Isoperimetric constants of metric probability spaces II

Seminar on Functional Analysis and Operator Algebras

December 11, 2020

Cheeger constant

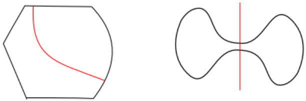
- Let μ be a Borel probability measure on $\mathbb{R}^n$. For every Borel subset A of $\mathbb{R}^n$, the Minkowski content of A with respect to μ is defined as

$$\mu^+(A) = \liminf_{t\to 0^+} \frac{\mu(A_t) - \mu(A)}{t}$$

where $A_t = \{x : d(X, A) < t\}$.

- The isoperimetric ratio of A is defined as follows:

$$\chi_\mu(A) := \frac{\mu^+(A)}{\min\{\mu(A), 1-\mu(A)\}}.$$



- Then, we define the Cheeger constant χ_μ of μ setting

$$\chi_\mu := \inf\{\chi_\mu(A) : A \ \text{Borel} \ \subset \mathbb{R}^n\}.$$

Poincaré constant and Cheeger constant

- We say that a Borel probability measure μ on $\mathbb{R}^n$ satisfies the Poincaré inequality with constant $\vartheta > 0$ if

$$\mathrm{Var}_\mu(f) \leqslant \vartheta^2 \int |\nabla f|^2 \, d\mu,$$

for all smooth functions f on $\mathbb{R}^n$, where

$$\mathrm{Var}_\mu(g) = \mathbb{E}_\mu(g^2) - (\mathbb{E}_\mu(g))^2$$

is the variance of g with respect to μ.

- The Poincaré constant ϑ_μ of μ is the smallest constant $\vartheta > 0$ for which the Poincaré inequality is satisfied for all f.

Theorem (Maz'ya, Cheeger)

Let μ be a Borel probability measure. Then, its Poincaré constant ϑ_μ satisfies

$$\vartheta_\mu \leqslant \frac{2}{\chi_\mu}.$$

Poincaré constant and Cheeger constant

- A Borel probability measure μ on $\mathbb{R}^n$ is called log-concave if for all compact subsets A, B of $\mathbb{R}^n$ and all $0 < \lambda < 1$ we have

$$\mu((1-\lambda)A + \lambda B) \geqslant \mu(A)^{1-\lambda}\mu(B)^{\lambda}.$$

Theorem (Buser, Ledoux)

Let μ be a log-concave probability measure on $\mathbb{R}^n$. Then, its Poincaré constant ϑ_μ satisfies

$$\frac{1}{2}\,\vartheta_\mu \leqslant \frac{1}{\chi_\mu} \leqslant c\,\vartheta_\mu.$$

Isotropic measures

- We say that a Borel probability measure μ on $\mathbb{R}^n$ is isotropic if $\mathrm{bar}(\mu) = \int_{\mathbb{R}^n} x d\mu(x) = 0$ and μ satisfies the isotropic condition

$$\int_{\mathbb{R}^n} \langle x, \theta \rangle^2 \, d\mu(x) = 1, \qquad \theta \in S^{n-1}.$$

- Similarly, we shall say that a log-concave function $f : \mathbb{R}^n \to [0, \infty)$ with barycenter $\mathrm{bar}(f) = 0$ is isotropic if $\int f(x)dx = 1$ and the measure $d\mu(x) = f(x)dx$ is isotropic.
- A convex body K of volume 1 in $\mathbb{R}^n$ with barycenter at the origin is called isotropic if

$$\int_K \langle x, \theta \rangle^2 \, dx = L_K^2$$

for some constant $L_K > 0$ (the isotropic constant of K) and all $\theta \in S^{n-1}$.
- One can check that K is isotropic if and only if the function $f_K := L_K^n \mathbf{1}_{\frac{1}{L_K} K}$ is an isotropic log-concave function.

