The Ribe Program

Assaf Naor
Courant Institute, NYU

Rigidity in Banach spaces

- Mazur-Ulam theorem (1932): Let X,Y be Banach spaces and

$$f : X \to Y$$

an onto isometry. Then f is affine.

- Kadec theorem (1960): Any two separable infinite dimensional Banach spaces are homeomorphic.

Work of Lindenstrauss and later Enflo in the 1960s showed that if we add the assumption that the homeomorphisms in the Kadec theorem are quantitatively continuous, then not all separable infinite dimensional Banach spaces are equivalent.

The situation has been clarified in an important theorem of M. Ribe (1976).

Finite representability

Definition (R. C. James): A Banach space X is finitely representable in a Banach space Y if there exists a constant K such that for every finite dimensional subspace F of X there exists a linear operator

$$T : F \to Y$$

satisfying

$$\forall x \in F, \qquad K\|x\| \geq \|Tx\| \geq \|x\|$$

Examples

- The function space $L_p(\mu)$ is finitely representable in the sequence space ℓ_p.

- Hilbert space is finitely representable in every infinite dimensional Banach space (Dvoretzky's theorem).

Local properties

If X is finitely representable in Y then it inherits all the quantitative linear properties of finite dimensional subspaces of Y.

An example of a consequence: if for every integer n and every $x_1, \ldots, x_n \in Y$ we have

$$\frac{\sum_{\epsilon_1,\ldots,\epsilon_n=\pm 1} \|\epsilon_1 x_1 + \cdots + \epsilon_n x_n\|}{2^n} \lesssim \sqrt{\|x_1\|^2 + \cdots + \|x_n\|^2}$$

(e.g., if $Y = L_p,\ p \geq 2$)

Then X has the same property.

A bijection $f : X \to Y$ is called a uniform homeomorphism if f, f^{-1} are both uniformly continuous. Equivalently, for all $x, y \in X$,

$$\alpha(\|x - y\|) \geq \|f(x) - f(y)\| \geq \beta(\|x - y\|)$$

where $\lim_{t \to 0} \alpha(t) = 0$

Ribe's theorem (1976)

If X and Y are uniformly homeomorphic then X is finitely representable in Y and vice versa.

Other proofs: Heinrich- Mankiewicz (1982), Bourgain's discretization theorem (1987); Giladi-N.-Schechtman (2011), Li-N. (2011).

The Ribe program (Bourgain, 1986)

- Linear local properties of Banach spaces are “metric properties”. Redefine them using only the notion of distance.
- Once this is done, the definition is meaningful for general metric spaces.
- Extend the linear theory to general metric spaces.
- Applications to geometric objects that have nothing to do with linear spaces.

Rigidity in mathematics

- Equivalence in a “weaker” category actually implies equivalence in a “stronger” category.
- A powerful statement about the stronger category.
- The philosophy of the Ribe program: this can say something deep about the weaker category; concepts, invariants, and theorems that make sense in the stronger category can maybe be formulated in the weaker category.

A rigidity theorem might lead to a new understanding of the weaker category that is motivated by insights that originally came up naturally in the presence of more structure.

Reversing in this way the usual implications of a rigidity result is not obvious, and, if possible, might require much work.

Rademacher type

A randomized triangle inequality: for every Banach space X and every $x_1, \ldots, x_n \in X$,

$$\sum_{i=1}^{n} \|x_i\| \geq \mathbb{E}_{\epsilon_1,\ldots,\epsilon_n \in \{-1,1\}} \left\| \sum_{i=1}^{n} \epsilon_i x_i \right\|$$

Definition: X is said to have Rademacher type $p \geq 1$ if

$$\left(\sum_{i=1}^{n} \|x_i\|^p \right)^{1/p} \gtrsim \mathbb{E} \left\| \sum_{i=1}^{n} \epsilon_i x_i \right\|$$

- Hilbert space has Rademacher type 2 (by parallelogram identity).
- Every Banach space has Rademacher type 1.
- The space $L_p(\mu)$ has Rademacher type $\min\{p, 2\}$ (by Khinchine's inequality).

Type of metric spaces?

By Ribe's theorem, Rademacher type p is preserved under uniform homeomorphisms.

Can we define this notion using only distances between points (using no linear structure)?

- Enflo (Hilbert's fifth problem in infinite dimensions, 1960's and 1970's).
- Gromov (1983).
- Bourgain-Milman-Wolfson (1986).

