A solution of Enflo's problem

Alexander Volberg joint work with Paata Ivanisvili and Ramon Van Handel

February 6, 2021

Embeddings and distortion

This is the talk about Ribe program popularized in many works of Bourgain, Naor, Lindenstrauss, Enflo, Pisier, Schechtman and many others. Ribe program: Do we really need linearity in the Banach space theory?

I will start with Bourgain's discretization theorem.

Suppose $(X, d_X), (Y, d_Y)$ are metric spaces and $D \geq 1$. We say that X embeds into Y with distortion D if there exists $f : X \to Y$ and $s > 0$ such that for all $x, y \in X$,

$$s\, d_X(x, y) \leq d_Y(f(x), f(y)) \leq s\, D\, d_X(x, y) \quad \forall x, y \in X\,.$$

The smallest D is called *embedding constant* and is denoted by $C_Y(X)$. If one of Y or X are linear normed spaces we can always think that $s = 1$, so $C_Y(X)$ is the infimum of D's such that

$$d_X(x, y) \leq d_Y(f(x), f(y)) \leq D\, d_X(x, y) \quad \forall x, y \in X\,.$$

The estimates of distortion of metric space embeddings is one of the hottest topic now at the junction of harmonic analysis and big data theory.

Discretization

Given X, Y, where X being n-dimensional linear normed space, Y Banach space, let $\delta = \delta_{X \to Y} > 0$, be the largest number such that *if for every δ-net $N_\delta(B_X)$ of the unit ball B_X there is a K-embedding of this net into Y, then there is $2K$-embedding of X into Y* and this embedding $X \to Y$ is **linear**.

May be such $\delta > 0$ does not exist for certain X, Y? It turns out it always exists, and moreover **is independent on geometry of** X and Y.

It can be chosen depending only on n. This is Bourgain's discretization theorem.

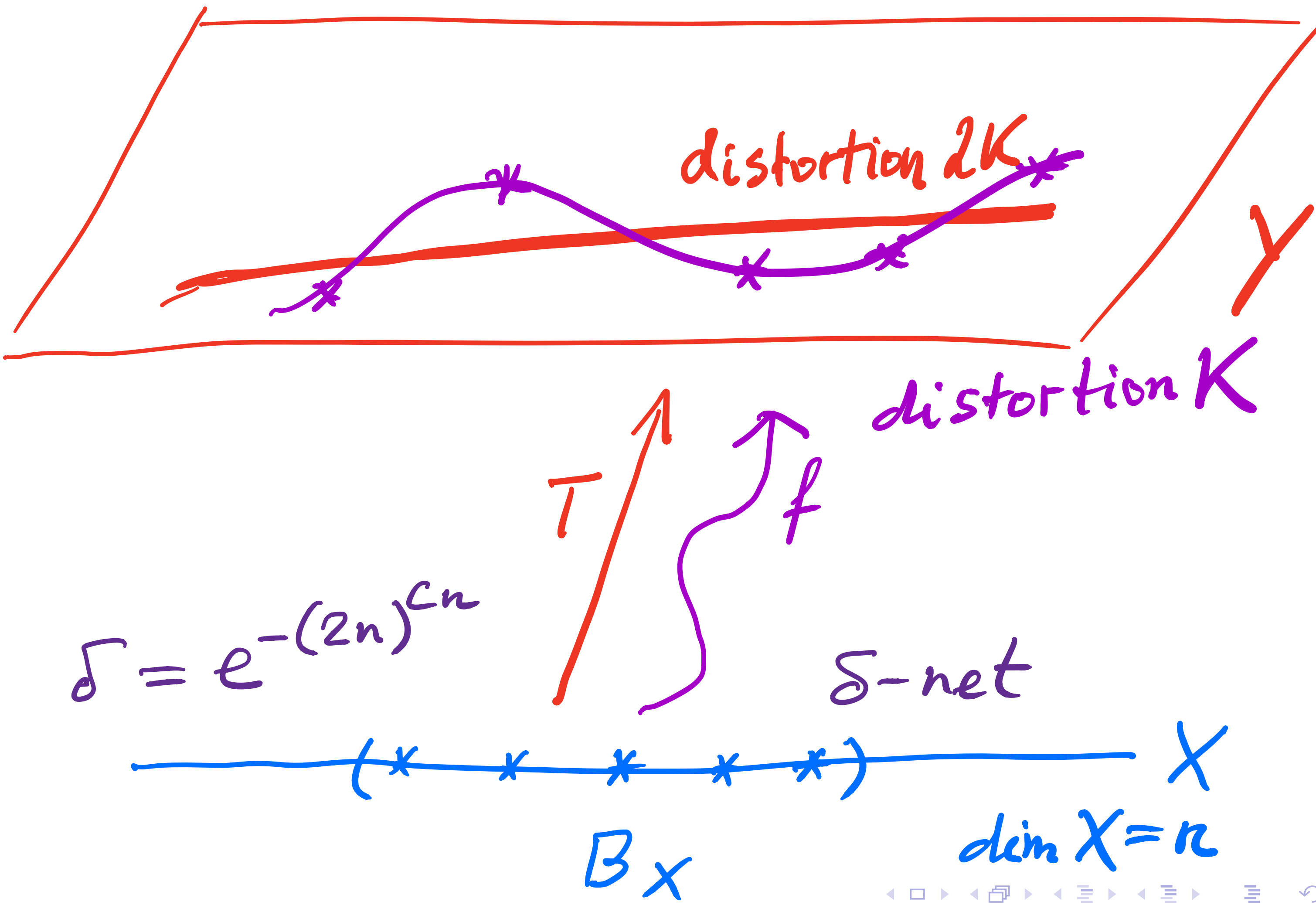
Bourgain discretization theorem

Theorem (Bourgain)

Let X be an n dimensional normed space, let Y be a Banach space. Then there exists **a linear map** $T : X \to Y$ *that realizes the following inequality*

$$C_Y(X) \leq 2 \sup_{\delta-nets} C_Y(N_\delta(B_X)), \text{ as soon as } \delta = e^{-(2n)^{Cn}}.$$

Figure of linearizability" á la Bourgain



Corollaries of Bourgain's discretization theorem

Corollary

Let X be a finite dimensional normed space, let Y be a Banach space. Let $f : X \to Y$ is a bi-Lipschitz map (not necessarily linear) with distortion D. Then there exists a linear embedding $T : X \to Y$ with distortion at most $2D$ (in fact even $(1+\varepsilon)D$).

Proof.

Let $n = \dim X < \infty$. Choose any $e^{-(2n)^{Cn}}$-net N in B_X. Map f embeds it with distortion D independent of the net. Use Bourgain's discretization theorem. Get linear embedding T with distortion $\leq 2D$. $\square$

What may happen for infinite dimension n?

Famous problem: are any two bi-Lipschitz equivalent Banach spaces X, Y linearly isomorphic?

Kadec: any two separable Banach spaces are homeomorphic.

In between: two uniformly homeomorphic Banach spaces X, Y can be NOT linearly isomorphic (Johnson–Lindenstaruss–Schechtman). But if $Y = \ell^p, 1 < p < \infty$, they are isomorphic.

Normed linear spaces are uniformly homeomorphic iff there exists invertible $F : X \to Y$ not necessarily linear such that F, F^{-1} are uniformly continuous. By Corson–Klee lemma this implies: there exists $f : X \to Y$

$$\forall x, y \in X,\ \|x - y\|_X \geq 1 \Rightarrow \|x - y\|_X \leq \|f(x) - f(y)\|_Y \leq D\|x - y\|_X$$

Corollaries of Bourgain's discretization theorem

Theorem (Martin Ribe)

If two Banach spaces X, Y are uniformly homeomorphic, then there exists D such that $\forall n < \infty, \forall X_0, dim X_0 = n$, X_0 linear subspace of X, there exists linear $T : X_0 \to Y$ with distortion at most D. Symmetrically for $\forall Y_0, dim Y_0 = n$, Y_0 linear subspace of Y....

For bi-Lipschitz equivalent X, Y we saw that it follows from Bourgain discretization theorem. But for uniformly homeomorphic X, Y it also easily follows from Bourgain discretization theorem combined with Corson–Klee lemma.

Corollaries of Bourgain's discretization theorem

Corollary (Martin Ribe)

Let X, Y be two Banach spaces that are bi-Lipschitz equivalent (or just uniformly homeomorphic). Then they have the same type p.