- Let f be a log-concave function with finite, positive integral. The covariance matrix $\mathrm{Cov}(f)$ is the matrix with entries

$$[\mathrm{Cov}(f)]_{ij} := \frac{\int_{\mathbb{R}^n} x_i x_j f(x)\, dx}{\int_{\mathbb{R}^n} f(x)\, dx} - \frac{\int_{\mathbb{R}^n} x_i f(x)\, dx}{\int_{\mathbb{R}^n} f(x)\, dx} \frac{\int_{\mathbb{R}^n} x_j f(x)\, dx}{\int_{\mathbb{R}^n} f(x)\, dx}.$$

- If f is the density of a measure μ we denote this matrix also by $\mathrm{Cov}(\mu)$.
- The isotropic constant of f is defined by

$$L_f := \left(\frac{\sup_{x\in\mathbb{R}^n} f(x)}{\int_{\mathbb{R}^n} f(x)dx}\right)^{\frac{1}{n}} [\det \mathrm{Cov}(f)]^{\frac{1}{2n}}.$$

(and given a log-concave measure μ with density f_μ we let $L_\mu := L_{f_\mu}$).

- It is easy to check that the isotropic constant L_μ is an affine invariant.

- Recall that a Borel probability measure μ on $\mathbb{R}^n$ is isotropic if $\mathrm{bar}(\mu) = \int_{\mathbb{R}^n} x d\mu(x) = 0$ and μ satisfies the isotropic condition

$$\int_{\mathbb{R}^n} \langle x, \theta \rangle^2 \, d\mu(x) = 1, \qquad \theta \in S^{n-1}.$$

- This implies that $\mathbb{E}_\mu(\|x\|_2^2) = n$.
- We also have $\mathrm{Cov}(\mu) = I_n$ is the identity and $L_\mu = \|f_\mu\|_\infty^{1/n}$ where f_μ is the density of μ.
- We define $L_n = \sup\{L_\mu : \mu$ is an isotropic log-concave probability measure on $\mathbb{R}^n\}$.
- The isotropic constant conjecture asks whether $L_n \leqslant C$ for an absolute constant $C > 0$.

KLS-Conjecture

For every isotropic log-concave probability measure μ on $\mathbb{R}^n$ one has $\chi_\mu \geqslant c$, where $c > 0$ is an absolute constant.

$$\chi_n := \inf\{\chi_\mu : \mu \text{ is an isotropic log-concave probability measure on } \mathbb{R}^n\} \geqslant c.$$

Theorem (Eldan-Klartag)

$$L_n \leqslant C/\chi_n.$$

- In other words, the KLS-conjecture is stronger than the isotropic constant conjecture, and any lower bound for χ_n provides an upper bound for L_n.

KLS-conjecture: bounds for χ_n

Theorem (Lee-Vempala)

If μ is a log-concave probability measure on $\mathbb{R}^n$ with covariance matrix A then

$$\chi_\mu \geqslant \frac{c}{\left(\mathrm{tr}(A^2)\right)^{1/4}}$$

where $c > 0$ is an absolute constant.

- If we make the additional assumption that μ is isotropic then we obtain the lower bound

$$\chi_\mu \geqslant c/\sqrt[4]{n}.$$

Theorem (announced by Y. Chen, 30/11/2020)

If μ is a log-concave probability measure on $\mathbb{R}^n$ then, for any integer $\ell \geqslant 1$,

$$\chi_\mu \geqslant \frac{1}{(c\ell(\log n + 1))^{\ell/2} n^{4/\ell} \sqrt{\varrho(\mu)}}$$

where $\varrho(\mu)$ is the spectral norm of $\mathrm{Cov}(\mu)$.

- Choosing $\ell = \sqrt{\frac{\log n}{\log\log n}}$ we get:

Theorem (announced by Y. Chen, 30/11/2020)

If μ is a log-concave probability measure on $\mathbb{R}^n$ then, for any integer $\ell \geqslant 1$,

$$\chi_\mu \geqslant \frac{1}{n^{c\sqrt{\frac{\log\log n}{\log n}}}\sqrt{\varrho(\mu)}}$$

where $\varrho(\mu)$ is the spectral norm of $\mathrm{Cov}(\mu)$.

