$$

*where $C_P(K)$, the Poincaré constant, is the relaxation time of the heat flow.*

Before proving this proposition, we need some tools from spectral theory.

**Theorem 1.3.4.** ***(Spectral theorem for Laplacian)*** *There exist an orthonomal basis $\{\varphi\}_i$ such that*

$$1 = \varphi_0, \varphi_1, \varphi_2, \ldots, \in L^2(K) \text{ smooth}$$

*and eigenvalues*

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \to \infty$$

*such that for all $i$*

$$\begin{cases} \Delta \varphi_i(x) = -\lambda_i \varphi_i(x) & x \in K \\ \langle \nabla \varphi_i(x), \nu(x) \rangle = 0 & x \in \partial K \end{cases} \tag{1.3.3}$$

**Remark 1.3.5.** $\varphi_i$ is the minimizer of the **Rayleigh quotient** , i.e.

$$\lambda_i = \inf_{0 \neq u \perp \varphi_0, \varphi_1, \ldots, \varphi_{i-1}} \frac{\int_K |\nabla u(x) \mathrm{d}\lambda(x)|^2}{\int_K u^2(x)\mathrm{d}\lambda(x)} =: R_K(u)\,, \tag{1.3.4}$$

where $u : K \to \mathbb{R}$ is $C^1$-smooth.

An example for a solution to the heat equation (1.3.1) is

$$u_t(x) = e^{-\lambda_i t}\varphi_i(x)\,.$$

In fact,

$$-\lambda_i u_t(x) = \frac{\mathrm{d}}{\mathrm{d}t}u_t(x) = e^{-\lambda_i t}\Delta\varphi_i = -\lambda_i u_t(x)\,.$$

Let us now define the **heat operator**

$$\begin{aligned} P_t : L^2(K) &\to L^2(K) \\ u_0 &\mapsto P_t(u_0) = u_t \end{aligned}$$

The heat operator is diagonal in the basis $\varphi_0, \varphi_1, \ldots$, i.e.

$$P_t(\varphi_i) = e^{-\lambda_i t}\varphi_i\,.$$

Note that $\lambda_1 = \frac{1}{C_P(K)}$ because

$$\lambda_1 = \inf\left\{\frac{\int_K |\nabla u(x)|^2 \mathrm{d}\lambda(x)}{\int_K u(x)^2 \mathrm{d}\lambda(x)} \text{ such that } \int_K u(x)\mathrm{d}\lambda(x) = 0\right\}$$

and therefore

$$\lambda_1 \int_K u(x)^2 \mathrm{d}\lambda(x) \leq \int_K \nabla u(x)^2 \mathrm{d}\lambda(x)$$

for all $u : K \to \mathbb{R}$ such that $\int_K u(x)\mathrm{d}\lambda(x) = 0$.

*Proof. (of Proposition 1.3.3)* Fix $u_0 : K \to \mathbb{R}$ with $\int_K u_0(x)\mathrm{d}\lambda(x) = 1$. We expand in the othonomal basis

$$u_0 = \sum_{i=0}^{\infty} a_i\varphi_i = 1 + \sum_{i=1}^{\infty} a_i\varphi_i\,,$$

where $a_i = \langle u_i, \varphi_i\rangle$ and $a_0 = \int_K u_0(x)\mathrm{d}\lambda(x) = 1$. Now,

$$u_t = P_t(u_0) = 1 + \sum_{i=1}^{\infty} a_i e^{\lambda_i t}\varphi_i\,,$$

where $\lambda_i \geq \lambda_1 = \frac{1}{C_P(K)}$. Therefore we have

$$\begin{aligned} ||u_t - 1||^2_{L^2(K)} &= \sum_{i=1}^{\infty} |a_i|^2 e^{-2\lambda_i t} \\ &\leq e^{-\frac{2t}{C_P(K)}} \sum_{i=1}^{\infty} |a_i|^2 \\ &= e^{-\frac{2t}{C_P(K)}} ||u_0 - 1||_{L^2(K)}\,. \end{aligned}$$

□

**Exercise 1.3.6.** Assume the existence of $\varphi_i = R_K(u)$, where $R_K$ is the Rayleigh quotient as in (1.3.4). Assume smoothness. Prove that for all $v$

$$R_K(u) \leq R_K(u + \varepsilon v)\,.$$

## 1.4 Conductance and isoperimetric inequality

The Poincaré constant can be interpreted as a measure of the speed at which a heat distribution converges to its average value on the domain.

In the context of analyzing the convergence properties of Markov chain Monte Carlo (MCMC) algorithms, the Poincaré constant is often used as a measure of how well-connected a given set or domain is. A higher Poincaré constant implies slower mixing behavior, meaning it takes longer for the MCMC algorithm to explore the state space and converge to the target distribution.

Intuitively, a domain with a low Poincaré constant allows heat to diffuse more easily throughout the domain, meaning that it is more connected and less fragmented. This means that random walks or MCMC chains on the domain are more likely to move around and explore different parts of the domain, leading to faster convergence. Conversely, a domain with a high Poincaré constant is more disconnected and has more barriers to movement, which can cause the MCMC chain to get stuck in local minima and slow down the convergence rate.

Therefore, the Poincaré constant can be seen as a measure of the conductance or connectivity of a set. It provides a quantitative measure of how easily a random walk or MCMC chain can explore the set, and can be used to compare the relative efficiency of different sets or domains in this regard.

**Definition 1.4.1.** For all open $K \subset \mathbb{R}^n$ we define the **Cheeger constant** (or **isoperimetric constant**) as

$$h_K = \inf_{A \subset K} \frac{\mathrm{Vol}_{n-1}(\partial A \cap K)}{\min\left\{\mathrm{Vol}_n(A), \mathrm{Vol}_n(K \setminus A)\right\}}\,, \tag{1.4.1}$$

wher $\partial A$ is smooth.

Poincaré inequality and the Cheeger's are strongly related.

**Theorem 1.4.2.** ***(Cheeger's inequality)*** *Under general assumptions it holds*

$$h_K^2 \cdot C_P(K) \leq 4\,.$$

**Theorem 1.4.3.** *[8] If $K \subset \mathbb{R}^n$ is convex then*

$$\frac{1}{9} \leq h_K^2 \cdot C_P(K)\,.$$

We will prove Theorem 1.4.3 with optimal transport techniques.

For metric-measure spaces and Cheeger's inequality refer to [3] and Chapter $3\frac{1}{2}$ of [7].

The Poincaré inequality can be derived from the isoperimetric inequality using the following argument. First, consider a smooth function $u$ that satisfies $\int_K u(x)\mathrm{d}\lambda(x) = 0$ (we can always assume that by subtracting a suitable constant). Let $B$ be a ball of volume equal to that of $K$, and consider the function $v(x) = \int_{\partial B} u(y)\mathrm{d}\sigma(y)$, where $\sigma$ denotes the surface measure on $\partial B$. By the isoperimetric inequality, $\sigma(\partial B) \leq \sigma(\partial K)$, which implies that $v$ is well-defined and belongs to $H_0^1(K)$. Moreover, one can show that $\nabla v(x) = \int_{\partial B} \nabla G(x,y)u(y)\mathrm{d}\sigma(y)$, where $G(x,y)$ is the Green's function for the Laplacian in $\mathbb{R}^n$.

Using this representation of $\nabla v$, one can show that

$$\int_K |u(x)|^2 \mathrm{d}\lambda(x) \leq C \int_K |\nabla v(x)|^2 \mathrm{d}\lambda(x)$$

where $C$ is a constant depending on the dimension $n$. Then, using the Cauchy-Schwarz inequality, one can obtain the Poincaré inequality for general $u$.

**Remark 1.4.4.** Let $K \subset \mathbb{R}^n$ be a convex body. Then the Cheeger constant of $K$ is attained and satisfies

$$\frac{1}{2} \leq h_K \,. \tag{1.4.2}$$

Furthermore, equality holds in the left inequality if and only if $K$ is a Euclidean ball of volume $\frac{1}{2}$.

### 1.4.1 Co-aerea inequality

In the following, consider a metric measurable space $(X, d, \mu)$ and wlog assume $\mu(X) = 1$. Let us denote $A \subset X$ a measurable set.

**Definition 1.4.5.** For $A \subset X$, the **surface area** is defined as

$$\mu^+(\partial A) = \lim_{\varepsilon \to 0} \inf \frac{\mu(N_\varepsilon(A)) - \mu(A)}{\varepsilon} \,,$$

where

$$N_\varepsilon(A) = \{x \in X : d(x, A) < \varepsilon\} \,,$$

is the neighborhood of $A$.

