Part III — Metric Embeddings

Based on lectures by A. Zsák
Notes taken by Daniel Ng

Lent 2020

Functional Analysis and Part II Probability and Measure are essential

Definitions, basic examples and motivations. Frechét embeddings, Aharoni's theorem (ℓ_∞, c_0), Euclidean distortion, Bourgain's embedding theorem (ℓ_2, L_2). Obstructions to embeddings. Poincaré's inequalities (L_1, L_2). Dimension reduction in L_2 (Johnson-Lindenstrauss Lemma). Lack of dimension reduction in L_1. Local theory of Banach spaces, Ribe programme. Bourgain's characterisation of super-reflexivity, metric type and cotype and/or metric Dvoretzky's Theorem). Coarse embeddings of ℓ_2 into Banach spaces, coarse embeddings into uniformly convex/uniformly smooth Banach spaces.

Books: Ostrowski's *Metric Embeddings*, Matousek's *Lectures in discrete geometry* (Ch15 - extended online notes), *Lectures in metric embeddings* (available online). Assaf Naor's survey article on the Ribe programme.

Related Part III courses: discrete analysis of Fourier series, some combinatorics.

Contents

1 Basic Definitions, Examples and Motivations

Definition. A *metric space* is a set M with a metric $d\colon M \times M \to \mathbb{R}$ such that (i) $d(x, y) \geq 0$ for all x, y, $d(x, x) = 0$ for all x, (ii) $d(x, y) = d(y, x)$ (symmetry), (iii) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality), (iv) $d(x, y) = 0 \implies x = y$. If d satisfies (i),(ii) and (iii) then it's a *semimetric*.

Example. Graph with the graph distance. A *graph* is a pair (V, E) where V is a set and $E \subset V^{(2)} = \{e \subset V : |e| = 2\}$. Elements of V are the *vertices* of G and elements of E are the *edges* of G. A *walk* in G is a sequence $x_0, x_1, ..., x_n$ of vertices such that $x_{i-1}x_i \in E$ for all $1 \leq i \leq n$. [Given $e = \{x, y\} \in E$, x, y are the *endvertices* of e, write $e = xy = yx$. We also write $x \sim y$]. The length of the walk is n. This is called a *walk from x_0 to x_n*. If $x_i \neq x_j$ whenever $1 < j - i < n$, this walk is called a *path* from x_0 to x_n. Say G is *connected* if for all $x, y \in V$ there exists a walk (or a path) in G from x to y. The *graph distance* is defined as $d_G(x, y) =$ the length of a shortest path in G from x to y. Some standard graphs: K_n is the *complete graph on n vertices*, all $\binom{n}{2}$ edges are present. Here

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y, \\ 0 & \text{if } x = y. \end{cases}$$

P_n is the *path of length n*, with $n + 1$ vertices $x_0, x_1, ..., x_n$, and $E = \{x_{i-1}x_i : 1 \leq i \leq n\}$. As a metric space, this is $\{0, 1, ..., n\}$ with $d(x, y) = |x - y|$. C_n is the *cycle of length n*. $V = \{x_1, ..., x_n\}$ and $E = \{x_i x_{i+1} : 1 \leq i < n\} \cup \{x_1 x_n\}$. B_n is the *rooted binary tree of depth n*. And finally, H_n is the *Hamming cube* $V = \{0, 1\}^n$, $x \sim y$ iff there exists exactly one coordinate i such that $x_i \neq y_i$. Then $d(x, y) = |\{i : x_i \neq y_i\}|$.

Example. Groups with the word metric. Let G be a group generated by some subset S. We always assume that $e \notin S$ and S is symmetric: $\forall x \in S$, $x^{-1} \in S$. The *word metric* is defined to be $d(x, y) = \min\{n : \exists a_1, ..., a_n \in S \, s.t. \, x^{-1}y = a_1...a_n\}$. The *Cayley Graph* $C(G, S)$ has vertex set G and $x \sim y$ iff $x^{-1}y \in S$. The graph distance on G is d.