- Since $\frac{\log\log n}{\log n} \to 0$, this implies that for any $\epsilon > 0$ and any isotropic log-concave probability measure μ on $\mathbb{R}^n$ we have

$$\chi_\mu \geqslant c/n^\epsilon.$$

- Therefore, $\chi_n \geqslant c/n^\epsilon$ and $L_n \leqslant Cn^\epsilon$ by the theorem of Eldan and Klartag.

- It is a result of E. Milman.

$$I_\mu(t) = \inf\{\mu^+(A) : A \text{ Borel},\ \mu(A) = t\}.$$

- Note that

$$\chi_\mu := \inf_{0<t<1} \frac{I_\mu(t)}{\min\{t, 1-t\}} = \inf_{0<t\leqslant 1/2} \frac{\min\{I_\mu(t), I_\mu(1-t)\}}{t}.$$

Theorem (E. Milman)

Let μ be a log-concave probability measure on $\mathbb{R}^n$. Then, the isoperimetric profile I_μ of μ is concave on $(0,1)$, and for every $t \in (0,1)$ we have $I_\mu(t) = I_\mu(1-t)$. As a consequence,

$$\chi_\mu = \inf_{0<t\leqslant 1/2} \frac{I_\mu(t)}{t} = 2I_\mu(1/2).$$

- This means that we can calculate the Cheeger constant of a log-concave probability measure μ by looking only at Borel sets A with $\mu(A) = 1/2$.

Eldan's stochastic localization

- Let μ be a log-concave probability measure on $\mathbb{R}^n$ and let $A := \mathrm{Cov}(\mu)$.
- We consider the stochastic differential equation

$$du_t = A_t^{-1/2} dW_t + A^{-1} a_t dt, \qquad u_0 = 0$$
$$dB_t = A^{-1} dt, \qquad B_0 = 0$$

where W_t is the Wiener process, and the density f_t, the mean a_t and the covariance matrix A_t of the probability measure μ_t are defined by

$$f_t(x) = \frac{e^{-\langle u_t, x\rangle - \frac{1}{2}\langle x, B_t x\rangle} f(x)}{\int_{\mathbb{R}^n} e^{-\langle u_t, y\rangle - \frac{1}{2}\langle y, B_t y\rangle} f(y) dy}, \quad a_t = \mathbb{E}_{\mu_t}(x), \quad A_t = \mathbb{E}_{\mu_t}((x - a_t) \otimes (x - a_t)).$$

- Eldan: "In some sense, the above is just the continuous version of the following iterative process: at every time step, normalize the measure to be isotropic, and multiply it by a linear function, equal to 1 at the origin, whose gradient has a random direction."

A rough description of what it is

- We fix a subset E of the original space with measure one half according to the original log-concave distribution (it suffices to consider such subsets to bound the Cheeger constant).
- Stochastic localization can be viewed as the continuous time version of a discrete process, where at each step, we pick a random direction and multiply the current density with a linear function along the chosen direction.
- Over time, the density can be viewed as a Gaussian density multiplied by a log-concave function, with the Gaussian gradually reducing in variance.
- When the Gaussian becomes sufficiently small in variance, then the overall distribution has good Cheeger constant, determined by the inverse of the Gaussian standard deviation.
- An important property of the infinitesimal change at each step is balance: the expected measure of any subset is the same as the original measure. Note that the measure of a set E is a random quantity that deviates from its original value of $\frac{1}{2}$ over time.
- The main question then is: what direction to use at each step so that (a) the measure of E remains bounded and (b) the Gaussian part of the density has small variance. What happens is that the simplest choice, namely a pure random direction chosen from the uniform distribution suffices.

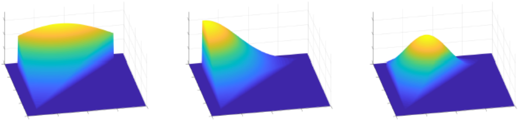
Eldan's stochastic localization

- The discrete time equivalent would be

$$f_{t+1}(x) = f_t(x)(1 + \sqrt{h}\langle x - a_t, w\rangle)$$

  for a sufficiently small $h > 0$ and a random Gaussian vector w in $\mathbb{R}^n$.
- Using the approximation $1 + y \sim e^{y - \frac{1}{2}y^2}$, we see that over time this process introduces a negative quadratic factor in the exponent, which will be the Gaussian factor.
- As time tends to ∞, the distribution tends to a more and more concentrated Gaussian and eventually a delta function, at which point any subset has measure either 0 or 1.
- The idea of the proof is to stop at a time that is large enough to have a strong Gaussian factor in the density, but small enough to ensure that the measure of a set is not changed by more than a constant.