**Exercise 1.4.6.** Show that $\mu^+$ is not a measure.

**Definition 1.4.7.** We say $f : X \to \mathbb{R}$ is **Lipschitz on balls** if

$$\forall B(x_0, r) \; \exists L \text{ such that } |f(x) - f(y)| \leq L d(x, y) \; \forall x, y \in B(x_0, r) \,. \tag{1.4.3}$$

**Remark 1.4.8.** The gradient (or accumulation point) of $f$ is defined as

$$|\nabla f(x)| = \lim_{\varepsilon \to 0} \sup_{0 < d(x,y) < \varepsilon} \frac{|f(y) - f(x)|}{d(x, y)} \,. \tag{1.4.4}$$

We say that $x \in X$ is an **isolated point** if $|\nabla f(x)| = 0$.

**Definition 1.4.9.** $C_P(X)$ is the **infimal** $C \geq 0$ such that for all $f : X \to \mathbb{R}$ Lipschitz on balls, it holds

$$\mathrm{Var}_\mu(f) \leq C \int_X |\nabla f(x)|^2 \mathrm{d}\lambda(x) \,.$$

**Lemma 1.4.10.** ***(Co-area inequality)*** *Let $f : X \to \mathbb{R}$ be a function Lipschitz on balls. Then it holds*

$$\int_0^\infty \mu^+\big(\partial\{f \geq t\}\big)\mathrm{d}t \leq \int_X |\nabla f(x)| \mathrm{d}\mu(x) \,, \tag{1.4.5}$$

*where $\{f \geq t\} = \{x \in X : f(x) = t\}$.*

**Remark 1.4.11.** For all $f : \mathbb{R}^n \to \mathbb{R}$ which are $C^1$, it holds

$$\int_{\mathbb{R}} \mathrm{Vol}_{n-1}\big(\{f = t\}\big)\mathrm{d}t = \int_{\mathbb{R}^n} |\nabla f(x)| \mathrm{d}\lambda(x) \,.$$

**Remark 1.4.12.** For all $f : X \to [0, \infty)$ which are $\mu$-integrable it holds

$$\int_X f(x)\mathrm{d}\mu(x) = \int_X \int_0^\infty \mathbb{1}_{\{f(x) \geq t\}} \mathrm{d}t \mathrm{d}\mu(x) = \int_0^\infty \mu\big(\{f \geq t\}\big)\mathrm{d}t \,.$$

**Remark 1.4.13.** Under some regularity assumptions, (1.4.5) is an equality.

*Proof.* We split the proof in two cases. Suppose that $f$ is Lipschitz on balls.
$\boxed{f \text{ bounded}}$ Let $h > 0$. We may add a constant to $f$ and make it non-negative. Set

$$f_h(x) = \sup_{d(x,y)<h} f(y)\,.$$

Clearly $f_h(x) \geq f(x)$. Also, for all $t \geq 0$ we have $\{f_h > t\} = N_h(\{f > t\})$. Therefore, we have

$$\int_X f_h(x)\mathrm{d}\mu(x) = \int_0^\infty \mu\big(\{f_h > t\}\big)\mathrm{d}t = \int_0^\infty \mu\big(N_h(\{f > t\})\big)\mathrm{d}t$$

and clearly

$$\int_X f(x)\mathrm{d}\mu(x) = \int_0^\infty \mu\big(\{f > t\}\big)\mathrm{d}t\,.$$

Therefore we get

$$\int_X \frac{f_h - f}{h}\mathrm{d}\mu = \int_0^\infty \frac{\mu\big(N_h(\{f > t\})\big) - \mu\big(\{f > t\}\big)}{h}\mathrm{d}t\,. \tag{1.4.6}$$

The argument of the integral of LHS of (1.4.6) is bounded by the Lipschitz constant of $f$ and is non-negative, and therefore by bounded convergence theorem

$$\int_X \frac{f_h - f}{h}\mathrm{d}\mu \leq \int_X \lim_{h\to 0^+}\sup \frac{f_h - f}{h}\mathrm{d}\mu\,.$$

By Fatou's Lemma we have for the RHS of (1.4.6)

$$\int_0^\infty \frac{\mu\big(N_h(\{f > t\})\big) - \mu\big(\{f > t\}\big)}{h}\mathrm{d}t \geq \int_0^\infty \lim_{h\to 0^+}\inf \frac{\mu\big(N_h(\{f > t\})\big) - \mu\big(\{f > t\}\big)}{h}\mathrm{d}t\,.$$

Thus we get

$$\begin{aligned}\int_X |\nabla f(x)|\mathrm{d}\mu(x) &\geq \int_X \lim_{h\to 0^+}\sup \frac{f_h - f}{h}\mathrm{d}\mu \\ &\geq \int_0^\infty \lim_{h\to 0^+}\inf \frac{\mu\big(N_h(\{f > t\})\big) - \mu\big(\{f > t\}\big)}{h}\mathrm{d}t \\ &\geq \int_0^\infty \mu^+\big(\{f > t\}\big)\mathrm{d}t\,,\end{aligned}$$

and we conclude.
$\boxed{f \text{ unbounded}}$ For $f$ unbounded, we want to truncate the function. Consider

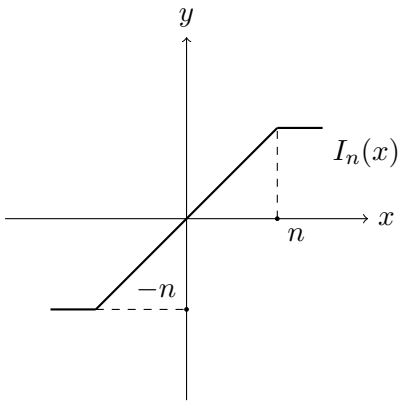
$$I_n(x) = \begin{cases} x & \text{if } x \in [-n, n] \\ n & \text{if } x > n \\ -n & \text{if } x < n \end{cases}$$

Note that the function $I_n$ is 1-Lipschitz. Set $f_n = I_n \circ f$. Then it holds

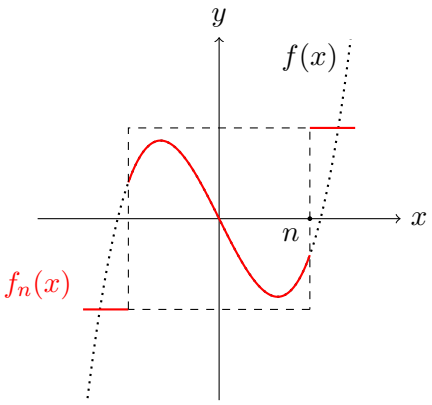
$$\big|f_n(x) - f_n(y)\big| = \big|I_n(f(x)) - I_n(f(y))\big| \leq \big|f(x) - f(y)\big|\,,$$

and therefore $|\nabla f_n| \leq |\nabla f|$. Hence it holds for all $n$

$$\begin{aligned}\int_X |\nabla f|\,\mathrm{d}\mu &\geq \int_X |\nabla f_n|\,\mathrm{d}\mu \\ &\geq \int_{-\infty}^{\infty} \mu^+\left(\partial\left\{f_n > t\right\}\right)\mathrm{d}t \\ &\geq \int_{-n}^{n} \mu^+\left(\partial\left\{f_n > t\right\}\right)\mathrm{d}t \\ &= \int_{-n}^{n} \mu^+\left(\partial\left\{f > t\right\}\right)\mathrm{d}t \overset{n\to\infty}{\longrightarrow} \int_{-\infty}^{\infty} \mu^+\left(\partial\left\{f > t\right\}\right)\mathrm{d}t\,.\end{aligned}$$




(a) Plot of $I_n$.




(b) Truncation of $f$, namely $f_n = I_n \circ f$.

□

### 1.4.2 From Co-aerea inequality to Cheeger's inequality

We first prove a $L^1 - L^1$ Poincaré inequality and then we show that this implies $L^2 - L^2$ Poincaré inequality.