The idea

Given $x_1, \ldots, x_n \in X$ define $f : \{-1,1\}^n \to X$ by $f(\epsilon) = f(\epsilon_1, \ldots, \epsilon_n) = \sum_{i=1}^n \epsilon_i x_i$

Using this notation, the Rademacher type p inequality becomes:

$$\mathbb{E}\left[\|f(\epsilon) - f(-\epsilon)\|\right]$$
$$\lesssim \left(\sum_{i=1}^{n} \mathbb{E}\left[\|f(\epsilon_1, \ldots, \epsilon_n) - f(\epsilon_1, \ldots, -\epsilon_i, \ldots, \epsilon_n)\|^p\right] \right)^{1/p}$$

$$\mathbb{E}\left[\|f(\epsilon) - f(-\epsilon)\|\right]$$

$$\lesssim \left(\sum_{i=1}^{n} \mathbb{E}\left[\|f(\epsilon_1, \ldots, \epsilon_n) - f(\epsilon_1, \ldots, -\epsilon_i, \ldots, \epsilon_n)\|^p\right]\right)^{1/p}$$

This inequality involves only distances, if we ignore the fact that f itself is defined using linear combinations of vectors.

- Definition. A metric space (X,d) has type p if for every $f : \{-1,1\}^n \to X$ we have

$$\mathbb{E}\left[d\left(f(\epsilon), f(-\epsilon)\right)\right]$$

$$\lesssim \left(\sum_{i=1}^{n} \mathbb{E}\left[d\left(f(\epsilon_1,\ldots,\epsilon_n), f(\epsilon_1,\ldots,-\epsilon_i,\ldots,\epsilon_n)\right)^p\right]\right)^{1/p}$$

- The implied constant may depend on X, but not on n or f.
- As in the case of Banach spaces, every metric space has type 1 (a randomized triangle inequality).

The problem

- Clearly, for Banach spaces the metric definition implies Rademacher type p, since Rademacher type p corresponds to a special class of functions f (linear functions).
- Do Banach spaces with Rademacher type p also have type p as metric spaces?
- Still open…

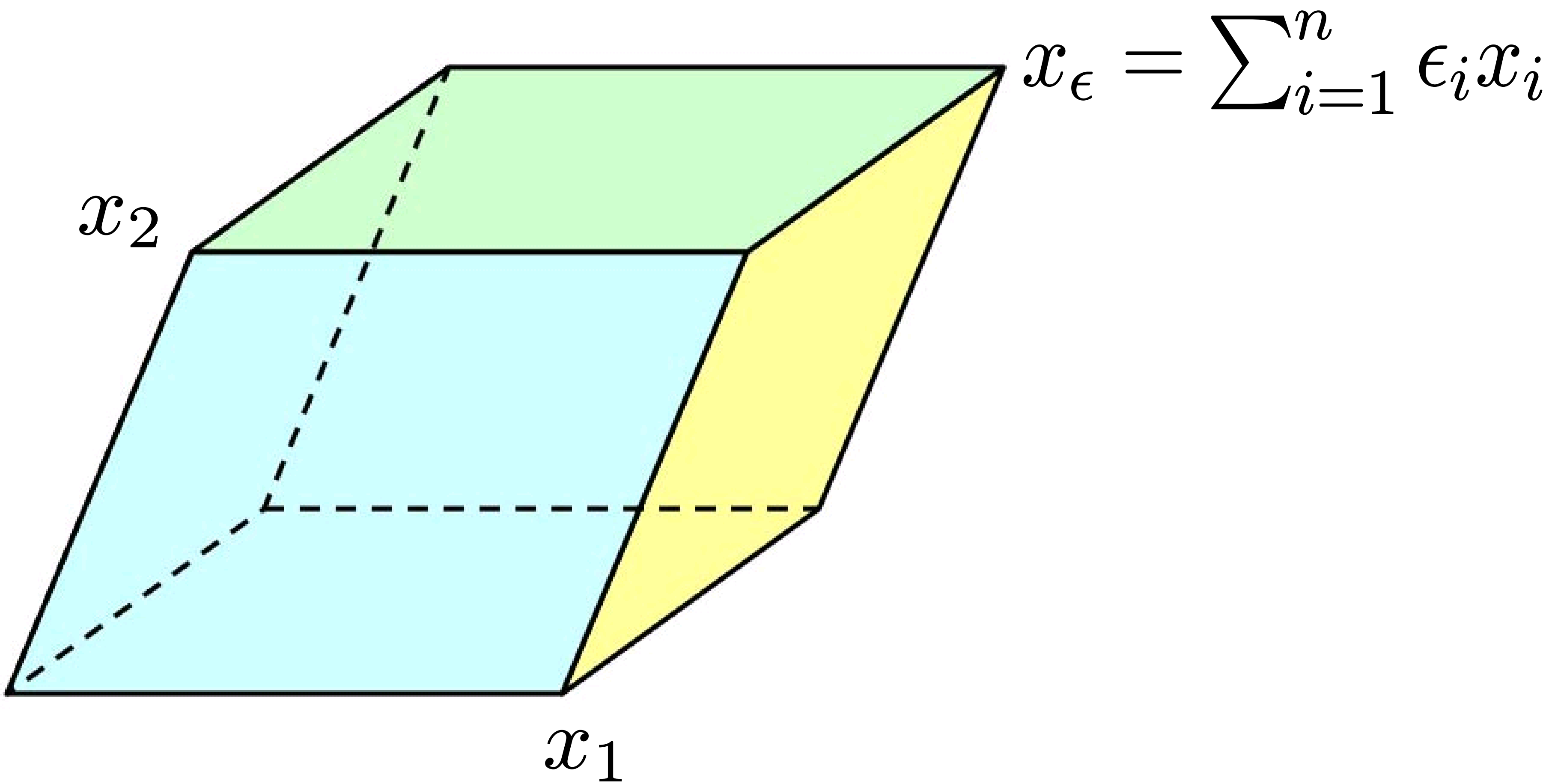
Bourgain-Milman-Wolfson, Pisier

Theorem. If a Banach space X has Rademacher type p then for every $\varepsilon>0$ it also has type $p-\varepsilon$ as a metric space.