Eldan's stochastic localization



Theorem

If f has compact support and its covariance matrix is invertible then the equation is well-defined and has a unique solution on the time interval $[0, T]$ for every $T > 0$. Moreover, for any $x \in \mathbb{R}^n$, $f_t(x)$ is a martingale with

$$df_t(x) = \langle x - a_t, A^{-1/2} dW_t \rangle \, f_t(x).$$

More log-concave than Gaussian

- A density $h : \mathbb{R}^n \to \mathbb{R}$ is called *more log-concave than Gaussian* if $h = \varphi \cdot f$ where φ is a Gaussian density and f is an integrable log-concave function.

Theorem (Cousins-Vempala)

Let μ be a log-concave probability measure on $\mathbb{R}^n$ with density proportional to $h(x) = e^{-\langle x, Bx\rangle} f(x)$, where $f : \mathbb{R}^n \to \mathbb{R}^+$ is an integrable log-concave function. For every $E \subset \mathbb{R}^n$ we have that

$$\mu^+(E) \geqslant \frac{1}{2}\|B\|_2^{-1/2} \min\{\mu(E), \mu(E^c)\}$$

for every $E \subset \mathbb{R}^n$, and hence $\chi_\mu \geqslant \frac{1}{2}\|B\|_2^{-1/2}$.

Bound for χ_μ via μ_t

- Let μ have compact support and let E be a subset of $\mathbb{R}^n$ with $\mu(E) = \frac{1}{2}$.
- Using the martingale property of $\mu_t(E)$ we see that

$$\begin{aligned}\mu^+(E) &= \mathbb{E}\left(\mu_t^+(E)\right) \\ &\geqslant \frac{1}{2}\mathbb{E}\big(\|B_t^{-1}\|_2^{-1/2}\min\{\mu_t(E),\mu_t(E^c)\}\big) \\ &\geqslant \frac{1}{2}\|B_t^{-1}\|_2^{-1/2}\cdot\frac{1}{4}\mathbb{P}\Big(\frac{1}{4}\leqslant\mu_t(E)\leqslant\frac{3}{4}\Big) \\ &= \frac{1}{4}\|B_t^{-1}\|_2^{-1/2}\mathbb{P}\Big(\frac{1}{4}\leqslant\mu_t(E)\leqslant\frac{3}{4}\Big)\min\{\mu(E),\mu(E^c)\}.\end{aligned}$$

- This gives a lower bound for χ_μ:

$$\chi_\mu \geqslant \frac{1}{4}\|B_t^{-1}\|_2^{-1/2}\inf_E \mathbb{P}\Big(\frac{1}{4}\leqslant\mu_t(E)\leqslant\frac{3}{4}\Big)$$

  where the infimum is over all $E\subset\mathbb{R}^n$ with $\mu(E)=\frac{1}{2}$.
- Note that $B_t = tA^{-1}$ and hence $\|B_t^{-1}\|_2^{-1/2} = \sqrt{t}\|A\|_2^{-1/2}$.

Two things to do

- We need to show that there exists $t > 0$ such that the Gaussian component $e^{-\langle x, B_t x\rangle}$ of f_t is large enough: this means to have a lower bound for $\|B_t^{-1}\|_2^{-1/2}$. We know that
$$\|B_t^{-1}\|_2^{-1/2} = \sqrt{t}\|A\|_2^{-1/2}.$$
- We need to study $\mu_t(E)$ in order to give a lower bound for
$$\mathbb{P}\Big(\frac{1}{4} \leqslant \mu_t(E) \leqslant \frac{3}{4}\Big)$$
for any $E \subset \mathbb{R}^n$ with $\mu(E) = \frac{1}{2}$.