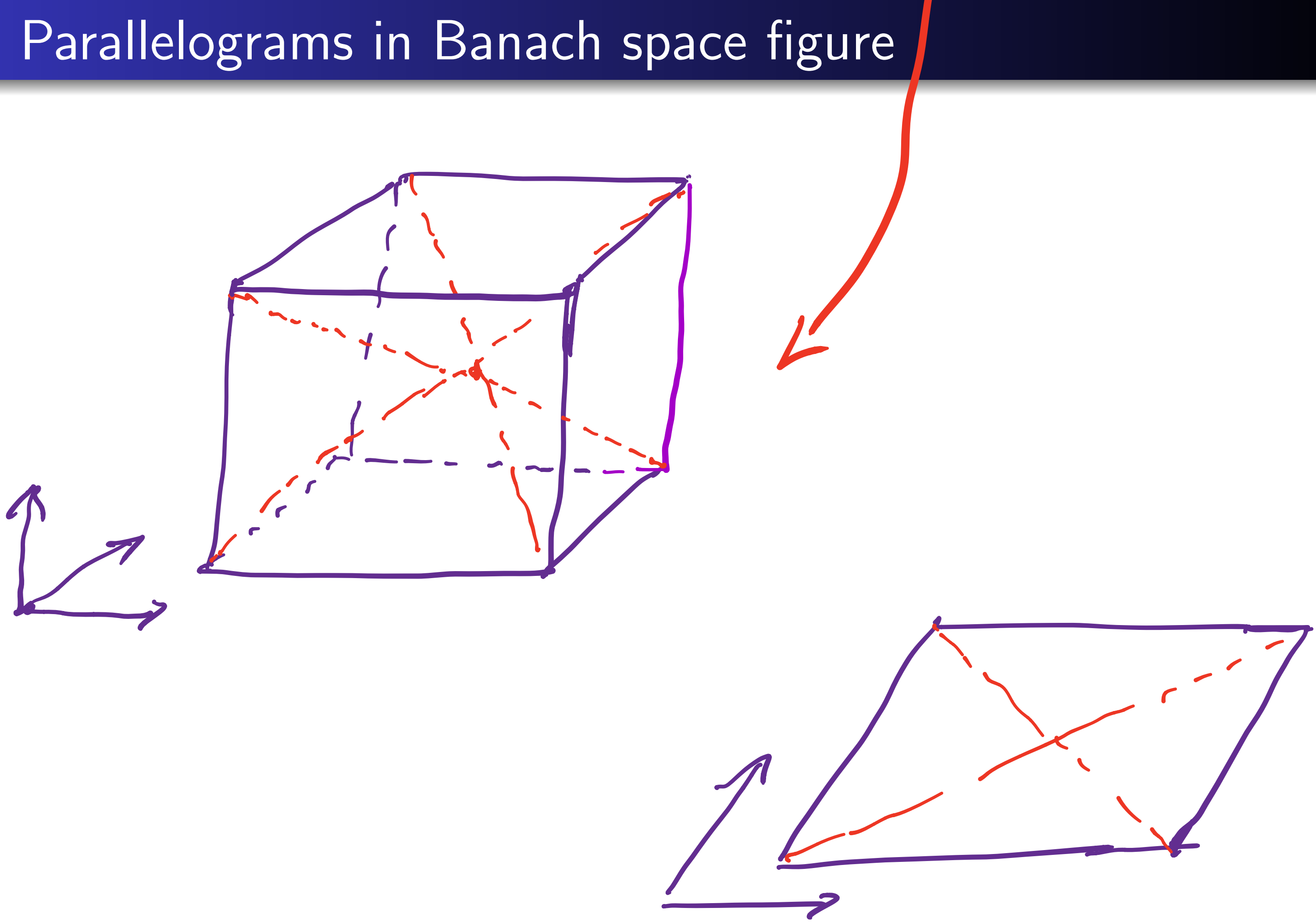
Definition

We call $p, 1 \le p \le 2$, the Rademacher type of Banach space X if the following **strengthening of triangle inequality** *holds*

$$\exists C < \infty \forall n < \infty, \forall x_1, \dots, x_n \in X, \mathbb{E}\| \sum_{i=1}^n \varepsilon_i x_i \|_X^p \le C \sum_{i=1}^n \|x_i\|_X^p .$$

It is easy to see that type can be only $p \le 2$, and that ALL Banach spaces have type 1 ($=$ triangle inequality, $C = 1$), so we say **non-trivial type** meaning $p > 1$. For example, ℓ^1 does not have non-trivial type.

Parallelograms in Banach space figure



But why this Ribe's corollary happened really?

What is the secret metric mechanism that exists and allows the type to be preserved under bi-Lipschitz maps?

The problem with Ribe's theorem is it is purely existential: it says that all local properties of infinite-dimensional Banach spaces (they are always linear properties) depend only on the metric structure of the Banach space (as they are invariant under uniform homeomorphisms), but it doesn't explain how to formulate these local notions metrically.

How to formulate local notions of Banach space metrically is important for several reasons:
(1) it makes it possible to extend these notions to more general metric spaces;
(2) it makes it possible to study embedding in Banach spaces of general metric spaces.
Bourgain initiated a program to find explicit metric descriptions of local properties of Banach spaces.

In the case of type, the natural conjecture was that Enflo's notion (which predates the Ribe program) is the right one in this context, and this is what we proved.

Local notions of p-type metrically: Enflo type

Hamming cube: It is a probability space $(\{-1,1\}^n, \mathbb{P})$, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$. We can think about ε_i's as independent r.v.having values ± 1 with probability $1/2, 1/2$. The integral $\int_{\{-1,1\}^n} f\, d\mathbb{P}$ will be denoted by $\mathbb{E}$. So $\mathbb{E} f(\varepsilon) = \frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} f(\varepsilon)$. Let $f : \{-1,1\}^n \to X$ into Banach space be given by $f(\varepsilon) = \sum_{i=1}^n \varepsilon_i x_i$, in other words, let f **be linear polynomial in** ε_i with Banach coefficients. Then type p means the existence of $C < \infty$ such that for all **X-valued linear polynomials** the following holds (ε^i changes ε in exactly only i-th place):

$$\mathbb{E}\, \|f(\varepsilon) - f(-\varepsilon)\|_X^p \le C \sum_{i=1}^n \mathbb{E} \|f(\varepsilon) - f(\varepsilon^i)\|_X^p\,.$$

There is **almost** no linearity here. That involves only metric notion of distance between couples of points: (**diagonals versus edges**)

$$\mathbb{E}\, d_X(f(\varepsilon), f(-\varepsilon))^p \le C \sum_{i=1}^n \mathbb{E} d_X(f(\varepsilon), f(\varepsilon^i))^p \quad f \text{ is linear}\,.$$

Enflo type p

Definition

Banach space X is called having Enflo type p, $1 \le p \le 2$, if

$$\mathbb{E}\,\|f(\varepsilon) - f(-\varepsilon)\|_X^p \le C \sum_{i=1}^n \mathbb{E}\|f(\varepsilon) - f(\varepsilon^i)\|_X^p$$

diagonals versus edges *inequality holds* **not** *just for linear f, but for* **all** $f : \{-1,1\}^n \to X$.