**Corollary 1.4.14.** *($L^1 - L^1$ **Poincaré inequality**) Let $f \in L^1(\mu)$ be a Lipschitz function. Assume that 0 is a median of $f$, i.e.*

$$\mu(\{f \geq 0\}) \geq \frac{1}{2} \qquad \text{and} \qquad \mu(\{f \leq 0\}) \geq \frac{1}{2}. \tag{1.4.7}$$

*Then*

$$h_X \cdot \int_X |f| \, \mathrm{d}\mu \leq \int_X |\nabla f| \, \mathrm{d}\mu. \tag{1.4.8}$$

**Exercise 1.4.15.** Replace the median (1.4.7) with expectation

**Exercise 1.4.16.** Show that (1.4.8) implies $L^p - L^p$ Poincaré inequality, namely

$$h_X \cdot \int_X |f|^p \, \mathrm{d}\mu \leq \int_X |\nabla f|^p \, \mathrm{d}\mu. \tag{1.4.9}$$

*Proof.* By Lemma 1.4.10

$$\begin{aligned} \int_X |\nabla f| \, \mathrm{d}\mu &\geq \int_\infty^\infty \mu^+(\partial\{f > t\}) \, \mathrm{d}t \\ &= \int_0^\infty \mu^+(\partial\{f > t\}) \, \mathrm{d}t. \end{aligned}$$

By isoperimetric inequality, for all $t > 0$

$$\mu^+(\partial\{f > t\}) \geq h_X \cdot \min\{\mu(\{f > t\}), 1 - \mu(\{f > t\})\} = h_X \cdot \mu(\{f > t\}),$$

and for all $t < 0$

$$\mu^+(\partial\{f > t\}) \geq h_X \cdot \mu(\{f \leq t\}).$$

Combining the previous equations we have

$$\begin{aligned}\int_X |\nabla f|\,\mathrm{d}\mu &\geq \int_{-\infty}^{\infty} \mu^+\left(\partial\{f>t\}\right)\mathrm{d}t\\ &\geq h_X\left(\int_{-\infty}^{0}\mu\left(\{-f\geq -t\}\right)\mathrm{d}t+\int_0^{\infty}\mu\left(\{f>t\}\right)\mathrm{d}t\right)\\ &= h_X\left(\int_X f_-\mathrm{d}\mu+\int_X f_+\mathrm{d}\mu\right)\\ &= h_X\int_X |f|\,\mathrm{d}\mu\,.\end{aligned}$$

□

**Remark 1.4.17.** If 0 is a median, then it holds

$$\int_X |f|\,\mathrm{d}\mu = \inf_{t\in\mathbb{R}}\int_X |f-t|\,\mathrm{d}\mu\,.$$

**Remark 1.4.18.** If $\int_X f\mathrm{d}\mu = 0$ then

$$\int_X f^2\mathrm{d}\mu = \inf_{t\in\mathbb{R}}\int_X (f-t)^2\mathrm{d}\mu,\,. \tag{1.4.10}$$

**Exercise 1.4.19.** Prove that Corollary 1.4.14 works for graphs.

**Exercise 1.4.20.** Prove that the inequality is tight, if $f=\mathbb{1}_A$ the limit is obtained for sets $A$ with finite perimeter.

**Lemma 1.4.21.** *1. Corollary 1.4.14 works for functions that are Lipschitz on balls.*

*2. It holds*

$$\left|\nabla f^2\right| \leq 2\,|f|\,|\nabla f|\,. \tag{1.4.11}$$

*3. If $f$ is Lipschitz on balls, the so is $f^2$.*

We are now ready to prove Theorem 1.4.2 given Lemma 1.4.21.

*Proof. (of Theorem 1.4.2)* Let $f\in L^2(\mu)$ be Lipschitz on balls. Assume first that 0 is a median. Set $f_+=\max\{f,0\}$ and $f_-=\max\{-f,0\}$. Then $f_+^2$ still has 0 as a median so by Corollary 1.4.14 and equation (1.4.11),

$$\begin{aligned}h_X\int_X f_+^2\mathrm{d}\mu &\leq \int_X \left|\nabla f_+^2\right|\mathrm{d}\mu \leq 2\int_{\{f_+>0\}}|f_+|\,|\nabla f_+|\,\mathrm{d}\mu\\ h_X\int_X f_-^2\mathrm{d}\mu &\leq 2\int_{\{f_->0\}}|f_-|\,|\nabla f_-|\,\mathrm{d}\mu\,.\end{aligned}$$

We sum both terms and apply Cauchy-Schwarz

$$h_X\int_X f^2\mathrm{d}\mu \leq 2\int_X |f|\,|\nabla f|\,\mathrm{d}\mu \leq 2\sqrt{\int_X f^2\mathrm{d}\mu\int_X |\nabla f|^2\,\mathrm{d}\mu}\,.$$

So

$$\frac{h_X^2}{4}\int_X f^2\mathrm{d}\mu \leq \int_X |\nabla f|^2\,\mathrm{d}\mu\,.$$

Now, by (1.4.10)

$$\mathrm{Var}_\mu(f) = \inf_{m\in\mathbb{R}}\int_X (f-m)^2\,\mathrm{d}\mu \leq \int f^2\mathrm{d}\mu \leq 4h_X^{-2}\int_X |\nabla f|^2\,\mathrm{d}\mu \tag{1.4.12}$$

We may also add a constant and make 0 a median. □

Finally, let us prove the auxiliary lemma.

*Proof. (of Lemma 1.4.21)*

1. Let $r > 0$. Define

$$I_r = \begin{cases} 1 & \text{if } x < r \\ 1 + r - x & \text{if } x \in [r, r+1] \\ 0 & \text{if } x > r \end{cases}$$

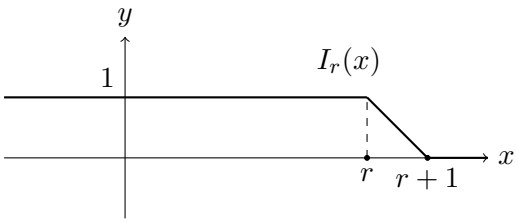



Figure 1.3: Plot of $I_r$.

Fix $a \in X$. Write $\eta_r(x) = I_r\big(d(a,x)\big)$ and set $f_r(x) = f(x)\eta_r(x)$. Therefore,

$$|\nabla f_r(x)| = \begin{cases} |\nabla f(x)| & \text{if } d(a,x) < r \\ 0 & \text{if } d(a,x) > r+1 \end{cases}$$

Moreover, if $r < d(a,x) < r+1$ then

$$|\nabla f_r(x)| \le |\nabla f(x)| + |f(x)|$$

We don't know the two values $d(a,x) = r$ or $r+1$. On the annulus $r < d(a,x) < r+1$ we have

$$\eta_r(x) = r + 1 - d(a,x).$$

So

$$\begin{aligned} |f_r(x) - f_r(y)| &\le |f(x)(\eta_r(x) - \eta_r(y))| + |\eta_r(y)(f(y) - f(x))| \\ &\le |f(x)|\, d(x,y) + |f(y) - f(x)|, \end{aligned}$$

and therefore

$$\frac{|f_r(x) - f_r(y)|}{d(x,y)} \le \frac{|f(x) - f(y)|}{d(x,y)} + |f(x)|.$$

Thus,

$$|\nabla f_r(x)| \le |\nabla f(x)| + |f(x)|.$$

The set where $\mu(\{d(a,x) = r\}) > 0$ is countable. Let us write $m_r$ the median of $f_r$ and let $r \to \infty$ outside this countable set

$$h_X \int |f_r - m_r| \, \mathrm{d}\mu \le \int |\nabla f_r| \, \mathrm{d}\mu \le \underbrace{\int_{B(a,r+1)} |\nabla f| \, \mathrm{d}\mu}_{\overset{r\to\infty}{\longrightarrow} \int_X |\nabla f| \mathrm{d}\mu} + \underbrace{\int_{r<d(a,x)<r+1} |f| \, \mathrm{d}\mu}_{\overset{r\to\infty}{\longrightarrow} 0 \text{ since } \int_X |f| \mathrm{d}\mu < \infty}.$$

Note that $f_r = f$ on $B(a,r)$. Pick a subsequence of $r$'s for which $m_r \to m \in \mathbb{R}$ (this limit exists since $m_r$ has a subsequence which converges by compactness and it cannot be infinite since $f \in L^1(\mu)$) To summarize $\exists m \in \mathbb{R}$

$$h_X \int |f - m| \, \mathrm{d}\mu \le \int |\nabla f| \, \mathrm{d}\mu.$$

2.

$$\begin{aligned}\left|\nabla f^2(x)\right| &= \lim_{\varepsilon\to 0}\sup_{y\in B(x,\varepsilon)}\frac{\left|f^2(y)-f^2(x)\right|}{d(x,y)}\\ &\le \lim_{\varepsilon\to 0}\sup_{y\in B(x,\varepsilon)}\frac{|f(y)-f(x)|}{d(x,y)}\cdot\lim_{\varepsilon\to 0}\sup_{y\in B(x,\varepsilon)}|f(y)+f(x)|\\ &= 2|\nabla f||f(x)|\ .\end{aligned}$$

3. If $f$ is Lipschitz on balls so $x,y\in B$

$$\left|f^2(x)-f^2(y)\right|\le|f(x)-f(y)|\cdot 2\sup_B|f|\le\tilde{L}\cdot d(x,y)\ .$$

**Remark 1.4.22.** Why Lipschitz on balls? Set $(X,d,\mu)=(\mathbb{R}^n,d_E,\mu)$ Then $\forall f\in\mathbb{R}^n\to\mathbb{R}$ which is a $C^1$-function

$$\mathrm{Var}_\mu(f)\le 4h_X^{-2}\int_{\mathbb{R}^n}|\nabla f|^2\,d\mu$$

**Example 1.4.23.** Let $X$ be an exponential random variable, i.e. it has distribution for all $t>0$

$$\mathbb{P}(X\ge t)=e^{-t}\,.$$

Prove that $\forall f$ which is $C^1$-continuous

$$\mathrm{Var}_\mu f(X)\le 4\mathbb{E}_\mu\left|f'(X)\right|^2\,. \tag{1.4.13}$$

Use Cheeger's inequality for $(\mathbb{R}^+,d(x,y)=|x-y|,\mu)$ where $\frac{\mathrm{d}\mu}{\mathrm{d}x}=e^{-x}=:\rho(x)$.