The geometric “puzzle”

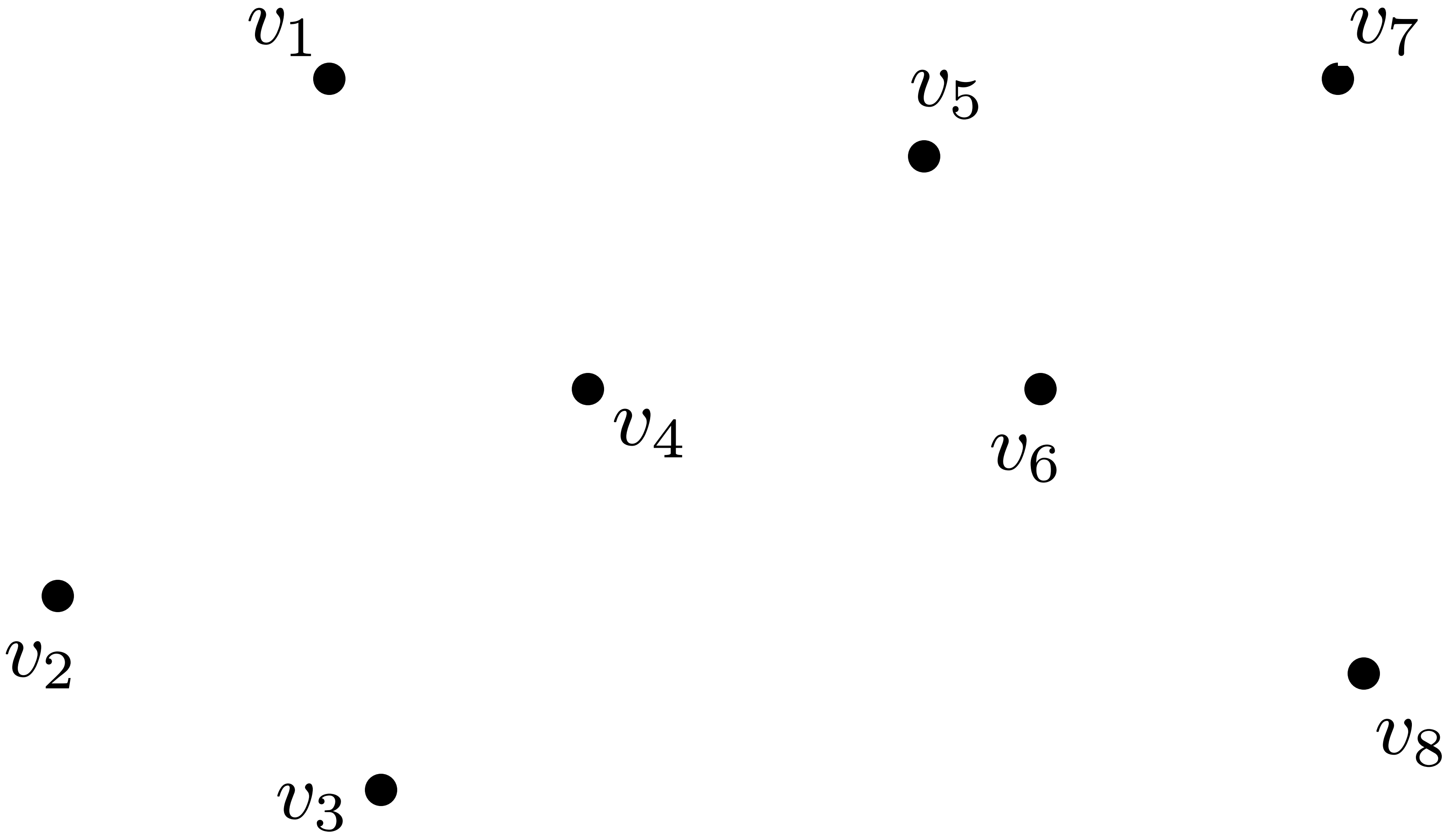
We are given a Banach space X for which we know the following inequality for all parallelopipeds:

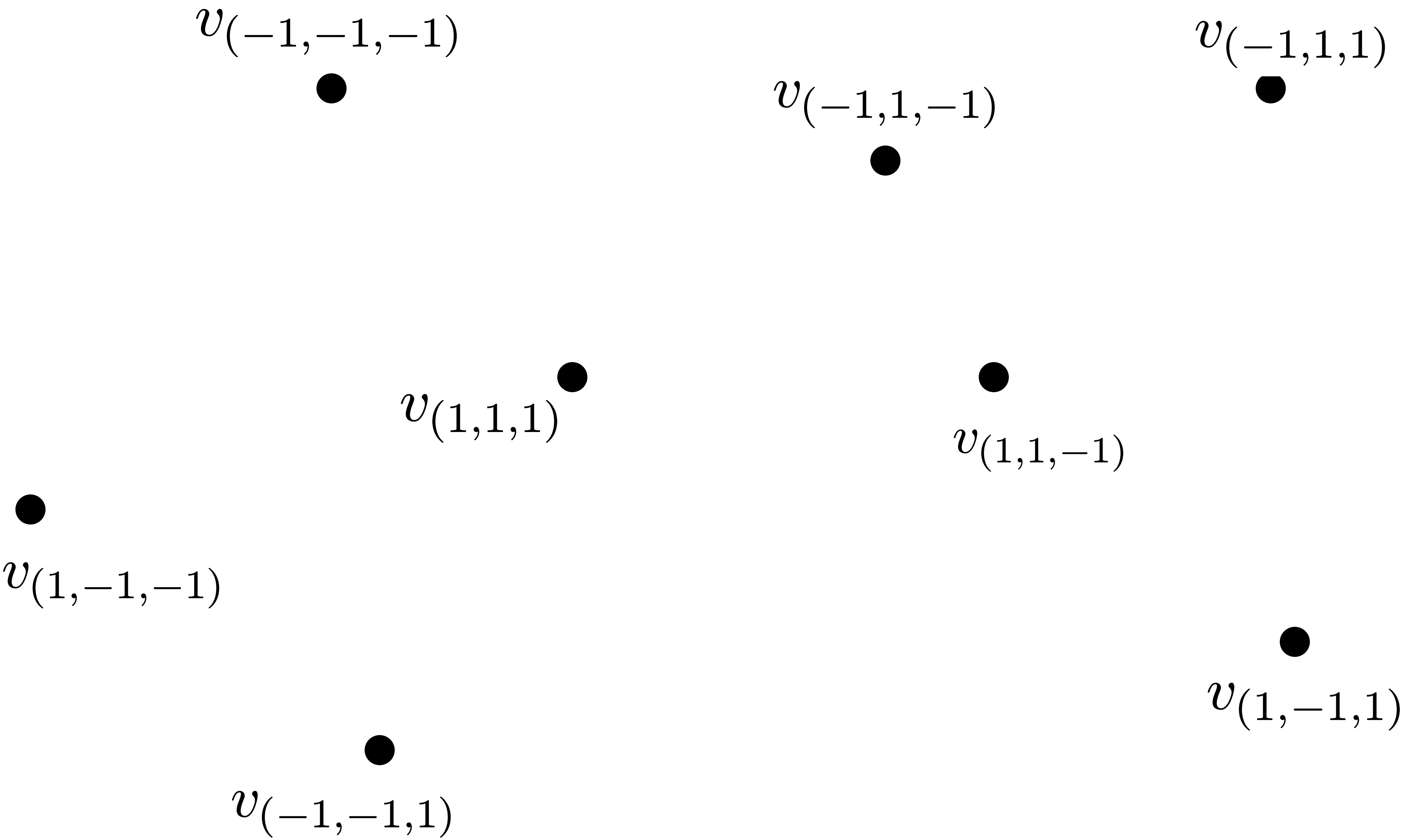
$$\frac{\sum diagonal}{2^n} \lesssim \left(\frac{\sum edge^p}{2^n}\right)^{1/p}$$