There is no linearity here **at all**. That involves only metric notion of distance between couples of points:

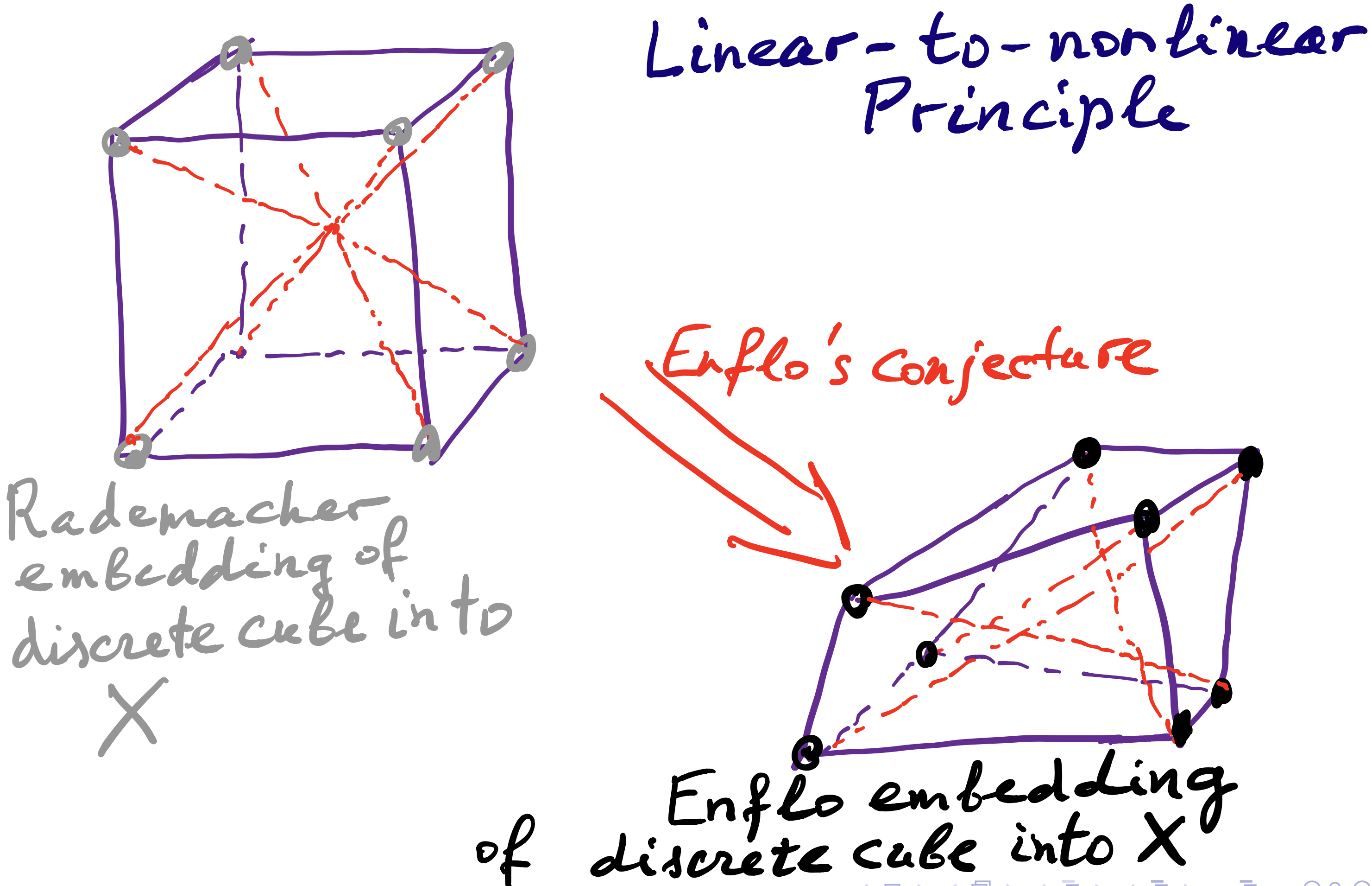
$$\mathbb{E}\,d_X(f(\varepsilon), f(-\varepsilon))^p \le C \sum_{i=1}^n \mathbb{E} d_X(f(\varepsilon), f(\varepsilon^i))^p \quad \forall f : \{-1,1\}^n \to X\,.$$

Linear implies non-linear principle. Enflo's problem: **does it**?

Geometric picture

Any "skewed" Hamming cube with vertices in Banach space X has the property that average of the p-th powers of lengths of **main diagonals** is controlled by of the p-th powers of lengths of all sides. This is the meaning of Enflo type p.
Obviously X of Enflo type $p \Rightarrow X$ is of Rademacher type p. In Enflo definition **all polynomials** $f : \{-1,1\}^n \to X$ participate, in Rademacher only **linear polynomials**.
Geometrically, in one definition "skewed" Hamming cubes with vertices in Banach space X participate, in another definition only "parallelogram" shaped Hamming cubes with vertices in Banach space X participate.
Enflo's problem: Does Rademacher type p implies Enflo type p? In other words, is it true that if
$\mathbb{E}\,\|f(\varepsilon) - f(-\varepsilon)\|_X^p \le C\sum_{i=1}^n \mathbb{E}\|f(\varepsilon) - f(\varepsilon^i)\|_X^p$ holds for **only linear polynomials with coefficients in Banach space** X, then it holds for **all polynomials with coefficients in Banach space** X independent of $\deg f$? May be with different but still finite C?

Figure: “Parallelogram”-shaped and general Hamming cubes embedded in Banach space



Yes: from linear polynomial to all polynomial

Theorem (Ivanishvili–Van Handel–Volberg, Annals of Math. (2) 192 (2020), no. 2, 665–678)

For any Banach space X, $C_p^{Enflo}(X) \leq \frac{\pi}{\sqrt{2}} C_p^{Rademacher}(X)$

To explain the method I need to recall Pisier–Poincaré inequalities.

Previous results

It was formulated by Enflo in 1978, and there were numerous partial results.

1) Enflo: if Rademacher constant is 1, then yes.
2) Bourgain–Milman–Wolfson: Rademacher type p implies Enflo type $p - \varepsilon$.
3) Schechtman–Naor, if X is a UMD space then yes.
4) Hytönen–Naor, if X is a UMD^+ space then yes.
5) Eskenazis, again certain class of Banach spaces, then yes.
6) Pisier.

Pisier–Poincaré inequalities

First Gaussian case: Let g, g' be independent gaussian $\mathbb{R}^n$-vectors. Let $f : \mathbb{R}^n \to X$ be a function with values in **an arbitrary Banach space** X. Let $1 \le p < \infty$. Then without any dependence on n we have

$$\left(\mathbb{E}\|f(g) - \mathbb{E}f(g)\|_X^p\right)^{1/p} \le \frac{\pi}{2}\left(\mathbb{E}\|\sum_{j=1}^n g_j' \frac{\partial f}{\partial x_j}(g)\|_X^p\right)^{1/p}.$$

Examples of applications: 1) Let $f : \mathbb{R}^n \to M_{d\times d}^{symm}$ is a matrix-valued function, and on $M_{d\times d}^{symm}$ we have norm $\|A\|_{\sigma_p} = (\mathrm{Tr}\,[|A|^p])^{1/p}$.
Then we have **concentration inequality for random matrices**.
Let $f : \mathbb{R}^n \to M_{d\times d}^{symm}$ is a test function, and on $M_{d\times d}^{symm}$. For large p, independently on f and d, using Pisier and also Fransoise Lust-Piquard non-commutative Khintchine inequality:

$$\mathbb{E}\,\|f(g) - \mathbb{E}f(g)\|_{\sigma_p} \le C^p\, p^{p/2}\, \|\Big[\Big(\sum_j \Big(\frac{\partial f}{\partial x_j}(g)\Big)^2\Big)^{1/2}\|_{\sigma_p}$$

More applications of Pisier–Poincaré inequality in Gaussian case

2) Scalar example, $f : \mathbb{R}^n \to \mathbb{R}$, p is large:

$$\left(\mathbb{E}|f(g) - \mathbb{E}f(g)|^p\right)^{1/p} \le C^p p^{1/2}\left(\mathbb{E}\|\nabla f(g)\|_{L^p(\ell^2)}\right)^{1/p}.$$

This leads to Gaussian concentration inequality with constants free of dimension: **Lipschitz functions of huge number of gaussian variables are very concentrated near their average**.
3) Scalar example, $p = 1$:

$$\mathbb{E}|f(g) - \mathbb{E}f(g)| \le \sqrt{\frac{\pi}{2}}\,\mathbb{E}\|\nabla f(g)\|\,.$$

This is Cheeger's gaussian inequality. Constant $\sqrt{\frac{\pi}{2}}$ is sharp and is attained on functions approximating the characteristic function of half-space.

One more application of Pisier's inequality

Let $A_1, \dots, A_n$ be in $M^{symm}_{d\times d}$. Non-commutative Khintchine inequality (NCK) says that

$$\mathbb{E}\|\sum_{i=1}^n g_i A_i\|_{op} \le \sqrt{\log d}\|\sum_{i=1}^n A_i^2\|_{op}^{1/2}.$$

(Margulis inequality.)
Then for any $f : \mathbb{R}^n \to M^{symm}_{d\times d}$, the following holds:

$$\mathbb{E}\|f - \mathbb{E}f\|_{op} \le^{Pisier} \frac{\pi}{2}\mathbb{E}\|\sum_{i=1}^n \tilde{g}_i D_i f(g)\|_{op} \le^{NCK}$$

$$\frac{\pi}{2}\sqrt{\log d}\,\mathbb{E}\|\sum_{i=1}^n (D_i f)^2\|_{op}^{1/2}.$$

For example f can be matrix valued degree 2 (or 100) polynomial of any number of variables.