**Remark 1.4.24.** In Example 1.4.23 4 is optimal. Set $f(x)=e^{\frac{x}{2}(1-\varepsilon)}$ with $\varepsilon>0$. To see this, plug in this function $f$ in (1.4.9), both the LHS and the RHS can be computed explicitly. Then let $\varepsilon$ tend to zero, and see that the ratio between RHS and LHS tends to 4.
In fact for all open subsets $\forall A\subseteq\mathbb{R}^+$

$$\mu^+(\partial A)\ge 1\cdot\min\{\mu(A),1-\mu(A)\}\,.$$

For $A=[a,b]$, $\mu^+(\partial A)=e^{-a}+e^{-b}$. Also, $A=(x,\infty)$ where $x>0$

$$\begin{aligned}e^{-x}=\mu^+(\partial A)&\ge\min\left\{\int_x^\infty e^{-t}dt,\int_0^x e^{-t}dt\right\}\\ &=\min\left\{e^{-x},1-e^{-x}\right\}.\end{aligned}$$

□

## 1.5 Isoperimetric inequality

The isoperimetric inequality is a fundamental result in geometry that establishes a relationship between the perimeter or boundary length of a set and its area or volume. It provides a quantitative measure of how "fat" or "thin" a shape is.

The study of isoperimetric inequalities dates back to ancient Greece, where mathematicians were interested in finding shapes with a given area that minimize the length of their boundaries. The problem can be stated as follows: Among all sets with a fixed area, which one has the smallest possible perimeter? The isoperimetric inequality provides an answer to this question by proving that among all sets in a given space with a fixed

measure (area or volume), the shape that minimizes the perimeter (boundary length) is a sphere or a ball.

The proof of the isoperimetric inequality can be approached using various mathematical techniques, including calculus of variations, geometric arguments, and integral geometry. One popular method to prove the isoperimetric inequality is through the Brunn-Minkowski inequality.

**Theorem 1.5.1.** ***(Isoperimetric inequality)*** *For all $A \subset \mathbb{R}^n$ open, bounded and with smooth boundary, it holds*

$$\frac{\mathrm{Vol}_{n-1}(\partial A)}{\mathrm{Vol}_n(A)^{\frac{n-1}{n}}} \geq \frac{\mathrm{Vol}_{n-1}(S^{n-1})}{\mathrm{Vol}_n(B^n)^{\frac{n-1}{n}}} \tag{1.5.1}$$

## 1.6 Brunn-Minkowski inequality

The Brunn-Minkowski inequality is a powerful geometric inequality that relates the volumes of two sets to the sum of their volumes. It provides a bridge between geometric properties and measure theory. By utilizing the Brunn-Minkowski inequality, we can establish the isoperimetric inequality.

The Brunn-Minkowski inequality states that for two sets $A$ and $B$ in Euclidean space, the volume of their Minkowski sum (sum of the sets scaled by positive coefficients) is greater than or equal to the sum of their individual volumes. This inequality captures the intuitive notion that "adding more material" to two sets increases their combined volume.

**Definition 1.6.1.** Let $A, B \subset X$. The **Minkowski sum** is defined as

$$A + B = \left\{x + y : x \in A \text{ and } y \in B\right\}.$$

In $X = \mathbb{R}^n$ we can discuss neighborhoods of a set $A$ $N_\varepsilon(A)$ as Minkowski sum $A + \varepsilon B^n$.

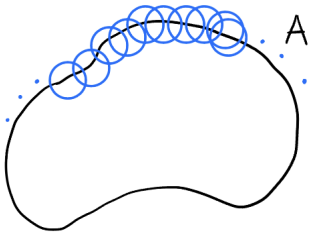



Figure 1.4: The resulting set, which represents the neighborhood $N_\varepsilon(A)$ consists of all the points obtained by "summing" the points in $A$ with $\varepsilon B^2$. Essentially, you would have a collection of scaled disks centered at each point of $A$

**Theorem 1.6.2.** ***(Brunn-Minkowski inequality)*** *For all non-empty and Borel subsets $A, B \subseteq \mathbb{R}^n$ it holds*

$$\mathrm{Vol}_n(A + B)^{1/n} \geq \mathrm{Vol}_n(A)^{1/n} + \mathrm{Vol}_n(B)^{1/n} \tag{1.6.1}$$

**Corollary 1.6.3.** *For all $A \subseteq \mathbb{R}^n$ open, bounded and with smooth boundary, it holds*

$$\mathrm{Vol}_{n-1}(\partial A) \geq \mathrm{Vol}_n(A)^{\frac{n-1}{n}} \cdot n \cdot \mathrm{Vol}_n(B^n)^{\frac{1}{n}},$$

*with equality for $A = B^n$.*

*Proof.*

$$\begin{aligned}\mathrm{Vol}_{n-1}(\partial A) &= \lim_{\varepsilon\to 0}\frac{\mathrm{Vol}_n(A+\varepsilon B^n)-\mathrm{Vol}_n(A)}{\varepsilon}\\ &\overset{(1.6.1)}{=} \lim_{\varepsilon\to 0}\frac{\left(\mathrm{Vol}_n(A)^{1/n}+\varepsilon\mathrm{Vol}_n(B^n)^{1/n}\right)^n-\mathrm{Vol}_n(A)}{\varepsilon}\\ &= n\mathrm{Vol}_n(A)^{\frac{n-1}{n}}\mathrm{Vol}_n(B^n)^{\frac{1}{n}}\,.\end{aligned}$$

□

We give the proof of Brunn-Minkowski theorem based on *elementary sets*. An elementary set is a finite union of non-overlapping boxes whose edges are parallel to the coordinate axes. See also [2], Section 1.2.2. We argue by induction over the number of boxes $m$.

*Proof. (of Theorem 1.5.1)* $\boxed{m=2}$ Let $A, B \subseteq \mathbb{R}^n$ are parallel boxes. Then $A+B$ is also a box and

$$A+B=\prod_{i=1}^n[0,a_i+b_i).$$

We need to prove

$$\prod_{i=1}^n(a_i+b_i)^{\frac{1}{n}} \geq \left(\prod_{i=1}^n a_i\right)^{\frac{1}{n}}+\left(\prod_{i=1}^n b_i\right)^{\frac{1}{n}},$$

which is equivalent to

$$1 \geq \left(\prod_{i=1}^n \frac{a_i}{a_i+b_i}\right)^{\frac{1}{n}}+\prod_{i=1}^n\left(\frac{b_i}{a_i+b_i}\right)^{\frac{1}{n}}.$$

By arithmetic-geometric means inequality shows that the RHS is less than or equal to

$$\frac{1}{n}\left(\sum_{i=1}^n\frac{a_i}{a_i+b_i}+\sum_{i=1}^n\frac{b_i}{a_i+b_i}\right)=1\,.$$

$\boxed{\text{(IH)}}$ Assume that for $m \geq 3$ the statement holds true for all pairs of elementary sets with complexity $\leq m-1$.
$\boxed{m-1\to m}$ Suppose there are at least 3 boxes at total. Say that $A$ consists of in least 2 boxes. Since they are not overlapping, we can separate them by a coordinate hyperplane $\mathcal{H}_n=\{x\in\mathbb{R}^n : x_n=\rho\}$ for some $\rho\in\mathbb{R}$. Define

$$A^+=A\cap\{x\in\mathbb{R}^n : x_n\geq\rho\} \quad \text{and} \quad A^-=A\cap\{x\in\mathbb{R}^n : x_n\leq\rho\}.$$

Since $A^+$ and $A^-$ are not overallping elementary sets, each one consists of a smaller number of boxes than $A$.