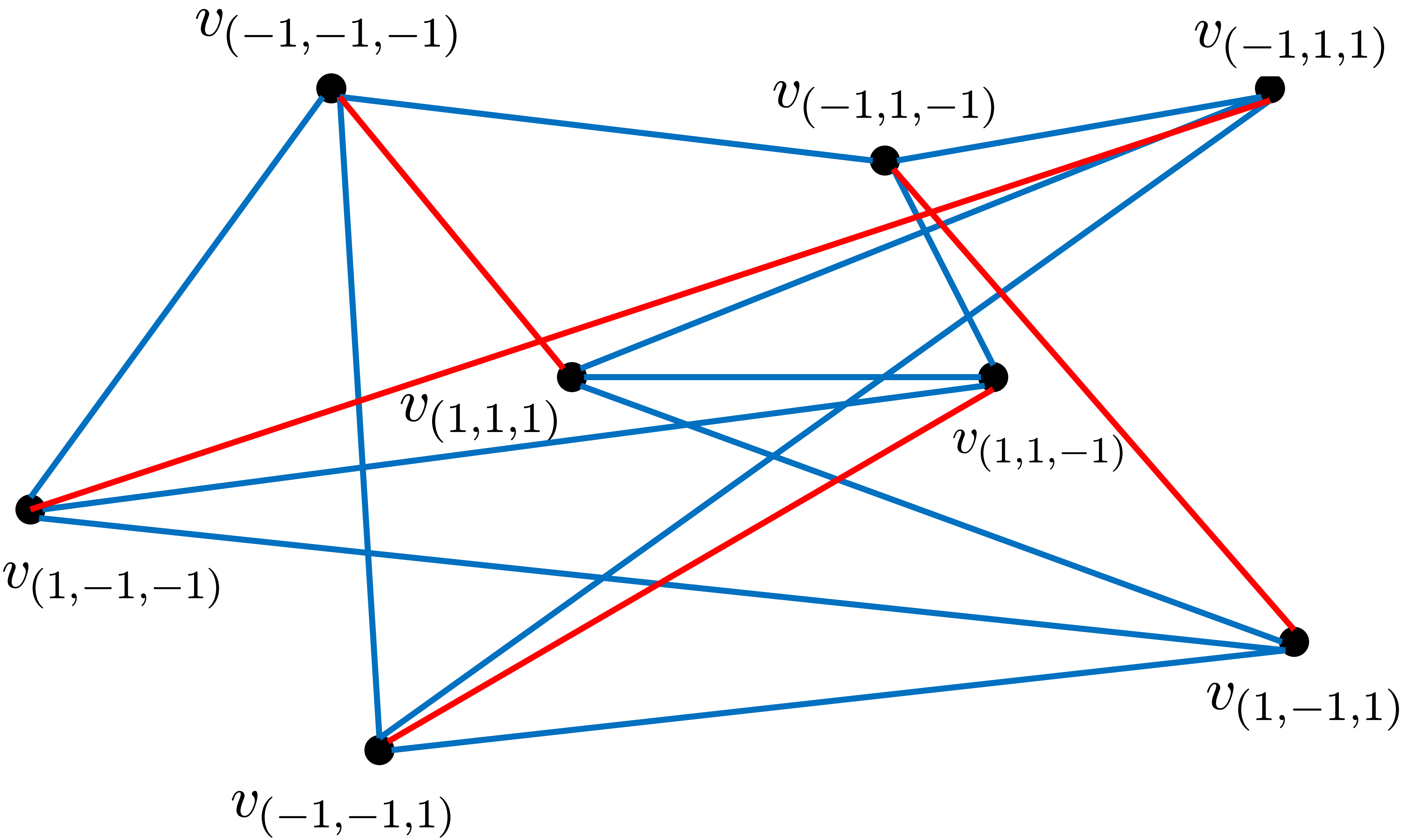


$$\frac{\sum diagonal}{2^n} \lesssim \left(\frac{\sum edge^p}{2^n} \right)^{1/p}$$

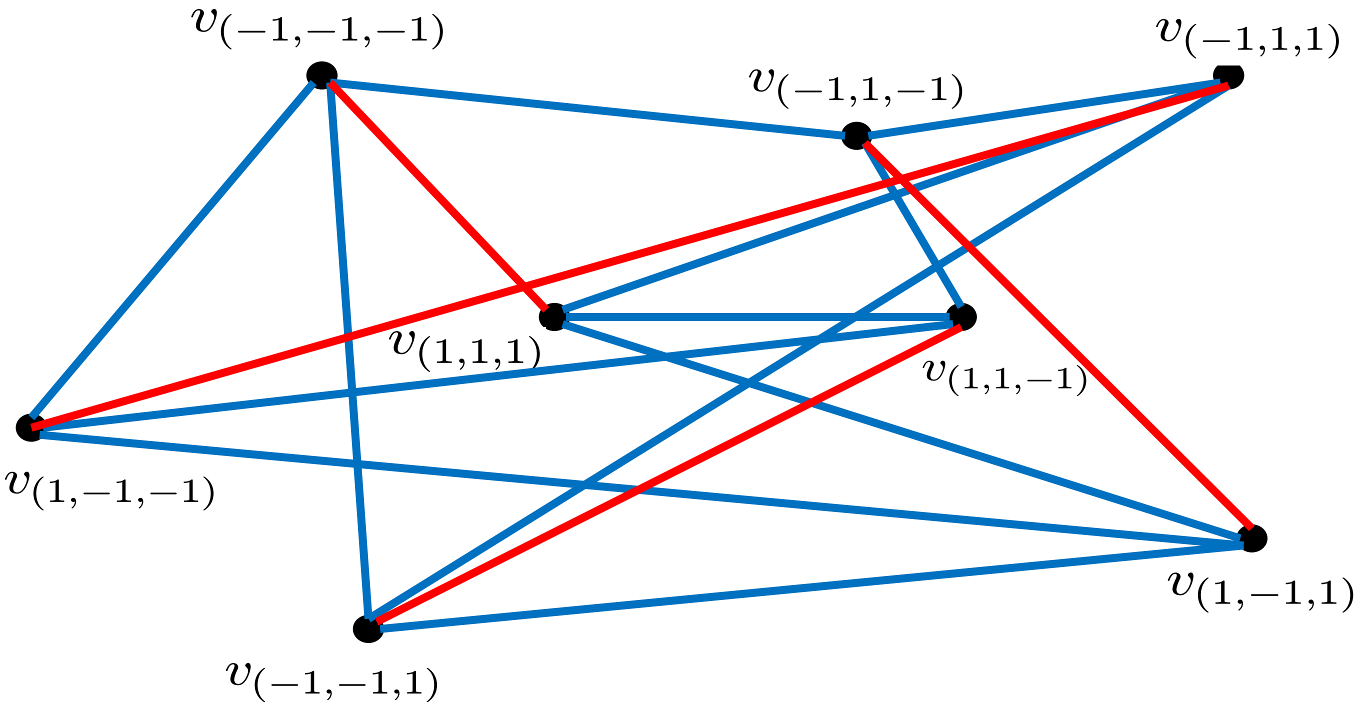
But now, we are given an arbitrary set of 2^n vectors $v_1, \ldots, v_{2^n} \in X$