Pisier–Poincaré inequality for Rademacher r. v.

Pisier has an analog of his Gaussian inequality for Rademacher situation: let $\{\delta_j\}_{j=1}^n$ are i.i.d standard Rademacher r. v. independent from $\{\varepsilon_j\}_{j=1}^n$. Then

$$\Big(\mathbb{E}\|f(\varepsilon)-\mathbb{E}f(\varepsilon)\|_X^p\Big)^{1/p}\le C(n)\Big(\mathbb{E}\|\sum_{j=1}^n\delta_j\frac{\partial f}{\partial x_j}(\varepsilon)\|_X^p\Big)^{1/p}\,.$$

Unfortunately, unlike in the Gaussian case the constant in Pisier's proof was dependent on n: $C(n)\le\log n+\log\log n+C_0$. Hytönen–Naor improved to $C(n)\le\log n+C_0$.

If constant were independent of n then that inequality plus the definition of Rademacher type would prove Enflo's conjecture immediately.

Talagrand's counterexample

Talagrand proved that $C(n) \approx \log n$ is possible. In his example $X = L^\infty(\{-1,1\}^n)$ and

$$f : \{-1,1\}^n \to L^\infty(\{-1,1\}^n), \quad f(\varepsilon, \tilde{\varepsilon}) := \max\left(0, \log \frac{\mathrm{dist}_{Hamming}(\varepsilon, \tilde{\varepsilon})}{\sqrt{n}}\right)$$

this is sharp: $C(n) \geq \frac{1}{2}\log n - C_0$.
All the previous works on Enflo's conjecture concentrated on Pisier's inequality

$$\left(\mathbb{E}\|f(\varepsilon) - \mathbb{E}f(\varepsilon)\|_X^p\right)^{1/p} \leq C(n)\left(\mathbb{E}\|\sum_{j=1}^n \delta_j \frac{\partial f}{\partial x_j}(\varepsilon)\|_X^p\right)^{1/p}.$$

and finding Banach spaces X, where one can prove $C(n) = O(1)$. The ultimate description of such X was not known till us.

Two questions

This raises two questions:

1. How to prove Enflo's conjecture by by-passing Talagrand's obstacle? May be one needs to modify Pisier's inequality?
2. What is the exact class of Banach spaces for which $C(n) = O(1)$ in Pisier's inequality for Rademacher r. v.?

Modification of Pisier's inequality for Rademacher variables. One more average, "skewed" Rademacher r. v.

Let

$$\mathbb{P}\{\xi_i(t) = \pm 1\} = \frac{1 \pm e^{-t}}{2}, \quad \delta_i(t) = \frac{\xi_i(t) - \mathbb{E}\xi_i}{(\operatorname{Var}\xi_i(t))^{1/2}}.$$

Theorem (Ivanisvili–Van Handel–Volberg)

$$\left(\mathbb{E}\|f(\varepsilon) - \mathbb{E}f\|_X^p\right)^{1/p} \leq \frac{\pi}{2}\int_0^\infty \left(\mathbb{E}\|\sum_{i=1}^n \delta_i(t) D_i f(\varepsilon)\|_X^p\right)^{1/p} d\mu(t),$$

for a certain concrete probability measure $d\mu(t)$ on $(0, \infty)$.

Corollary

By central limit theorem one gets Pisier gaussian inequality.

What are Banach spaces where Pisier's original inequality has $C(n) = O(1)$?

Theorem (Ivanisvili–Van Handel–Volberg)

Pisier's original inequality

$$\Big(\mathbb{E}\|f(\varepsilon) - \mathbb{E}f\|_X^p\Big)^{1/p} \le C\Big(\mathbb{E}\|\sum_{i=1}^n \delta_i D_i f(\varepsilon)\|_X^p\Big)^{1/p},$$

holds with constant independent of n if and only if X has finite co-type. (Here $\{\delta_i\}$ are i.i.d standard Rademacher r. v. independent of $\{\varepsilon_i\}$.)

How it all started. Cheeger inequality on Hamming cube. Improving Lust-Piquard's constant

We saw gaussian (isoperimetric) inequality in the form of Cheeger's inequality:

$$\mathbb{E}|f(g) - \mathbb{E}f(g)| \le \sqrt{\frac{\pi}{2}}\mathbb{E}[|\nabla f|_{\ell^2_n}],$$

where $f : \mathbb{R}^n \to \mathbb{R}$, and g is standard gaussian vector in $\mathbb{R}^n$. Constant is sharp. What about our Hamming cube? Francoise Lust-Piquard proved on cube

$$\mathbb{E}|f(\varepsilon) - \mathbb{E}f(\varepsilon)| \le \frac{\pi}{2}\mathbb{E}[|\nabla f|_{\ell^2_n}].$$

We came to Enflo's conjecture technique via improving Lust-Piquard's constant:

Theorem

$$\mathbb{E}|f(\varepsilon) - \mathbb{E}f(\varepsilon)| \le \left(\frac{\pi}{2} - 10^{-10}\right)\mathbb{E}[|\nabla f|_{\ell^2_n}].$$

Why L^1-Poincaré inequality on Hamming cube?

1. It is closely related to Erdös–Rényi graphs and Margulis' sharp threshold theorem;
2. Poincaré inequalities on Hamming cube, scalar or X-valued, are closely related to singular integral theory on Hamming cube, still in making;
3. Francoise's proof of L^1-Poincaré inequality on Hamming cube was via **quantum random variables**. Which was amazing. At least for me.
4. Our approach on previous slides gives very short proof of Talagrand's conjecture on improved Poincaré inequality for Boolean functions (proved in 2020 by Eldan–Gross), see below.

Example of singular integral problem on Hamming cube

Theorem (Ivanishvili–Van Handel–Volberg)

For any functions $f_i : \{-1,1\}^n \to X$, $i = 1, \dots, n$, $p \in [1, \infty)$, we have

$$\mathbb{E}\|\sum_{i=1}^n \Delta^{-1} D_i f_i\|_X^p \le C(q,p)\mathbb{E}\|\sum_{i=1}^n \delta_i f_i\|_X^p \qquad (**)$$

if and only if $(X, \|\cdot\|)$ be a Banach space of finite co-type q.

(Put $f_i = D_i f$ to come back to Pisier's original form of inequality.)

Theorem (I–vH–V.)

Let X be a Banach space of finite co-type q. Let $1 \le p < \infty$, $\varepsilon > 0$. Then for any functions $f_i : \{-1,1\}^n \to X$, $i = 1, \dots, n$:

$$\mathbb{E}\|\sum_{i=1}^n \Delta^{\frac{1}{\max(p,q)} - 1 - \varepsilon} D_i f_i\|^p \le C(q,p,\varepsilon)\mathbb{E}\|\sum_{i=1}^n \delta_i f_i\|^p \qquad (**)$$

Gradient via square root of Laplacian. Riesz transforms from above. Unexpected effect with $X \in UMD$.

Inequality below of type (*) for $L^p(X)$ always implies (**) for $L^{p'}(X^*)$

$$\mathbb{E}\|\sum_{i=1}^{n} \delta_i \Delta^{-a} D_i f\|_X^p \le C_p \,\mathbb{E}\|f\|_X^p, \quad \frac{1}{2} \le a \le 1 \quad (*)$$

1) holds for $1 < p < \infty$, X is UMD, **gaussian case** (Pisier, 1986);
2) holds for $2 \le p < \infty$, $X = \mathbb{R}$, **Hamming cube case** (Bakry-**25 pages of probability**, Lust-Piquard 98-quantum r. v., us–**Bellman function proof**);
3) does not hold $1 < p < 2$, $X = \mathbb{R}$, **Hamming cube case**.
4) **Does it hold for** $X \in UMD$**,** $2 \le p < \infty$ for **Hamming cube case**? The answer is very unexpected: NO.
5) **Scalar question**. Let $1 < p < 2$, is it true that

$$\mathbb{E}|\{\Delta^{-1/p} D_i f\}|_{\ell^2}^p \le C\,\mathbb{E}|f|^p\,.$$

With $\Delta^{-1/2} D_i f$ false!!!, but morally $\Delta^{-1/p} D_i f \le \Delta^{-1/2} D_i f$ $p < 2$.