Find a second hyperplane $\tilde{\mathcal{H}}_n=\{x\in\mathbb{R}^n : x_n=s\}$ such that if

$$B^+=B\cap\{x\in\mathbb{R}^n : x_n\geq s\} \quad \text{and} \quad B^-=B\cap\{x\in\mathbb{R}^n : x_n\leq s\},$$

then

$$\frac{\mathrm{Vol}_n(A^+)}{\mathrm{Vol}_n(A)}=\frac{\mathrm{Vol}_n(B^+)}{\mathrm{Vol}_n(B)}=:\lambda.$$

Clearly $0 < \lambda < 1$. Since $A + B \supseteq (A^+ + B^+) \cup (A^- + B^-)$ we have

$$\begin{aligned}
\mathrm{Vol}_n(A+B) &\geq \mathrm{Vol}_n\left(A^+ + B^+\right) + \mathrm{Vol}_n\left(A^- + B^-\right) \\
&\overset{\text{(IH)}}{\geq} \left(\mathrm{Vol}_n(A^+)^{\frac{1}{n}} + \mathrm{Vol}_n(B^+)^{\frac{1}{n}}\right)^n + \left(\mathrm{Vol}_n(A^-)^{\frac{1}{n}} + \mathrm{Vol}_n(B^-)^{\frac{1}{n}}\right)^n \\
&= \left(\lambda^{\frac{1}{n}}\mathrm{Vol}_n(A)^{\frac{1}{n}} + \lambda^{\frac{1}{n}}\mathrm{Vol}_n(B)^{\frac{1}{n}}\right)^n + \left((1-\lambda)^{\frac{1}{n}}\,\mathrm{Vol}_n(A)^{\frac{1}{n}} + (1-\lambda)^{\frac{1}{n}}\,\mathrm{Vol}_n(B)^{\frac{1}{n}}\right)^n \\
&= \left(\mathrm{Vol}_n(A)^{\frac{1}{n}} + \mathrm{Vol}_n(B)^{\frac{1}{n}}\right)^n.
\end{aligned}$$

We can now approximate a Borel set by an elementary set (finite union of boxes) and theorem is proved. □

**Remark 1.6.4.** The Brunn-Minkowksi inequality can also be written in the following multiplicative form. For all $A, B \subseteq \mathbb{R}^n$ and $0 < \lambda < 1$

$$\mathrm{Vol}_n\big((1-\lambda)A + \lambda B\big) \geq \mathrm{Vol}_n(A)^{1-\lambda} \cdot \mathrm{Vol}_n(B)^{\lambda}. \tag{1.6.2}$$

In fact, by arithmetic-geometric inequality, the LHS of (1.6.2) can be bounded by

$$\begin{aligned}
\mathrm{Vol}_n\big((1-\lambda)A + \lambda B\big) &\geq \left((1-\lambda\mathrm{Vol}_n(A)^{\frac{1}{n}} + \mathrm{Vol}_n(B)^{\frac{1}{n}}\right)^n \\
&\geq \mathrm{Vol}_n(A)^{1-\lambda}\mathrm{Vol}_n(B)^{\lambda}.
\end{aligned}$$

## 1.7 *log*-concave random vectors

**Definition 1.7.1.** A random vector $X$ in $\mathbb{R}^n$ is **log-concave** if it is supported in some affine subspace $E \subseteq \mathbb{R}^n$ and has a density $\rho$ in $E$ which is log-concave $\forall x, y \in E, 0 < \lambda < 1$,

$$\rho\left(\lambda x + (1-\lambda)\, y\right) \geq \rho\left(x\right)^{\lambda} \rho\left(y\right)^{1-\lambda} \tag{1.7.1}$$

i.e. $\rho = e^{-H}$ where $H$ is convex.

**Remark 1.7.2.** If the random vector $X$ is absolutely continuous, then it is enough to say that $X$ is log-concave if it is supported on convex open set and satisfies (1.7.1).

**Example 1.7.3.** Gaussian vector and Uniform measure on a convex set.

**Corollary 1.7.4.** *Let $K \subseteq \mathbb{R}^n$ be convex, open, bounded set. Let $X$ be a random vector distributed uniformly in $K$. Let $E \subseteq \mathbb{R}^n$ be a subspace. Then* $\mathrm{Proj}_E(X)$ *is log-concave random vector.*

*Proof.* Wlog assume $\mathrm{Vol}_n(K) = 1$. Write $f_E$ for density of $\mathrm{Proj}_E X$ i.e. for every $x \in E$

$$f_E\left(x\right) = \mathrm{Vol}_n\left(K \cap \left(E^{\perp} + x\right)\right).$$

We need $f_E\left(\frac{x+y}{2}\right) \geq \sqrt{f_E\left(x\right) f_E\left(y\right)}$. But since

$$K \cap \left(\frac{x+y}{2} + E^{\perp}\right) \supseteq \frac{K \cap \left(x + E^{\perp}\right) + K \cap \left(y + E^{\perp}\right)}{2}$$

and by Brunn-Minkowksi

$$\underbrace{\mathrm{Vol}_n\left(K \cap \left(\frac{x+y}{2} \cap E^{\perp}\right)\right)}_{=f_E\left(\frac{x+y}{2}\right)} \geq \sqrt{\underbrace{\mathrm{Vol}_n\left(K \cap \left(x + E^{\perp}\right)\right)}_{f_E(X)} \underbrace{\mathrm{Vol}_n\left(K \cap \left(y + E^{\perp}\right)\right)}_{f_E(y)}}.$$

□

**Exercise 1.7.5.** Let $X$ be a log-concave random vector and $E \subseteq \mathbb{R}^n$ a subspace. Prove that $\mathrm{Proj}_E(X)$ is log-concave.

**Exercise 1.7.6.** Deduce that if $X, Y$ are independent log-concave random vectors then $X + Y$ is log-concave.

**Exercise 1.7.7.** Prove that the pointwise product of log-concave measure is log-concave.

**Exercise 1.7.8.** Prove that log-concavity is closed under weak limits.

**Corollary 1.7.9.** *Let $X$ be a* log*-concave random vector. Let $K \subseteq \mathbb{R}^n$ convex set with $\mathbb{P}(X \in K) > 0$. Then the condition on $X\big|_K$ is still* log*-concave.*

# Chapter 2

# Convex localization

Our goal is to improve the Poincaré constant $C_P$ of Theorem 1.2.1. We focus on the following version of Poincaré inequality, due to L. E. Payne, Hans F Weinberger, in [11].

**Theorem 2.0.1.** ***(Poincaré inequality, Payne-Weinberger constant)*** *Let $K \subseteq \mathbb{R}^n$ be a convex set. Let $\mu$ be a* log*-concave measure on $K$. Then for all $f$ $C^1$-functions with $\int_K f \mathrm{d}\mu = 0$ it holds*

$$\int_K f^2 \mathrm{d}\mu \leq \frac{\mathrm{diam}^2(K)}{\pi^2} \int_K |\nabla f|^2 \mathrm{d}\mu \,.$$

We will prove Theorem 2.0.1 using **convex localization**, a method based on **hyperplane bisections**. The idea behind is very intuitive: we want to divide one big problem in two slightly smaller problems and work separately with each subproblem.

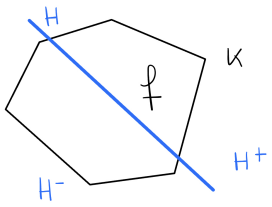



Figure 2.1: We find a hyperplane $H$ that divides $K$ such that the integral in the left part and the integral in the right part are also null.

So, our goal is to find a hyperplane $H$ that divides the convex set $K$ such that $\int_{K\cap H^-} f \mathrm{d}\mu = \int_{K\cap H^+} f \mathrm{d}\mu = 0$ and prove Theorem 2.0.1 for each $K \cap H^{\pm}$.

## 2.1 Hyperplane bisection

The following topological observations plays a key role in our discussion.

**Lemma 2.1.1.** *Let $E_0 \subseteq \mathbb{R}^n$ be an affine subspace of dimension $n-2$. Let $f : K \to \mathbb{R}$ be a continuous and p-integrable function. Let $\mu$ be an absolutely continuous measure. Then there exists an hyperlplane $H$ containing $E_0$ such that*

$$\int_{K\cap H^-} f \mathrm{d}\mu = \int_{K\cap H^+} f \mathrm{d}\mu = 0$$

*Proof.* We have $E_0^{\perp} \cong \mathbb{R}^2$. For all $\theta \in E_0^{\perp}$ unit vector, let us define $H(\theta) = E_0 + \mathbb{R}\theta$. Note that each line is attained twice ($\theta$ and $-\theta$).
Let

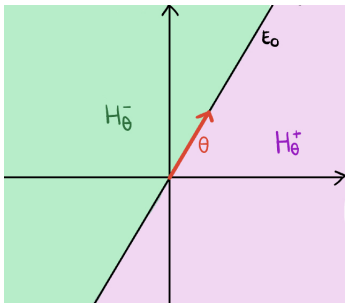



Figure 2.2: $H(\theta)^+$ and $H(\theta)^-$.

$$H(\theta)^+ = \{\text{all points in } E_0^\perp \text{ to the right of } \theta\} \times E_0^\perp\,,$$
$$H(\theta)^- = \{\text{all points in } E_0 \text{ to the left of } \theta\} \times E_0\,.$$