$v_{(-1,-1,-1)}$
$v_{(-1,1,1)}$
$v_{(-1,1,-1)}$
$v_{(1,1,1)}$
$v_{(1,1,-1)}$
$v_{(1,-1,-1)}$
$v_{(1,-1,1)}$
$v_{(-1,-1,1)}$



$v_{(-1,-1,-1)}$
$v_{(-1,1,1)}$
$v_{(-1,1,-1)}$
$v_{(1,1,1)}$
$v_{(1,1,-1)}$
$v_{(1,-1,-1)}$
$v_{(1,-1,1)}$
$v_{(-1,-1,1)}$



$$
\begin{aligned}
&8\left(\frac{\|v_{(-1,-1,-1)}-v_{(1,1,1)}\|+\|v_{(-1,-1,1)}-v_{(1,1,-1)}\|+\|v_{(-1,1,1)}-v_{(1,-1,-1)}\|+\|v_{(-1,1,-1)}-v_{(1,-1,1)}\|}{8}\right)^p\\
\lesssim\quad &\|v_{(-1,-1,-1)}-v_{(1,-1,-1)}\|^p+\|v_{(-1,-1,-1)}-v_{(-1,1,-1)}\|^p+\|v_{(-1,-1,-1)}-v_{(-1,-1,1)}\|^p+\|v_{(1,1,-1)}-v_{(1,1,1)}\|^p\\
+\quad &\|v_{(-1,1,1)}-v_{(1,1,1)}\|^p+\|v_{(-1,1,1)}-v_{(-1,-1,1)}\|^p+\|v_{(-1,1,-1)}-v_{(1,1,-1)}\|^p+\|v_{(-1,1,1)}-v_{(-1,1,-1)}\|^p\\
+\quad &\|v_{(1,1,-1)}-v_{(1,-1,-1)}\|^p+\|v_{(1,1,1)}-v_{(1,-1,1)}\|^p+\|v_{(1,-1,1)}-v_{(-1,-1,1)}\|^p+\|v_{(-1,1,-1)}-v_{(-1,1,1)}\|^p
\end{aligned}
$$

Pisier's approach (1987)

Every $f : \{-1,1\}^n \to X$ has a Fourier expansion

$$f(x) = \sum_{A \subseteq \{1,\ldots,n\}} \widehat{f}(A) W_A(x)$$

where

$$W_A(x) = \prod_{i \in A} x_i$$

$$\widehat{f}(A) = \mathbb{E}\left[f(\epsilon) W_A(\epsilon)\right] = \frac{1}{2^n} \sum_{\epsilon_1,\ldots,\epsilon_n \in \{-1,1\}} f(\epsilon) \prod_{i \in A} \epsilon_i$$

Vector-valued heat semigroup

For $t \in \mathbb{R}$ and $f : \{-1,1\}^n \to X$

$$T_t f = \sum_{A \subseteq \{1,\ldots,n\}} e^{-t|A|} \widehat{f}(A) W_A$$

Then for t>0, $$\|f\|_1 \geq \|T_t f\|_1 \geq e^{-nt} \|f\|_1$$

where

$$\|f\|_1 = \|f\|_{L_1(\{-1,1\}^n, X)} = \mathbb{E}\left[\|f(\epsilon)\|\right]$$

Duality

Fix s>0 that will be determined later.