Square root of Laplacian via gradient. Riesz transforms from below. Towards quantum r. v.

We come to results of type $(+)$:

$$\|\Delta^{1/2} f\|_X^p \le C_p \,\mathbb{E}\|\sum_i \delta_i D_i f(\varepsilon)\|_X^p, \quad (+)$$

for the "simplest case" of $\{\varepsilon_i\}, \{\delta_i\}$ are all i.i.d. Rademacher ± 1 random variables. Scalar case: Bakry, Lust-Piquard. Generalization/reformulation (just put $f_1 = \cdots = f_n = f$ below):

$$\|\sum_i \Delta^{-1/2} D_i f_i\|_X^p \le C_p \,\mathbb{E}\|\sum_i \delta_i D_i f_i(\varepsilon)\|_X^p,$$

Conjecture. It holds for any X of finite co-type, $1 < p < \infty$. In particular, for any $X = L^q, q < \infty$.

Theorem

1) $1 < p \le 2$, $X = L^q, 1 \le q \le 2$, then holds. 2) For $p > 2$ this holds for $X = L^q$ if $2 \le q \le p$. We do not know the rest.

Regime $1 < p \leq 2$, $X = L^q, 1 \leq q \leq 2$

We use our formula with unbalanced Rademacher r. v.: modification of Pisier's formula.

Regime $2 < p < \infty, X = L^q, 2 \le q \le p$. X-valued Riesz transforms on discrete cube

Theorem

1) $1 < p \le 2$, $X = L^q, 1 \le q \le 2$. 2) For $p > 2$ this holds for $X = L^q$ if $2 \le q \le p$. Namely

$$\|\sum_i \Delta^{-1/2} D_i f_i\|_X^p \le C_p \mathbb{E}\|\sum_i \delta_i D_i f_i(\varepsilon)\|_X^p,$$

We do not know the rest of regimes.

The proof below is just for $X = \mathbb{R}$, **the simplest case–the scalar case**. Of course co-type $q = 2$ for this case $X = \mathbb{R}$. Still the scalar case–a theorem about the usual functions of Rademacher variables–**required quantum random variables.**

So for usual real valued functions of Rademacher (Bernoulli) r. v. ε_i, we want to prove the following:

$$\mathbb{E}_\varepsilon|\sum_i \Delta^{-1/2} D_i f_i(\varepsilon)|^p \le C_p \mathbb{E}_{\delta,\varepsilon}|\sum_i \delta_i D_i f_i(\varepsilon)|^p \quad (+1)$$

1. Quantum r. v. $2^n \times 2^n$ matrices Q_A, P_B

Let

$$Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; P = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \; U = iQP,$$

They have anti-commutative relationship

$$QP = -PQ\,.$$

Let $Q_j = I \otimes \dots Q \otimes I \dots \otimes I$, $P_j = I \otimes \dots P \otimes I \dots \otimes I$, on j-th place. These are independent non-commutative random variables in the sense of trace = sum of diagonal elements divided by 2^n. Then $Q_A = \Pi_{i \in A} Q_i$. For any $f = \sum_{A \subset [n]} \hat{f}(A) \varepsilon^A$, the reasoning non-commutative scheme dictates to assign on non-commutative object, a matrix from $\mathcal{M}_{2^n}$ given by

$$T_f = \sum_{A \subset [n]} \hat{f}(A) Q_A\,.$$

Such matrices form commutative sub-algebra $M_{2^n} \subseteq \mathcal{M}_{2^n}$.

Now one considers algebra generated by Q_j, P_j (this is algebra of all matrices $\mathcal{M}_{2^n}$). We have a projection $\mathcal{P}$ from multi-linear polynomials in P_j, Q_j (notice $P^2 = I, Q^2 = I$) that kills everything except terms having only $Q's$.
Small (really easy) algebra shows that $\mathcal{P}$ can be written as $\rho\,\mathrm{Diag}\,\rho^*$, where ρ is a unitary operator, and Diag, is the operator on matrices that just kills all matrix elements except the diagonal. This *Diag* is obviously the contraction on Schatten-von Neumann class S_p for any $p \in [1, \infty]$ (obvious for Hilbert–Schmidt ($p = 2$) class and for bounded operators ($p = \infty$)–so interpolation does that). Here $\rho = r \otimes \cdots \otimes r$, where r is "Hadamard gate":

$$r = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Operators ∂_j, D_j, such that $\partial_j \varepsilon_i = \delta_{ij}$ and $D_j = \varepsilon_j \partial_j$, can be considered acting on M_{2^n} and on $\mathcal{M}_{2^n}$ in a canonical way.

Non-commutative semigroup is given on elementary Q_A, P_A by:

$$\mathcal{R}(\theta) Q_A = \Pi_{j\in A}(Q_j \cos\theta + P_j \sin\theta), \mathcal{R}(\theta) P_A = \Pi_{j\in A}(P_j \cos\theta - Q_j \sin\theta)\,.$$

Operator $\mathcal{R}(\theta)$ is an automorphism of algebra $\mathcal{M}_{2^n}$ preserving all S_p norms. $\mathcal{R}(\theta)(T)$ is given by $R(\theta)^* T R(\theta)$, where $R(\theta)$ is a unitary matrix which is n-fold tensor product of

$$\rho_\theta = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix}$$

Semigroup $\mathcal{R}(\theta)$ can be written down as

$$\mathcal{R}(\theta) T = e^{\theta \mathcal{D}}(T), \quad \text{where } \mathcal{D}(T) = \sum_i P_i D_i(T), \forall T \in M_{2^n}\,. \tag{1}$$

Pisier's lemma:

Lemma

The odd function $\frac{sgn\theta}{t(\theta)} = \frac{sgn\theta}{(-\log\cos\theta)^{1/2}}$ *on* $[-\pi/2, \pi/2]$ *is such that*
a) $\phi(\theta) - \cot(\theta/2)$ *is bounded and*

$$b)\ \forall m \geq 0,\ \int_{-\pi}^{\pi} \cos^m\theta \sin\theta \,\frac{sgn\theta}{t(\theta)}\, d\theta = c\,\frac{1}{\sqrt{m+1}}\,.$$

Notice (here $t(\theta) := (-\log\cos\theta)^{1/2}$):

$$D_j \Delta^{-1/2} Q_A = \mathcal{P}\Big(\int_{-\pi/2}^{\pi/2} \frac{\operatorname{sgn}(\theta)}{t(\theta)} e^{\theta \mathcal{D}} P_j \cdot [\partial_j Q_A] \Big) \, . \qquad (2)$$

In fact, if $j \notin A$ both sides are zero. Now let $j \in A$. Then

$$e^{\theta\mathcal{D}} P_j \partial_j Q_A = e^{\theta\mathcal{D}} P_j Q_{A\setminus j} =$$

$$\prod_{s\in A, s<j} (\cos\theta Q_s + \sin\theta P_s)(\sin\theta\, Q_j - \cos\theta P_j) \prod_{s\in A, s>j} (\cos\theta Q_s + \sin\theta P_s)$$

Thus

$$\cos^{|A|-1}\theta \sin\theta \cdot Q_A = \mathcal{P}\big(e^{\theta\mathcal{D}} P_j \cdot [\partial_j Q_A]\big) \, .$$

And now we integrate against $\frac{\operatorname{sgn}(\theta)}{t(\theta)}$ and use Pisier's lemma.