Control on $\mu_t(E)$

Lee-Vempala

For any $E \subset \mathbb{R}^n$ with $\mu(E) = \frac{1}{2}$ and any $t \geqslant 0$ we have

$$\mathbb{P}\Big(1/4 \leqslant \mu_t(E) \leqslant 3/4\Big) \geqslant \frac{9}{10} - \mathbb{P}\Big(\int_0^t \|A^{-1/2}A_sA^{-1/2}\|_2 ds \geqslant 1/64\Big).$$

- We define
$$\chi'_n = \inf_{\mu} \chi_\mu \sqrt{\varrho(\mu)},$$
where the infimum is over all log-concave probability measures with compact support and $\varrho(\mu) = \|\mathrm{Cov}(\mu)\|_2$.
- At this point, Chen has a new bound:

Theorem (Chen)

Assume that there exist $0 < \beta \leqslant \frac{1}{2}$ and $\alpha \geqslant 1$ such that $\chi'_k \geqslant \frac{1}{\alpha k^\beta}$ for all $k \leqslant n$. Then, if we set $q = \lceil \frac{1}{\beta} \rceil + 1$ and $T_2 = \frac{1}{cq\alpha^2(\log n)n^{2\beta-\beta/q}}$ we have that

$$\mathbb{P}\Big(\int_0^{T_2} \|A^{-1/2}A_sA^{-1/2}\|_2 ds \geqslant 1/64\Big) < \frac{1}{4}.$$

Theorem (Chen)

Assume that there exist $0 < \beta \leqslant \frac{1}{2}$ *and* $\alpha \geqslant 1$ *such that* $\chi_k' \geqslant \frac{1}{\alpha k^\beta}$ *for all* $k \leqslant n$. *Then, if we set* $q = \lceil \frac{1}{\beta} \rceil + 1$ *and* $T_2 = \frac{1}{cq\alpha^2 (\log n) n^{2\beta - \beta/q}}$ *we have that*

$$\mathbb{P}\Big(\int_0^{T_2} \|A^{-1/2} A_s A^{-1/2}\|_2 ds \geqslant 1/64 \Big) < \frac{1}{4}.$$

Theorem

Assume that there exist $0 < \beta \leqslant \frac{1}{2}$ *and* $\alpha \geqslant 1$ *such that* $\chi_k' \geqslant \frac{1}{\alpha k^\beta}$ *for all* $k \leqslant n$. *Then, if we set* $q = \lceil \frac{1}{\beta} \rceil + 1$ *we have that*

$$\chi_n' \geqslant \frac{1}{c\sqrt{q}\alpha\sqrt{\log n}\, n^{\beta - \beta/(2q)}}.$$

Using the new bound

Theorem

Assume that there exist $0 < \beta \leqslant \frac{1}{2}$ and $\alpha \geqslant 1$ such that $\chi_k' \geqslant \frac{1}{\alpha k^{\beta}}$ for all $k \leqslant n$. Then, if we set $q = \lceil \frac{1}{\beta} \rceil + 1$ we have that

$$\chi_n' \geqslant \frac{1}{c\sqrt{q}\alpha\sqrt{\log n}n^{\beta-\beta/(2q)}}.$$

- Let μ have compact support and let E be a subset of $\mathbb{R}^n$ with $\mu(E) = \frac{1}{2}$.
- Using the martingale property of $\mu_t(E)$ and the fact that $B_t = tA^{-1}$ we see that

$$\begin{aligned}\mu^+(E) = \mathbb{E}\left(\mu_{T_2}^+(E)\right) &\geqslant \frac{1}{2}\mathbb{E}\big(\|B_{T_2}^{-1}\|_2^{-1/2}\min\{\mu_{T_2}(E), \mu_{T_2}(E^c)\}\big)\\ &\geqslant \frac{1}{2}\|B_{T_2}^{-1}\|_2^{-1/2}\cdot\frac{1}{4}\mathbb{P}\Big(\frac{1}{4}\leqslant \mu_{T_2}(E)\leqslant\frac{3}{4}\Big)\\ &\geqslant \frac{1}{4}\|B_{T_2}^{-1}\|_2^{-1/2}\cdot\frac{1}{4}\\ &= \frac{1}{8}\sqrt{T_2}\|A\|_2^{-1/2}\min\{\mu(E), \mu(E^c)\}.\end{aligned}$$

- This gives a lower bound for χ_μ: $\chi_\mu \geqslant \frac{1}{8}\sqrt{T_2}$ in the isotropic case, because then $A = I_n$ is the identity.