There exists $g^* : \{-1,1\}^n \to X^*$ such that

$1 \geq \|g^*(\epsilon)\|$ for all $\epsilon \in \{-1,1\}^n$

and

$$\begin{aligned} \|T_s f\|_1 &= \mathbb{E}\left[g^*(\epsilon)(T_s f(\epsilon))\right] \\ &= \sum_{A \subseteq \{1,\ldots,n\}} e^{-s|A|}\widehat{g^*}(A)\left(\widehat{f}(A)\right) \end{aligned}$$

Linearization via interpolation between two hypercubes

Define for every t>0,

$$g_t^* : \{-1,1\}^n \times \{-1,1\}^n \to X^*$$

by

$$g_t^*(\epsilon,\epsilon') = g^*(e^{-t}\epsilon + (1-e^{-t})\epsilon')$$

$$= \sum_{A\subseteq\{1,\ldots,n\}} \widehat{g^*}(A) \prod_{i\in A}(e^{-t}\epsilon_i + (1-e^{-t})\epsilon_i')$$

$$g_t^*(\epsilon, \epsilon') = \sum_{i=1}^{n} \epsilon_i' \sum_{\substack{A \subseteq \{1,\dots,n\} \\ i \in A}} (1 - e^{-t}) e^{-(|A|-1)t} \widehat{g^*}(A) W_{A \setminus \{i\}}(\epsilon) + \Phi_t^*(\epsilon, \epsilon')$$

where for all $h_1, \dots, h_n : \{-1, 1\}^n \to X$,

$$\mathbb{E}_{\epsilon, \epsilon'} \left[\Phi_t^*(\epsilon, \epsilon') \left(\sum_{i=1}^{n} \epsilon_i' h_i(\epsilon) \right) \right] = 0$$

For $f : \{-1,1\}^n \to X$ and $i \in \{1,\ldots,n\}$ define $D_i f : \{-1,1\}^n \to X$ by

$$
\begin{aligned}
& D_i f(\epsilon) \\
= \; & f(\epsilon_1,\ldots,\epsilon_{i-1},1,\epsilon_{i+1},\ldots,\epsilon_n) - f(\epsilon_1,\ldots,\epsilon_{i-1},-1,\epsilon_{i+1},\ldots,\epsilon_n) \\
= \; & 2 \sum_{\substack{A\subseteq\{1,\ldots,n\} \\ i\in A}} \widehat{f}(A) W_{A\setminus\{i\}}(\epsilon)
\end{aligned}
$$

$$\begin{aligned}
&\mathbb{E}_{\epsilon,\epsilon'}\left[g_t^*(\epsilon,\epsilon')\left(\sum_{i=1}^{n}\epsilon_i' D_i f(\epsilon)\right)\right] \\
&= \quad 2(e^t-1)\sum_{i=1}^{n}\sum_{\substack{A\subseteq\{1,\ldots,n\}\\ i\in A}} e^{-|A|t}\widehat{g^*}(A)\left(\widehat{f}(A)\right) \\
&= \quad 2(e^t-1)\sum_{A\subseteq\{1,\ldots,n\}} |A|e^{-|A|t}\widehat{g^*}(A)\left(\widehat{f}(A)\right)
\end{aligned}$$

So, for all t>0,

$$\left(\max_{\epsilon,\epsilon' \in \{-1,1\}^n} \|g_t^*(\epsilon,\epsilon')\|\right) \mathbb{E}_{\epsilon,\epsilon'} \left[\left\|\sum_{i=1}^{n} \epsilon_i' D_i f(\epsilon)\right\|\right]$$

$$\geq 2(e^t - 1) \sum_{A \subseteq \{1,\ldots,n\}} |A| e^{-|A|t} \widehat{g^*}(A) \left(\widehat{f}(A)\right)$$

Since $1 \geq \|g^*(\epsilon)\|$ for every $\epsilon \in \{-1,1\}^n$ also $1 \geq \|g_t^*(\epsilon,\epsilon')\|$ for all $\epsilon, \epsilon' \in \{-1,1\}^n$.

This follows from convexity and tensorization.

So, we proved that for all t>0,

$$\frac{1}{2(e^t-1)}\mathbb{E}_{\epsilon,\epsilon'}\left[\left\|\sum_{i=1}^{n}\epsilon_i' D_i f(\epsilon)\right\|\right]$$

$$\geq \sum_{A\subseteq\{1,\ldots,n\}} |A|e^{-|A|t}\widehat{g^*}(A)\left(\widehat{f}(A)\right)$$

Integration of this inequality:

$$\left(\int_s^\infty \frac{dt}{2(e^t-1)}\right)\mathbb{E}_{\epsilon,\epsilon'}\left[\left\|\sum_{i=1}^n \epsilon_i' D_i f(\epsilon)\right\|\right]$$

$$\geq \sum_{A\subseteq\{1,\ldots,n\}} e^{-s|A|}\widehat{g^*}(A)\left(\widehat{f}(A)\right)$$

$$= \|T_s f\|_1$$

$$\geq e^{-sn}\|f\|_1$$

Pisier's inequality

So,

$$\frac{e^{sn}}{2}\log\left(\frac{e^s}{e^s-1}\right)\mathbb{E}_{\epsilon,\epsilon'}\left[\left\|\sum_{i=1}^{n}\epsilon_i' D_i f(\epsilon)\right\|\right] \geq \|f\|_1$$

Optimal choice of s is $s \approx \frac{\log\log n}{n\log n}$

$$\mathbb{E}_{\epsilon}\left[\|f(\epsilon)-f(-\epsilon)\|\right]$$

$$\lesssim \quad (\log n)\mathbb{E}_{\epsilon,\epsilon'}\left[\left\|\sum_{i=1}^{n}\epsilon_i' D_i f(\epsilon)\right\|\right]$$