Clearly $H(\theta)^+$ and $H(\theta)^-$ are half spaces and

$$H(\theta)^+ = H(-\theta)^+ \quad H(-\theta)^+ = H(\theta)^-\,.$$

Let $\theta \in S^1$. Consider the map

$$\theta \mapsto \int_{K\cap H(\theta)^+} f\mathrm{d}\mu - \int_{K\cap H(\theta)^-} f\mathrm{d}\mu\,.$$

First, by dominated convergence theorem, this is a continuous map. Also, it is an odd map in $\theta$. Therefore, by intermediate value theorem, the map vanishes, i.e. it exists $\vartheta$ such that

$$\int_{K\cap H(\vartheta)^+} f\mathrm{d}\mu - \int_{K\cap H(\vartheta)^-} f\mathrm{d}\mu = 0.$$

We conclude by recalling that it also holds

$$\int_{K\cap H(\theta)^+} f\mathrm{d}\mu + \int_{K\cap H(\theta)^-} f\mathrm{d}\mu = 0.$$

□

Let $K$ be a convex set in $\mathbb{R}^n$. We apply Lemma 2.1.1 and partition it in two subset $K^+$ and $K^-$ that are disjoint up to measure zero and satisfy

$$\int_{K^+} f\mathrm{d}\mu = \int_{K^-} f\mathrm{d}\mu = 0.$$

So it's enough to prove Theorem 2.0.1 separately on each piece $K^\pm$, i.e.

$$\int_{K^\pm} f^2\mathrm{d}\mu \le \frac{\mathrm{diam}^2(K)}{\pi^2}\int_{K^\pm} \left|\nabla f\right|^2\mathrm{d}\mu\,.$$

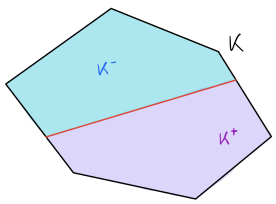



Figure 2.3: We partition $K$ in two disjoint subsets.

In a second step we bisect separately each convex piece and we get

$$K = K_1 \cup K_2 \cup K_3 \cup K_4 \text{ disjoint up to measure } 0$$

with each $K_i$ convex and $\int_{K_i} f \mathrm{d}\mu = 0$ for $i = 1, \dots, 4$.

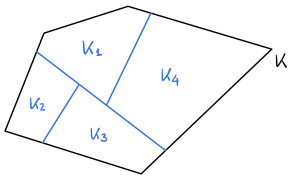



Figure 2.4: We partition $K$ in four disjoint subsets.

At step $n$, we will have $2^n$ pieces with

$$K = \bigcup_{i=1}^{2^n} K_i \,,$$

with $K_i$ convex, disjoint (up to measure 0) and

$$\int_{K_i} f \mathrm{d}\mu = 0 \quad \forall i = 1, \dots, 2^n \,.$$

We need also an other ingredient: we want to make these small pieces almost one-dimensional.
Chose the $(n-2)$-dimensional affine subspaces $E_1, E_2, \dots \subset \mathbb{R}^n$ such that

1. at step $i$ of the bisection process, we apply Lemma 2.1.1 with $E_i$ and
2. we want all pieces to be approximately one-dimensional.

This is called **needle decomposition**. For references see [6, 9, 10].

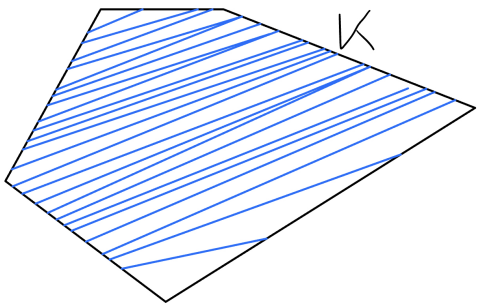



Figure 2.5: Needle decomposition of $K$.

## 2.2 Needle decomposition

**Definition 2.2.1.** Let $\delta > 0$. A convex set $P \subseteq \mathbb{R}^n$ is **$\delta$-needle** if it exists a line $\ell \subset \mathbb{R}^n$ such that $P \subseteq \ell + \delta B^n$.

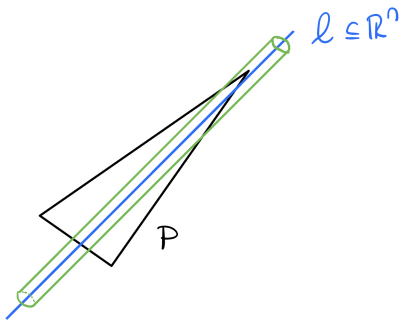



Figure 2.6: $\delta$-needle.

**Proposition 2.2.2.** *Let $\delta > 0$. Then it exists $N \geq 1$ and $(n-2)$-dimensional subspaces $E_1, \dots, E_N \subseteq \mathbb{R}^n$ such that when applying the hyperplane bisection process, we obtain*

$$K = \bigcup_{i=1}^{2^N} K_i \,,$$

*and $K_i$ are $\delta$-needles for all $i = 1, \dots, 2^N$.*

We define the **affine Grassmannian**

$$\mathtt{AG}_{n,k} = \{k\text{-dimensional affine subspaces in } \mathbb{R}^n\}$$

equipped with the topology of sub-basis of open sets, $(n-k)$-dimensional, relatively open balls $B \subseteq \mathbb{R}^n$. Set

$$\mathcal{U} = \{E \in \mathtt{AG}_{n,k} \text{ that pierce it, i.e.} \#(E \cap B) = 1\}$$

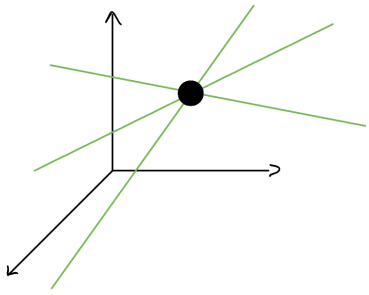


Figure 2.7: $\mathtt{AG}_{3,1}$.

*Proof.* Pick a dense subsequence $E_0, E_1, \dots \in \mathtt{AG}_{n,n-2}$. Fix $\delta$. We will show that it exists $N \geq 1$ such that the claim holds. Set $\tilde{\delta} = \frac{\delta}{10}$.

Step 1 It exists $N \geq 1$ such that for any relatively open 2-dimensional disc $D \subset R^n$ of radius at least $\tilde{\delta}$, which is contained in $K \subseteq \mathbb{R}^n$ (where $K$ is a bounded convex set), there exists an index $i \in \{1, \dots, N\}$ such that $\#(E_i \cap D) = 1$. We present a compactness argument and argue by contradiction:

$$\forall N \; \exists D \subseteq \mathbb{R}^2 \text{ 2-dim disc of radius at least } \tilde{\delta} : \bigcup_{i=1}^{N} E_i \text{ does not pierce } D. \tag{2.2.1}$$

Set

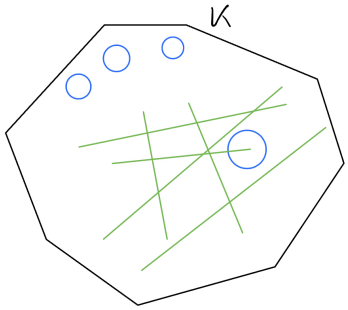



Figure 2.8: Graphical representation of (2.2.1).

$$\Omega_N = \{\text{2-dim discs, relatively open, compact in } K, \text{ of radius at least } \tilde{\delta}\}.$$

We claim that $\Omega_N$ is compact.

The subset $\Omega_N$ of all discs pierced by $\bigcup_{i=1}^{N} E_i$ is open and therefore the complement is compact. Also, by compactness, $\Omega_n$ is decreasing in $n$ and therefore $\bigcap_{n=1}^{\infty} \Omega_n \neq \emptyset$. Hence it exists $D \subseteq \Omega$ which is not pierced by any $E_i$. We get a contradiction to the fact that $E_1, E_2, \ldots$ is a dense subsequence.

Step 2 With this choice of $E_1, E_2, \ldots$ any convex piece arising in this localization (hyperplane bisection, needle decomposition) process can not contain a 2-dimensional disc of radius at least $\tilde{\delta}$. Let $H_1, \ldots, H_N$ be halfspaces and set $K_i = K \cap (H_1 \cap H_2 \cap \cdots \cap H_i)$. We know that for all $i$ it holds $E_i \subseteq \partial H_i$. This means that the interior of the convex set $K_i$ is disjoint from $\partial H_1 \cup \cdots \cup \partial H_i$, so it is disjoint from $E_1 \cup \cdots \cup E_N$. So $K_i$ can not contain a 2-dimensional disc of radius at least $\tilde{\delta}$.