Formula (2) gives us (for any $T = T_f \in M_{2^n}$), $D_j\Delta^{-1/2}(\varepsilon_j \partial_j f) =$

$$D_j\Delta^{-1/2} f {=} \mathcal{P}\Big(\int_{-\pi/2}^{\pi/2} \frac{\operatorname{sgn}(\theta)}{(-\log\cos\theta)^{1/2}} e^{\theta\mathcal{D}} P_j \cdot [\partial_j T_f] \Big) \,. \qquad (3)$$

Having proved formula (3) we can finish the proof of our end-point inequality. Let $f_1, \dots, f_n$, we consider $T_1 = T_{f_1}, \dots, T_n = T_{f_n}$. Then

$$\sum_j D_j\Delta^{-1/2} f_j = \mathcal{P}\Big(\int_{-\pi/2}^{\pi/2} \frac{\operatorname{sgn}(\theta)}{(-\log\cos\theta)^{1/2}} e^{\theta\mathcal{D}} \sum_j P_j \cdot [\partial_j T_j] \Big) \,.$$

Now using the facts that $\mathcal{P}$ is a contraction in S_p, that our singular integral is bounded in S_p, $1 < p < \infty$, and that semigroup $e^{\theta\mathcal{D}}$ is bounded in S_p we get that

$$\|\sum_j D_j \Delta^{-1/2} f_j\|_p \leq C\, p\, \|\sum_j P_j \cdot [\partial_j T_j]\|_{S^p}\,.$$

Even though our problem is about scalar functions we used here 1) non-commutative model, 2) we used here Burkholder's and Bourgain's estimates of the boundedness of the Hilbert transform $\int_{-\pi/2}^{\pi/2} \frac{\mathrm{sgn}\,(\theta)}{t(\theta)} \star$ on X-valued $L^p(X)$ spaces. (Actually constant p here is due to Bourgain's $L_p(S_p)$ Hilbert transform estimate, it is a subtle place.)
We are almost done–unfortunately LHS is about functions, but RHS is about operators.

8. Quantum r. v. Towards non-commutative Khintchine inequality

Funny trick:

$$\|\sum_j P_j \cdot [\partial_j T_j]\|_{S^p} = \|\sum_j \varepsilon_j P_j \cdot [\partial_j T_j]\|_{S^p}$$

for any signs $\varepsilon_j = \pm 1$. This is because

$$\|\sum_j P_j \cdot [\partial_j T_j]\|_{S^p} = \|Q_i \cdot \sum_j P_j \cdot [\partial_j T_j] \cdot Q_i\|_{S^p},$$

and Q_i travels freely–and then **anti-commutes** with P_i and disappears.

9. Quantum r. v. Non-commutative Khintchine inequality

Now using non-commutative Khintchine inequality of Lust-Piquard we get for $p \geq 2$:

$$\| \sum_j D_j \Delta^{-1/2} f_j \|_p \leq C\, p \,\| \sum_j P_j \cdot [\partial_j T_j] \|_{S^p} \leq Cp \mathbb{E}_\varepsilon \| \sum_j \varepsilon_j P_j \cdot [\partial_j T_j] \|_{S^p} \leq$$

$$C\, p^{3/2} \big(\| (\sum_i (\partial_i T_i)^* (\partial_i T_i))^{1/2} \|_{S^p} +$$

$$\| (\sum_i P_i (\partial_i T_i)(\partial_i T_i)^* P_i)^{1/2} \|_{S^p} \big)\,.$$

The second term is equal to the first one, because $\partial_i T_i$ does not have Q_i in it, so P_i can be carried be through to make $P_i^2 = I$.
But if $T_i = T_{f_i}$ then

$$\| (\sum_i (\partial_i T_i)^* (\partial_i T_i))^{1/2} \|_{S^p} = \| (\sum_i |D_i f_i|^2)^{1/2} \|_p \,.$$

10. Quantum r. v. $L^p(L^q), 2 \le q \le p$, on Hamming cube

Nothing should depend on n, N below. Consider the space (algebra) of matrices M, where M is block diagonal with n blocks each $N \times N$. Call blocks $m_1, ..., m_n$.
Consider the following norm on M (it is a norm):
$\|M\|_{S^p(\ell^q)} := \|[(m_1 * m_1)^{q/2} + ... + (m_n * m_n)^{q/2}]^{1/q}\|_{S_p}$.
Now what about NCK (non-commutative Khintchine) in this norm? Namely, let $p \ge 2$. Let $M_1, ..., M_k$ be a collection of such block diagonal matrices.
Then $M_i * M_i$ are also block diagonal, $\sum_i M_i * M_i$ is also such and $[\sum_i M_i * M_i]^{1/2}$ is also a block diagonal matrix as above.
So we can ask the question: consider $p \ge 2$ and consider $\mathbb{E}_\varepsilon \|\varepsilon_1 M_1 + ... + \varepsilon_k M_k\|_{S_p(\ell^q)}$. Is it bounded (independent of n, N, k) by $C_{p,q} \left(\|[\sum_i M_i * M_i]^{1/2}\|_{S_p(\ell^q)} + \|[\sum_i M_i M_i *]^{1/2}\|_{S_p(\ell^q)}\right)$?
We know that this is true in the regime $2 \le q \le p$, and in the regime $p \le q \le 2$. We are most interested in the regime $2 \le p \le q$, where we do not know the answer.

A. Margulis sharp threshold theorem on Erdös–Rényi graphs and L^1-Poincaré type inequalities

Poincaré inequalities on Hamming cube: analysis, combinatorics, probability

B. Hamming cube

Consider the Hamming cube $\{-1,1\}^n$ of an arbitrary dimension $n \geq 1$. For any $f : \{-1,1\}^n \to \mathbb{R}$ define the discrete gradient

$$|\nabla f|^2(x) = \sum_{y \sim x} \left(\frac{f(x) - f(y)}{2} \right)^2,$$

where the summation is over all neighbor vertices of x in $\{-1,1\}^n$. Set

$$\mathbb{E}f = \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} f(x).$$

C. Discrete surface measure

Let $f = \mathbf{1}_A$. Then

$|\nabla \mathbf{1}_A|^2(x) := w_A(x) :=$ number of neighbors of x of opposite color .

Consider also

$$h_A(x) := w_A(x)\mathbf{1}_A(x)\,.$$

We are interested in estimates of

$$\int \sqrt{w_A(x)}\,d\mu(x), \quad \int \sqrt{h_A(x)}\,d\mu(x)$$

from below (μ can be μ_p, $0 < p < 1$) by p and $\mu(A)$ but not on n. Why? Why this is even possible? Why square root?

This is the portion of boundary edges of a subset A of a Hamming cube:

$$\int w_A(x)\, d\mu(x)\,.$$

Given $t = \mu(A)$ what is the minimal number of boundary edges and what are extremal sets? Harper: The edge-extremal sets are given by the first points in the lexicographical order of Hamming cube. In particular, when $t = 2^{-k}$, these are sub-cubes: $A = \{(1, \dots, 1) \times \{-1, 1\}^{n-k}\}$. For $t = \frac{1}{2}$ these are exactly faces of the cube. For such sets

$$\int w_A(x)\, d\mu(x) = 2t \log \frac{1}{t}, \quad t = \mu(A),$$

which is very far from the estimate, as $\int \sqrt{w_A(x)}\, d\mu(x) \geq I(t) \asymp t\sqrt{\log \frac{1}{t}}$, $t \approx 0$.

E. Square root estimate implies Harper's estimate

By Cauchy inequality

$$\int w_A(x)\, d\mu(x) \ge \frac{(\int \sqrt{w_A(x)}\, d\mu(x))^2}{\mu(\partial A)} \ge c\, \frac{t^2 \log \frac{1}{t},}{t} \quad \textit{if} \;\; t = \mu(A),$$

so it follows from

$$\int \sqrt{w_A(x)}\, d\mu(x) \ge c\, t \sqrt{\log \frac{1}{t}}\,.$$

F. Bobkov's estimates for $\sqrt{w_A}$

$$\int_{\partial_A} \sqrt{w_A}\, d\mu_p \geq I(\mu_p(A)),$$

$I(t) := \Phi'(\Phi^{-1}(t))$, $\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du$.
But $\sqrt{w_A}$ is essentially $|\nabla \mathbf{1}_A|$. What if we want to have **functional** inequality of the type above? One such thing exists as Bobkov inequality **for any** f**,** $0 \leq f \leq 1$**, on Hamming cube**:

$$\int \sqrt{|\nabla f|^2 + I^2(f)}\, d\mu \geq I(\int f\, d\mu).$$

In fact, this is discrete generalization of Bobkov's **gaussian** inequality. When $\mu(A) = 1/2$, we get L^1-Poincaré inequality for characteristic functions on Hamming cube:

$$\mathbb{E}|\mathbf{1}_A - 1/2| = 1/2 = \sqrt{\frac{\pi}{2}} \cdot \frac{1}{\sqrt{2\pi}} = \sqrt{\frac{\pi}{2}}\, \mathbb{E}|\nabla \mathbf{1}_A| = \sqrt{\frac{\pi}{2}}\, \mathbb{E}\sqrt{w_A}\,.$$

PDE $M_{xx} M_{yy} - M_{xy}^2 + \frac{M_y M_{yy}}{y} = 0$, $M(x, y) = -\sqrt{y^2 + I^2(x)}$.