If X has Rademacher type p then for every fixed $\epsilon \in \{-1,1\}^n$,

$$\mathbb{E}_{\epsilon'}\left[\left\|\sum_{i=1}^{n}\epsilon_i' D_i f(\epsilon)\right\|\right] \lesssim \left(\sum_{i=1}^{n}\|D_i f(\epsilon)\|^p\right)^{1/p}$$

So, by Pisier's inequality

$$\mathbb{E}_\epsilon\|f(\epsilon)-f(-\epsilon)\|$$

$$\lesssim \quad (\log n)\left(\sum_{i=1}^{n}\mathbb{E}_\epsilon\left[\|f(\epsilon)-f(\epsilon_1,\ldots,-\epsilon_i,\ldots,\epsilon_n)\|^p\right]\right)^{\frac{1}{p}}$$

Not quite the Bourgain-Milman-Wolfson theorem because of the logarithmic term, but also we got power p instead of power p-ε.

One can manipulate this inequality to get rid of the logarithmic term at the cost of changing the power to p-ε.

Improve Pisier's inequality?

- Talagrand (1993): the logarithmic term in Pisier's inequality is needed in general.
- N.-Schechtman (2002): the logarithmic term can be replaced by a constant depending on the geometry of X but not on the dimension n if X is a UMD Banach space (martingale differences are unconditional; includes L_p spaces, p>1). Thus for a wide class of spaces we know that their Rademacher type is the same as their type as metric spaces.

Completion of the Ribe program for Rademacher type

- Open: can the logarithmic term in Pisier's inequality be removed if X has nontrivial Rademacher type?
- Mendel-N. (2007): a different definition of type p of metric spaces (scaled Enflo type p) which is equivalent to Rademacher type p.

Unique obstruction

When does a metric space X have type p>1?

$$\frac{\sum diagonal}{2^n} \lesssim \left(\frac{\sum edge^p}{2^n} \right)^{1/p}$$

Obvious obstruction: there exists K>0 such that for all n there is $f:\{-1,1\}^n \to X$ satisfying

$$K\|\epsilon-\epsilon'\|_1 \geq d(f(\epsilon),f(\epsilon')) \geq \|\epsilon-\epsilon'\|_1$$

$$\frac{\sum diagonal}{2^n} \gtrsim n$$

$$\left(\frac{\sum edge^p}{2^n}\right)^{1/p} \lesssim n^{1/p}$$

Theorem (Bourgain-Milman-Wolfson, Pisier):

The only obstruction for a metric space not to have nontrivial type is that it contains hypercubes of arbitrarily high dimension with bi-Lipschitz distortion O(1).

Further impact of the Ribe program: metric theories motivated by linear insights

The (often surprising) successes of the Ribe program motivate a powerful variant of this program: a variety of useful theorems on general metric spaces that are analogues of important results in Banach space theory.

The ultrametric skeleton theorem (Mendel-N., 2011)

For every $\varepsilon>0$ there exists $c_\varepsilon > 0$ such that if (X,d) is a compact metric space and μ is a Borel probability measure on X then X has a compact subset S that embeds into an ultrametric space with distortion $O(1/\varepsilon)$ and there is a Borel probability measure ν supported on S such that

$$\forall x \in X,\ \forall r \geq 0, \quad \left(\mu(B(x, c_\epsilon r))\right)^{1-\epsilon} \geq \nu(B(x,r))$$

- Sharp solution of Bourgain-Figiel-Milman nonlinear Dvoretzky problem (1986). Bartal-Linial-Mendel-N. (2002), Mendel-N. (2006), N.-Tao (2010).
- Solution of Tao’s nonlinear Dvoretzky problem for Hausdorff dimension (Mendel-N., 2006).
- Solution of Urbanski’s problem (Keleti-Mathe-Zindulka, 2012).
- Talagrand’s majorizing measures theorem.
- Best known lower bounds for the randomized k-server problem.
- Only known way to construct approximate distance oracles with constant query time, other proximity data structures.

- Bourgain's super-reflexivity theorem (1986).
- Markov convexity (Lee-Peres-N., 2006, and Mendel-N., 2008), geometry of trees, Lipschitz quotients.
- Metric cotype (Mendel-N., 2006). Metric Maurey-Pisier theorem, coarse embeddings, metric dichotomies.
- Markov type (Ball, 1990), N.-Peres-Schramm-Sheffield (2005), Johnson-Lindenstrauss Lipschitz extension problem (1983). Applications to group theory (Austin-N.-Peres, 2008, N.-Peres, 2009+2010).
- Bi-Lipschitz embedding theory: Bourgain (1985), Arora-Lee-N. (2007), Sparsest Cut Problem (Linial-London-Rabinovich, Goemans-Linial).
- Nonlinear spectral gaps and construction of super-expanders: V. Lafforgue (2007) and Mendel-N. (2009). Applications to coarse geometry.