Step 3 We need to show that all pieces in the decomposition are $\delta$-needles. Take a piece $K_i$. We know that it does not contain a 2-dimensional disc of radius at least $\tilde{\delta}$. Take $x, y \in \bar{K}$ as far as possible. Then $||x - y|| = \text{diam}(K_i)$. Let $\ell$ be a line passing trough $x$ and $y$. We need to show $K_i \subseteq \ell + \delta B^n$.

If $|x - y| < \delta$, there is nothing to prove. If $|x - y| \geq \delta$: given a triangle $xyz$ with basis and height both of at least $\delta$, then the triangle contains a ball of radius at least $\frac{\delta}{10}$. □

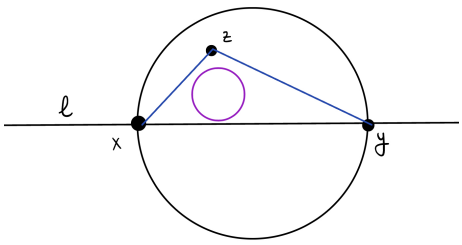



Figure 2.9: Graphical representation of the case $|x - y| \geq \delta$.

**Example 2.2.3.** Take the unit square $K = \left[-\frac{1}{2}, \frac{1}{2}\right]^2$, with the Lebesgue measure and $f(x) = x_1$. The needle decomposition looks like Figure 2.10.

**Example 2.2.4.** Figure 2.11 shows the needle decomposition of the unit circle.

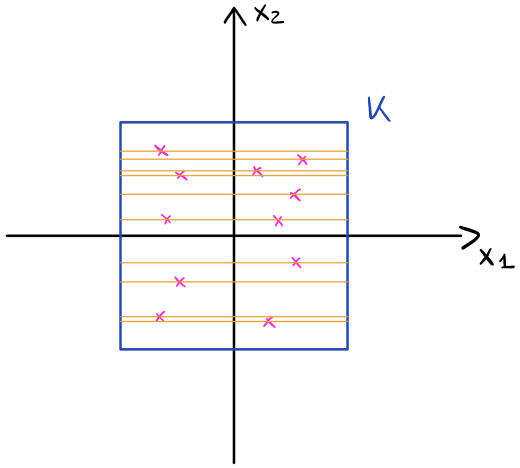



Figure 2.10: Needle decomposition of $K = \left[-\frac{1}{2}, \frac{1}{2}\right]^2$.

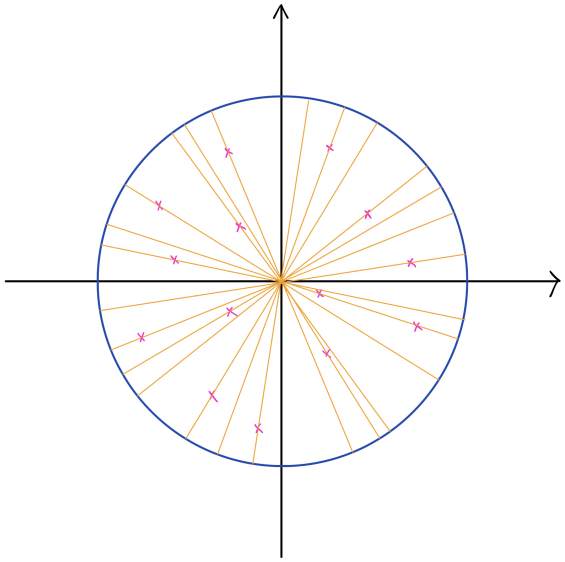


Figure 2.11: Needle decomposition of unit square.

So, what we proved it the following result.

**Theorem 2.2.5.** ***(Needle decomposition I)*** *Let $K \subseteq R^n$ be a convex, open, bounded set. Let $\mu$ be an absolutely continuous measure on $K$, let $f : K \to \mathbb{R}$ be a continuous and $\mu$-integrable function with $\int_K f \mathrm{d}\mu = 0$. Let $\delta > 0$. Then it exists $N$ and there exist convex bodies $K_1, \ldots, K_{2^N} \subseteq \mathbb{R}^n$ such that*

1. *$K = K_1 \cup \cdots \cup K_{2^N}$ is a disjoint union up to measure 0,*
2. *for all $i$, $K_i$ is a $\delta$-needle,*
3. *for all $i$, $\int_{K_i} f \mathrm{d}\mu = 0$.*

**Lemma 2.2.6.** *It suffices to prove Theorem 2.0.1 in 1 dimension.*

*Proof.* Given $K, \mu, f$ as above. Let $\varepsilon > 0$. It suffices to prove

$$\int_K f^2 \mathrm{d}\mu \leq \frac{\mathrm{diam}^2(K)}{\pi^2} \int_K |\nabla f|^2 \mathrm{d}\mu + \varepsilon \,.$$

By uniform continuity of $f^2$ and $|\nabla f|^2$ on $\bar{K}$, there is a $\delta > 0$ such that for all $x, y \in K$, if $|x-y| < \delta$, then

$$\begin{aligned} \left|f(x) - f(y)\right| &< \frac{\varepsilon}{4}\,, \\ \left|f^2(x) - f^2(y)\right| &< \frac{\varepsilon}{4}\,, \\ \left||\nabla f(x)|^2 - |\nabla f(y)|^2\right| &< \frac{\varepsilon \pi^2}{4\mathrm{diam}^2(K)}\,. \end{aligned}$$

We use needle decomposition $\delta$ and we obtain $K = K_1 \cup \cdots \cup K_{2^N}$, $K_i$ all $\delta$-needles and all satisfy $\int_{K_i} f \mathrm{d}\mu = 0$. So, it is enough to prove

$$\int_{K_i} f^2 \mathrm{d}\mu \leq \frac{\mathrm{diam}^2(K_i)}{\pi^2} \int_{K_i} |\nabla f|^2 \mathrm{d}\mu + \varepsilon\,,$$

and sum over $i$.
Let $X$ be a random vector with law $\mu_i$, log-concave. Set $X_i = X\big|_{K_i}$, which is log-concave by Corollary 1.7.9. We need to show

$$\mathbb{E}[f^2(X_i)] \leq \frac{\mathrm{diam}^2(K_i)}{\pi^2} \mathbb{E}[|\nabla f(X_i)|^2] + \varepsilon\,. \tag{2.2.2}$$

We know that $K_i \subseteq \ell_i + \delta B^n$, where $\ell_i$ is a line. Set $Y_i = \mathrm{Proj}_{\ell_i}(X_i)$, which is log-concave by Exercise 1.7.5. Also, $\mathbb{P}(|X_i - Y_i| \leq \delta) = 1$. In order to prove (2.2.2), by the choice of $\delta$, it is enough to prove

$$\mathbb{E}[f^2(Y_i)] \leq \frac{\mathrm{diam}^2(K_i)}{\pi^2} \mathbb{E}[|\nabla f(Y_i)|^2] + \varepsilon\,.$$

Hence, it is enough to show

$$\mathrm{Var}(f(Y_i)) \leq \frac{\mathrm{diam}(\mathrm{supp}(Y_i))^2}{\pi^2} \mathbb{E}\left[|\partial_\theta f|^2(Y_i)\right]\,,$$

where $\theta$ is the direction of $f$. □

**Remark 2.2.7.** There is no dependence on dimension in this Poincaré inequality.

**Remark 2.2.8.** So far, we get $C\mathrm{diam}^2(K)$, where $C \leq 1$. We need to improve it to $\frac{1}{\pi^2}$. This will require us the use of ODE's and spectral theorems.

## 2.3 Reverse Hölder inequality

Let us see another application of the needle decomposition, obtained by Bourgain [4].

**Theorem 2.3.1.** ***(Reverse Hölder inequality for polynomials)*** *Let $X$ be a log-concave random vector in $\mathbb{R}^n$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree at most $d$ (all monomials of degree $\leq d$). Let $0 < p \leq q$. Then*

$$\mathbb{E}\left[|f(X)|^p\right]^{\frac{1}{p}} =: ||f(X)||_p \overset{(*)}{\leq} ||f(X)||_q \leq C||f(X)||_p\,, \tag{2.3.1}$$

*where $C > 0$ does not depend on $X$, $C = C_{p,q,d} \leq (10d(q+1))^d$.*

**Remark 2.3.2.** The inequality $(*)$ is standard Hölder inequality.

**Remark 2.3.3.** For now, the best known constant is $C_{p,q,d} = \left(C\frac{q}{p}\right)^d$.

**Example 2.3.4.** If $f$ is a quadratic polynomial, then

$$\mathbb{E}[|f(X)|^4] \leq C\big(\mathbb{E}[|f(X)|]^4\big)\,,$$

for $C > 0$ universal constant.