Margulis normalized the integral $\int w_A(x)\,d\mu(x)$, he considered

$$\mu(\partial A)\int_{\partial A} w_A(x)\,d\mu(x) \geq \text{ and } \quad \mu_p(\partial_+ A)\int_{\partial_+ A} h_A\,d\mu_p \geq c(p,\mu(A)).$$

By Schwarz inequality this is bounded $\geq \int_{\partial A}\sqrt{w_A}\,d\mu$, $\geq \int_{\partial_+ A}\sqrt{h_A}\,d\mu$. And Talagrand made the above Margulis estimates more precise by introducing **square root**:

$$\int_{\partial_+ A}\sqrt{h_A}\,d\mu \geq c_p\mu_p(A)(1-\mu_p(A))\sqrt{\log\left(\frac{1}{\mu_p(A)(1-\mu_p(A))}\right)}$$

The "surface measures" 1) $\int_{\partial A}\sqrt{w_A}\,d\mu$, 2) $\int_{\partial_+ A}\sqrt{h_A}\,d\mu$ turned out to be the desired discrete analogs of the gaussian perimeter $\gamma_n^+(\partial A)$.

H. Margulis–Russo lemma and its consequences

For **monotone** subsets A of Hamming cube

$$\frac{d}{dp}\mu_p(A) = \frac{1}{p}\int h_A \, d\mu_p \, .$$

Suppose $h_A(x) \geq k$, $x \in \partial_+ A$. Then Talagrand's estimate gives:

$$\frac{d}{dp}\mu_p(A) \geq \frac{\sqrt{k}}{p}\int \sqrt{h_A}\, d\mu_p \geq \frac{\sqrt{k}}{p}\mu_p(A)(1-\mu_p(A))\, .$$

$$\int_{p_1}^{p_2} \frac{d}{dp}\log\frac{\mu_p(A)}{1-\mu_p(A)} \geq \int_{p_1}^{p_2}\frac{\sqrt{k}}{p}dp \geq \sqrt{k}(p_2 - p_1)\, .$$

If $p_1 < p_2$ and $\mu_{p_1}(A) = \varepsilon = 0.1$, $\mu_{p_2}(A) = 1-\varepsilon = 0.9$, then $\sqrt{k}(p_2 - p_1) \leq 2\log\frac{1-\varepsilon}{\varepsilon} = 2\log 9$.

$$\sqrt{k}(p_2 - p_1) \le 2\log\frac{1-\varepsilon}{\varepsilon} = 2\log 9\,.$$

This means that on interval of p' of size $\asymp \frac{1}{\sqrt{k}}$ the probability of A ($=\mu_p(A)$) changes from 0.9 to 0.1–very sharp change if k is big. This is called **sharp threshold theorem of Margulis**.
Let G be a fixed connected graph with very large n number of edges. Let us delete edges independently with probability p. This Erdös–Renyi random graphs are in one to one correspondence with vertices of Hamming cube $(\{-1,1\}^n, \mu_p)$, where 1 on i-th place of vertex means that i-th edge is deleted. Let G has **connectivity** k–G rests connected if less than k edges are deleted. Let A be vertices corresponding to disconnected graphs. It is obviously monotone.
Exercise: If $x \in \partial_+ A$ then it has at least k connected neighbors, that is $h_A(x) \ge k$.

$$\sqrt{k}(p_2 - p_1) \le 2\log\frac{1-\varepsilon}{\varepsilon} = 2\log 9\,.$$

This means that on interval of p' of size $\asymp \frac{1}{\sqrt{k}}$ the probability of A ($=\mu_p(A)$) changes from 0.9 to 0.1–very sharp change if k is big. This is called **sharp threshold theorem of Margulis**.

Let G be a fixed connected graph with very large n number of edges. Let us delete edges independently with probability p. This Erdös–Renyi random graphs are in one to one correspondence with vertices of Hamming cube $(\{-1,1\}^n, \mu_p)$, where 1 on i-th place of vertex means that i-th edge is deleted. Let G has **connectivity** k–G rests connected if less than k edges are deleted. Let A be vertices corresponding to disconnected graphs. It is obviously monotone.

Exercise: If $x \in \partial_+ A$ then it has at least k connected neighbors, that is $h_A(x) \ge k$.

Conclusion: if connectivity a priori is k then on an interval of "rupture" probability $\approx k^{-1/2}$ the network goes from large probability of being disrupted to small probability.

J. Coming back to Poincaré inequalities on Hamming cube

$$\int |f - \int f d\mu|^q \, d\mu \le C_q \int |\nabla f|^q \, d\mu, \quad 1 \le q < \infty\,.$$

New notation

$$\mathbb{E}|f - \mathbb{E}f|^q \le C_q \, \mathbb{E}|\nabla f|^q\,.$$

The sharp constant is known only for $q = 2$: $C_2 = 1$. Extremal functions are characteristic functions of the faces.

Theorem (Ivanisvili–Volberg)

For $1 < q \le 2$, any $n \ge 1$ and any $f : \{-1,1\}^n \to \mathbb{R}$ we obtain $(\mathbb{E}|f|^q - |\mathbb{E}f|^q) \le C_q^q \|\nabla f\|_q^q$, where $C(q)^{-1}$ is the smallest positive zero of the confluent hypergeometric function ${}_1F_1(q/2(1-q), 1/2, x^2/2)$.

Approach is based on a certain duality between the classical square function estimates on Euclidean space and the gradient estimates on the Hamming cube. Constant $C(2) = 1, C(1+) = 0$. The latter is not good. What is known about $C(1)$?

K. Cheeger's inequality and Ben Efraim–Lust-Piquard

On gaussian space

$$\int \left| f - \int f d\gamma_n \right| d\gamma_n \le \sqrt{\frac{\pi}{2}} \int |\nabla f| \, d\mu, \quad 1 \le q < \infty$$

and the constant is sharp. Proved by Cheeger and then by Maurey–Pisier and differently by Ledoux. On Hamming cube Ben Efraim–Lust-Piquard proved

$$\mathbb{E}|f - \mathbb{E}f| \le \frac{\pi}{2} \mathbb{E}|\nabla f| \,.$$

The method of the proof was **absolutely vertiginous** and I got hooked.
The idea was to embed this commutative problem about functions to non-commutative problem about operators.

L. Lust-Piquard's idea of $\mathbb{E}|f-\mathbb{E}f| \le \frac{\pi}{2}\mathbb{E}|\nabla f|$

$\{-1,1\}^n$ is Hamming cube of dimension n. Pauli matrices

$$Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad P = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \quad U = iQP = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Then Q is the matrix of mult. by x_1 in $L^2(\{-1,1\})$ in basis $\{\mathbf{1}, \mathbf{x_1}\}$. Alg. spanned by Q is isom. $L^\infty(\{-1,1\})$–mult. oper. on $L^2(\{-1,1\})$, and $\mathcal{M}_2$ is non-comm. alg. spanned by P, Q. $\mathcal{M}_{2^n} = \mathcal{M}_2 \otimes \dots \mathcal{M}_2$- alg. of all $2^n \times 2^n$ matrices. M_{2^n} is commutative sub. alg. generated by $Q_j := I \otimes \dots Q \otimes \dots I$, which are oper. of mult. on x_j in $L^2(\{-1,1\}^n)$. $\mathcal{E}_n : \mathcal{M}_{2^n} \to M_{2^n}$ $(\mathrm{trace}(Q_A \mathcal{E}_n(S)) = \mathrm{trace}(Q_A S))$. Consider multipl. oper. M_f on $L^2(\{-1,1\}^n)$. $R_\theta := \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix}$, $\mathcal{R}_\theta := R_\theta^{\otimes n}$.

Theorem

Then $\cos^\Delta \theta(f) = \mathcal{E}_n\left(\mathcal{R}_\theta^* M_f \mathcal{R}_\theta\right) = \mathcal{E}_n\left(e^{\theta\mathcal{D}} M_f\right)$, $\mathcal{D}$ generator of semi-group flow $T \to \mathcal{R}_\theta^* T \mathcal{R}_\theta$ on matrices.