- It is now a calculus matter to show:

Theorem (Chen)

There exists an absolute constant $c > 0$ such that, for any log-concave probability measure μ with compact support on $\mathbb{R}^n$ and any integer $\ell \geqslant 1$,

$$\chi_\mu \geqslant \frac{1}{(c\ell(\log n + 1))^{\ell/2} n^{4/\ell}\sqrt{\varrho(\mu)}}.$$

- We apply the previous result recursively. First we set

$$\alpha_1 = 4 \quad \text{and} \quad \beta_1 = \frac{1}{2}.$$

- For every $\ell \geqslant 1$ we define

$$\alpha_{\ell+1} = 2c\alpha_\ell\beta_\ell^{-1/2} \quad \text{and} \quad \beta_{\ell+1} = \beta_\ell - \beta_\ell^2/4.$$

- Then, we have

$$\frac{1}{\ell+1} \leqslant \beta_\ell \leqslant \frac{4}{\ell} \quad \text{and} \quad \alpha_\ell \leqslant (4c^2\ell)^{\ell/2}.$$

Recursion

- We start with the simplest known lower bound (of Kannan-Lovász and Simonovits):

$$\chi_n' \geqslant \frac{1}{\alpha_1 n^{\beta_1}}.$$

- Assuming that

$$\chi_k' \geqslant \frac{1}{\alpha_\ell (\log k + 1)^{\ell/2} k^{\beta_\ell}}$$

  we get

$$\begin{aligned}
\chi_n' &\geqslant \frac{1}{c\sqrt{q}\alpha_\ell (\log n + 1)^{\ell/2} \sqrt{\log n}\, n^{\beta_\ell - \beta_\ell/(2q)}} \\
&\geqslant \frac{1}{2c\alpha_\ell \beta_\ell^{-1/2} (\log n + 1)^{(\ell+1)/2} n^{\beta_\ell - \beta_\ell^2/4}} \\
&= \frac{1}{\alpha_{\ell+1} (\log n + 1)^{(\ell+1)/2} n^{\beta_{\ell+1}}}.
\end{aligned}$$

- Let μ be an isotropic log-concave probability measure on $\mathbb{R}^n$. Since $\mathbb{E}_\mu(\|x\|_2^2) = n$, from Markov's inequality we see that

$$\mu(\{\|x\|_2 \geqslant 5\sqrt{n}\}) \leqslant \frac{1}{25}.$$

- Consider the measure ν obtained by truncating μ on the ball $B := (5\sqrt{n})B_2^n$. For every $E \subset \mathbb{R}^n$ with $\mu(E) = \frac{1}{2}$ we have that

$$\begin{aligned}\mu^+(E) &= \nu^+(E)\mu(B) \geqslant \chi_\nu \min\{\nu(E), \nu(E^c)\}\mu(B)\\ &= \chi_\nu \min\{\mu(E \cap B), \mu(B \cap E^c)\} \geqslant \chi_\nu \min\{\mu(E) - \mu(B^c), \mu(E^c) - \mu(B^c)\}\\ &\geqslant \frac{1}{2}\chi_\nu \min\{\mu(E), \mu(E^c)\},\end{aligned}$$

because $\mu(E) - \mu(B^c) = \mu(E^c) - \mu(B^c) \geqslant \frac{1}{2} - \frac{1}{25} \geqslant \frac{1}{4}$.

- Since $\varrho(\nu) \leqslant C$ (this is simple), having given a lower bound for the Cheeger constant of log-concave probavility measures with compact support, we have settled the general case (for isotropic measures, but an analogous general result can be also deduced).