**Lemma 2.3.5.** *It suffices to prove Theorem 2.3.1 in dimension* 1.

*Proof.* By approximation, we may assume that $X$ is compact supported. Set

$$g(x) := |f(x)|^p - \underbrace{\int_{\mathbb{R}^n} |f(x)|^p \mathrm{d}\mu(x)}_{=:\lambda^p}\,,$$

where $\mu$ is the log-concave law of $X$. Clearly $\int_{\mathbb{R}^n} g(x)\mathrm{d}\mu(x) = 0$. Let $\varepsilon > 0$. We need to prove

$$\int_{\mathbb{R}^n} |f|^2 \mathrm{d}\mu \leq C_{p,q,d}^q \lambda^q (1+\varepsilon)\,.$$

We apply the needle decomposition for some $\delta > 0$ obtained by $\varepsilon$ and uniform continuity of $|f|^p$ and $|f|^q$. We get a decomposition $K = \bigcup_{i=1}^{2^n} K_i$ of $\delta$-needles with $\int_{K_i} g\mathrm{d}\mu = 0$.

Let $X_i = X\big|_{K_i}$. Then $\mathbb{E}[g(X_i)] = 0$, i.e. $\mathbb{E}[|f(X_i)|^p]^{\frac{1}{p}} = \lambda$ for all $i$. We know $K_i \subseteq \ell_i + \delta B^n$. We set $Y_i = \mathrm{Proj}_{\ell_i}(X_i)$, which is log-concave and one-dimensional. We know

$$\mathbb{E}[|f(Y_i)|^p] \leq \mathbb{E}[|f(X_i)|^p]\left(1+\frac{\varepsilon}{10}\right)^q\,,$$
$$\mathbb{E}[|f(Y_i)|^q p] \geq \mathbb{E}[|f(X_i)|^p]\left(1-\frac{\varepsilon}{10}\right)^q\,.$$

Assuming Theorem 2.3.1 in dimension 1, we have

$$\mathbb{E}[|f(Y_i)|^q]^{\frac{1}{q}} \leq C_{p,q,d}\mathbb{E}[|f(Y_i)|^p]^{\frac{1}{p}}$$

and since the error is multiplicative

$$\mathbb{E}[|f(X_i)|^q]^{\frac{1}{q}} \leq C_{p,q,d}\mathbb{E}[|f(X_i)|^p]^{\frac{1}{p}}\frac{1+\frac{\varepsilon}{10}}{1-\frac{\varepsilon}{10}}\,.$$

So for any $i$

$$\frac{1}{\mu(K_i)}\int_{K_i} |f|^2\mathrm{d}\mu \leq (1+\varepsilon)^q C_{p,q,d}^q \lambda^q\,,$$

and summing over $i$ we get

$$\begin{aligned}\int_K |f|^q \mathrm{d}\mu &= \sum_{i=1}^{2^n}\int_{K_i}|f|^q\mathrm{d}\mu \\ &\leq \sum_{i=1}^{2^n}\mu(K_i)(1+\varepsilon)^q C_{p,q,d}^q\lambda^q \\ &= (1+\varepsilon)^q C_{p,q,d}^q\left(\int_K |f|^p\mathrm{d}\mu\right)^{\frac{q}{p}}\,.\end{aligned}$$

Let $\varepsilon \to 0$ and we conclude. □

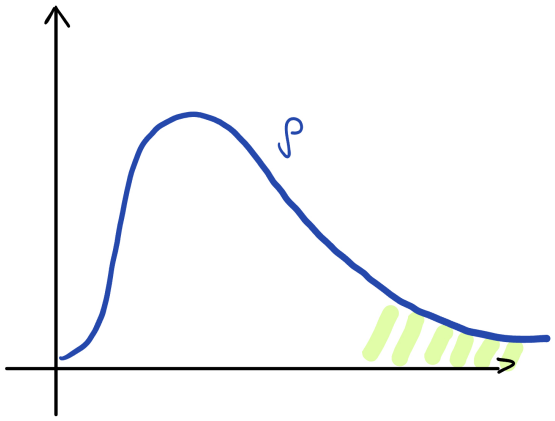



Figure 2.12: $\rho$ decays with a "light tail".

**Where does the reverse Hölder inequality come from?** Let $X$ have density $\rho$ in $[0,\infty)$. Suppose you know that $\mathbb{E}[X^{10}]^{\frac{1}{10}} \approx \mathbb{E}[X] = 1$. The RHS of (2.3.1) says that $\rho$ has a "light tail" (see Figure 2.12).

The next proposition shows the equivalency between the exponential tail and the linear growth of moments.

**Proposition 2.3.6.** *Let $X$ be a random vector with median 1. Let $A \geq 1$.*

1. *Assume that $\mathbb{P}(|X| \geq t) \leq 2e^{-\frac{t}{A}}$ for all $t > 0$. Then*

$$\mathbb{E}[||X||_p] \leq 5A(p+1) \quad \forall p > 0\,.$$

2. *Assume $\mathbb{E}[||X||_p] \leq Ap$. Then for all $t > 0$*

$$\mathbb{P}(|X| \geq t) \leq 2e^{-\frac{t}{A}}\,,$$

*where $C > 0$ is an universal constant.*

**Example 2.3.7.** In dimension 1, log-concave distributions are rather simple. Suppose we know $\rho(0) = 1$ and $\rho(1) = e^{-1}$. Then $\rho(t) \leq e^{-t}$. In fact, by (1.7.1) we know $1 = \frac{1}{t} \cdot t + \left(1 - \frac{1}{t}\right) \cdot 0$ and therefore

$$\frac{1}{e} = \rho(1) \geq \rho(t)^{\frac{1}{t}} \cdot \rho(0)^{1-\frac{1}{t}} = \rho(t)^{\frac{1}{t}} \quad \Rightarrow \rho(t) \leq e^{-t}\,.$$

# Bibliography


[1] Grégoire Allaire. *A la recherche de l'inégalité perdue*. 2012. URL: https://hal.science/hal-01111806/document.

[2] S. Artstein-Avidan, A. Giannopoulos, and V.D. Milman. *Asymptotic Geometric Analysis, Part I*. Mathematical Surveys and Monographs. 2015.

[3] Serguei G. Bobkov and Christian Houdré. *Some connections between isoperimetric and Sobolev-type inequalities*. Memoirs of the American Mathematical Society, 1997.

[4] J. Bourgain. "On the distribution of polynomials on high dimensional convex sets". In: *Geometric Aspects of Functional Analysis*. Ed. by Joram Lindenstrauss and Vitali D. Milman. Springer Berlin Heidelberg, 1991, pp. 127–137.

[5] Lawrence C Evans. *Partial Differential Equations*. American Mathematical Society, 2010.

[6] M. Gromov and V. D. Milman. "Generalization of the spherical isoperimetric inequality to uniformly convex Banach spaces". In: *Compositio Mathematica* 62.3 (1987), pp. 263–282.

[7] Mikhail Gromov. *Metric Structures for Riemannian and Non-Riemannian Spaces*. Springer, 2007.

[8] Michel Ledoux. "Isoperimetry and Gaussian analysis". In: *Geometric and Functional Analysis* 7.2 (1996), pp. 455–481.

[9] L. Lovasz, R. Kannan, and M. Simonovits. "Isoperimetric Problems for Convex Bodies and a Localization Lemma". In: *Discrete & computational geometry* 13.3-4 (1995), pp. 541–560.

[10] F. Nazarov, M. Sodin, and A. Volberg. *The geometric Kannan-Lovasz-Simonovits lemma, dimension-free estimates for volumes of sublevel sets of polynomials, and distribution of zeroes of random analytic functions*. 2001. arXiv: math/0108212 [math.CA].

[11] L. E. Payne and H. F. Weinberger. "An optimal Poincaré inequality for convex domains". In: *Archive for rational mechanics and analysis* 5.1 (1960), pp. 286–292.

[12] Henri Poincaré. "Sur les équations aux dérivées partielles de la physique mathématique". In: *Bulletin de la Société Mathématique de France* 18 (1890), pp. 93–96.

# Index


### Symbols

$\delta$-needle 23
*log*-concave 19

### A

accumulation point 11
affine Grassmannian 24

### B

Brunn-Minkowski inequality 17

### C

Cheeger constant 10
Cheeger's inequality 10
Co-area inequality 11

### E

elementary set 18

### F

Fourier law 8

### H

heat equation 7
    Neumann boundary condition 7
heat operator 9
hyperplane bisection 21

### I

isolated point 11
Isoperimetric inequality 17

### L

Lipschitz on balls 11

### M

metric-measure space 5
Minkowski sum 17

### N

Needle decomposition 23

### P

Poincaré inequality 6
    $L^1 - L^1$ 13
    $L^p - L^p$ 13
    Payne-Weinberger constant 21

### R

Rayleigh quotient 8
Reverse Hölder inequality for polynomials 27

### S

Spectral theorem 8
surface area 11