K1. Lust-Piquard's idea of $\mathbb{E}|f - \mathbb{E}f| \le \frac{\pi}{2}\mathbb{E}|\nabla f|$

Theorem

Let f be a scalar function on Hamming cube, M_f is a multiplication on f operator. Then $(\cos\theta)^{\Delta}(f) = \mathcal{E}_n\big(\mathcal{R}_\theta^* T \mathcal{R}_\theta\big) = \mathcal{E}_n\big(e^{\theta\mathcal{D}} M_f\big)$, $\mathcal{D}$ generator of semi-group flow $T \to \mathcal{R}_\theta^* T \mathcal{R}_\theta$ on matrices.

Theorem

$L^\infty(\{-1,1\}^n) = M_{2^n} \subset \mathcal{M}_{2^n}$ *extends to isometry*
$L^1(\{-1,1\}^n) \to \sigma_1(\mathcal{M}_{2^n}, 2^{-n}\,\mathrm{trace})$ *(trace class operators).*

Theorem

Projection $\mathcal{E}_n : \mathcal{M}_{2^n} \to M_{2^n}$ is contraction in σ_1 (trace) norm.

Theorem (Ben Efraim–Lust-Piquard)

$\| \int_0^{\pi/2} \varphi(\theta) \frac{d}{d\theta}(\cos\theta)^{\Delta} f \, d\theta \|_{L^1} \le \|\nabla f\|_{L^1} \int_0^{\pi/2} |\varphi(\theta)|.$

T1. Definition of influences for Boolean functions

For any function on discrete cube $f : \{-1,1\}^n \to \mathbb{R}$ Poincaré inequality is

$$Var(f) := \mathbb{E}|f - \mathbb{E}f|^2 \leq \mathbb{E}\|Df\|^2, \quad Df = (D_1 f, \dots, D_n f)\,.$$

For linear functions it is very good, for Boolean functions ($f : \{-1,1\}^n \to \{-1,1\}$ or $f : \{-1,1\}^n \to \{0,1\}$) it is quite bad. For Boolean functions $Inf_i f :=$

$$\mathbb{E}|D_i f| = \mathbb{E}|D_i f|^2 = \Pr(\text{of}\, x \in \{-1,1\}^n \,\text{such that the flip of}\, x_i \,\text{changes}\, f)\,.$$

So if $f(x) = maj_n(x) = 2\mathbf{1}_{\sum x_j \geq n/2} - 1$ we can see $Inf_i f \asymp n^{-1/2}$, $\mathbb{E}\|Df\|^2 = \sum Inf_i f \asymp n^{1/2}$ but $Var(f) = 1$. Very much off.

T2. Better Poincaré inequality for booleans

KKL (Kahn–Kalai–Linial, 1988) and Funny corollaries,

$$Var(f) \log \frac{1}{\max_i Inf_i f} \le C\mathbb{E}\|Df\|^2 .$$

Corollary

Let f be Boolean. Then $\max_i Inf_i f \ge c\frac{\log n}{n} Var(f)$.

$Var(f) \le \mathbb{E}\|Df\|^2 = \sum_{i=1}^n Inf_i f \Rightarrow \max_i Inf_i f \ge \frac{1}{n} Var(f)$.

Corollary

Let f be Boolean and monotone (voting function). Let $\mathbb{E}f \ge -0.99$. Then candidate 1 can bribe selected $o(n)$ (actually $O(\frac{n}{\log n})$) votes in such a way that $\mathbb{E}f_{bribed} \ge 0.99$.

Here f_{bribed} means that there exists $J \subset \{1, 2, \dots, n\}$, $J| \le C\frac{n}{\log n}$, $\bar{J} := \{1, 2, \dots, n\} \setminus J$, such that $f_{bribed} = f_{\bar{J}|(1,\dots,1)\, on\, J}$.

T3. Another better Poincaré inequality

Theorem (Falik–Samorodnitsky, Rossignol, 2005)

Any function $f : \{-1,1\}^n \to \mathbb{R}$

$$Var(f) \log \frac{Var(f)}{\sum_{i=1}^n (\mathbb{E}|D_i f|)^2} \le C\mathbb{E}\|Df\|^2 .$$

Heat semigroup, $P_t f = \sum_{S\subset\{1,\dots,n\}} e^{-t|S|}\hat{f}(S)x^s$, we get

$$\frac{d}{dt} Var(P_t f) = -\mathbb{E}\|DP_t f\|^2 \le -\frac{1}{c} Var(P_t f) \log \frac{Var(P_t f)}{\sum(\mathbb{E}|D_i P_t f|)^2},$$

We integrate inequality (deleting $P_t f$ in denominator): $t \le 1$

$$Var(P_t f) \le Var(f) \Big(\frac{\sum(\mathbb{E}|D_i P_t f|)^2}{Var(f)}\Big)^{ct}$$

Any f!

T4. Keller–Kindler theorem

Theorem (Keller–Kindler, 2012)

Any function **Boolean** $f : \{-1,1\}^n \to \{0,1\}$

$$Var(P_t f) \leq C Var(f) \Big(\sum (\mathbb{E}|D_i P_t f|)^2 \Big)^{ct}.$$

Proof.

We saw that for any real f, $Var(P_t f) \leq Var(f) \left(\frac{\sum (\mathbb{E}|D_i P_t f|)^2}{Var(f)} \right)^{ct}$. If $Var(f) \geq (\sum (\mathbb{E}|D_i P_t f|)^2)^{1/2}$, we are immediately done.

If $Var(f) \leq (\sum (\mathbb{E}|D_i P_t f|)^2)^{1/2}$ we use that for **Boolean** f **by hypercontractivity**,
$Var(P_t f) \leq Var(f) Var(f)^{ct} \leq Var(f) (\sum (\mathbb{E}|D_i P_t f|)^2)^{ct/2}$ and we are also done.

□

T5. Ultimate better Poincaré inequality for boolean f. Talagrand's conjecture 1997

Theorem (Eldan–Gross, solution of Talagrand's conjecture, 2020)

Let f be boolean. Then

$$Var(f)\sqrt{\log\frac{e}{\sum_{i=1}^n (\mathbb{E}|D_i f|)^2}} \le C\mathbb{E}\|Df\|\,.$$

Compare with F-S, R and **it is not a square root**:

$$Var(f)\log\frac{Var(f)}{\sum_{i=1}^n (\mathbb{E}|D_i f|)^2} \le C\mathbb{E}\|Df\|^2\,.$$

Talagrand himself proved that there exists $a \in (0, 1/2]$ such that

$$Var(f)\Big(\log\frac{e}{\sum_{i=1}^n (\mathbb{E}|D_i f|)^2}\Big)^a\Big(\log\frac{e}{Var(f)}\Big)^{\frac{1}{2}-a} \le C\mathbb{E}\|Df\|\,.$$

T6. Short proof of Talagrand's conjecture

1) From Keller–Kindler theorem above it follows immediately that for Boolean f

$$t \ge t_* := 1/\log \frac{c}{\sum(\mathbb{E}|D_i f|)^2} \Rightarrow Var(P_t f) \le \frac{1}{2} Var(f)\,.$$

2) For $f : \{-1,1\}^n \to \{0,1\}$, $\mathbb{E}|f - \mathbb{E}f| = 2(Var(f) - Var(\mathbb{E}f))$. Similarly, $\mathbb{E}|f - P_t f| = 2(Var(f) - Var(P_t f))$. Hence,

$$Var(f) = \frac{1}{2}\mathbb{E}|f - P_t f| + Var(P_t f)\,.$$

3) $\mathbb{E}|f - P_t f| = \int_0^t \frac{e^{-s}}{\sqrt{1-e^{-2s}}} \mathbb{E}_{\varepsilon,\xi} |\sum_{i=1}^n \delta_i(t) D_i f(\varepsilon)|\, ds$. Our formula. So $\mathbb{E}|f - P_t f| \le C \int_0^t \frac{1}{\sqrt{s}} \mathbb{E}_\varepsilon \|Df(\varepsilon)\| ds = C\sqrt{t}\, \mathbb{E}\|Df\|$.
4) And here is the proof of Talagrand's conjecture, combine 2) then 3) then 1) and get

$$t = t_* \Rightarrow Var(f) \le C \sqrt{\frac{1}{\log \frac{c}{\sum(\mathbb{E}|D_i f|)^2}}}\, \mathbb{E}\|Df\| + \frac{1}{2} Var(f)\,.$$