Example. Cut semimetrics. A *cut* on a set M is a partition of M into S and $M \setminus S$. The corresponding *cut semimetric* is

$$d_S(x, y) = \begin{cases} 0 & x, y \text{ are in the same part} \\ 1 & \text{otherwise.} \end{cases}$$

Definition. A *normed space* is a real or complex vector space V with a *norm* on V, i.e. a function $\|\cdot\|\colon V \to \mathbb{R}$ such that (i) $\|x\| \geq 0$ for all $x \in V$, (ii) $\|\lambda x\| = |\lambda|\|x\|$ for all λ scalars and $x \in V$, (iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$, (iv) $\|x\| = 0 \implies x = 0$. Then $d(x, y) = \|x - y\|$ is a metric on V. If V is complete then it is called a *Banach space*. If $\|\cdot\|$ satisfies (i),(ii) and (iii) then it is called a *seminorm*.

Example. Classical sequence spaces.

- $\ell_p^n = (\mathbb{R}^n, \|\cdot\|_p)$ for $1 \leq p \leq +\infty$, with $\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$. Here e_i is the standard ith basis vector. If $p = \infty$ the norm is $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$.

- $\ell_p = \{(x_i)_{i=1}^{\infty} : \sum_{i=1}^{n} |x_i|^p\}$ for $1 \le p < +\infty$, with $\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p}$. $\ell_\infty = \left\{(x_i)_{i=1}^{\infty} : \sup_{i\ge 1} |x_i| < \infty\right\}$, with $\|x\|_\infty = \sup_{i\ge 1} |x_i|$. More generally, for a set S, $\ell_\infty(S) = \{x \colon S \to \mathbb{R} : x \text{ is bounded}\}$. The norm is $\|x\|_\infty = \sup_{s\in S} |x(s)|$. Note $c_0 = (x_i)_{i=1}^{\infty} : x_i \to 0\}$ is a closed subspace of ℓ_∞.

Example. Classical function spaces. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space.

- For $1 \le p < \infty$, $L_p(\mu) = \{f \colon \Omega \to \mathbb{R} : f \text{ measurable}, \int_\Omega |f|^p \, d\mu < \infty\}$ equipped with $\|f\|_p = \left(\int_\Omega |f|^p \, d\mu\right)^{1/p}$.
- For $p = \infty$, $L_\infty(\mu) = \{f \colon \Omega \to \mathbb{R} : f \text{ measurable}, \exists N \in \mathcal{F}, \mu(N) = 0, f \text{ bounded on } \Omega \backslash N\}$, equipped with $\|f\|_\infty = \operatorname{ess\,sup}(f) = \inf\{\sup_{\Omega\backslash N} |f| : N \in \mathcal{F}, \mu(N) = 0\}$.
- In the case $\Omega = [0,1]$, μ= Lebesgue measure, we write L_p for $L_p(\mu)$.
- For compact space K, $C(K) = \{f \colon K \to \mathbb{R} : f \text{ cts}\}$ is a closed subspace of $\ell_\infty(K)$, e.g. $C([0,1])$.

Example. Hilbert Space. An *inner product space* (IPS) is a vector space V equipped with an inner product $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ (symmetric bilinear, positive definite). Then V becomes a normed space with $\|x\| = \langle x, x \rangle^{1/2}$. If V is complete wrt $\|\cdot\|$, then it's called a *Hilbert space*.

Definition. Let $f \colon M \to N$ be a map between metric spaces. Then f is *isometric* or an *isometric embedding* if $d(f(x), f(y)) = d(x,y)$ for all $x, y \in M$. We say f is a *bilipschitz embedding* if $\exists a, b > 0$ such that

$$a d(x,y) \le d(f(x), f(y)) \le b d(x,y) \qquad \forall x, y \in M. \tag{1}$$

The *distortion* of f is $\operatorname{dist}(f) = \min\{\frac{b}{a} : (1) \text{ holds for } f\}$.

Remark.

(i) If $a = b$, then f is a *scaled isometric embedding.*

(ii) Definition makes sense for semimetrics.

(iii) If (1) holds, then f is Lipschitz with Lipschitz constant $\operatorname{Lip}(f) \le b$, where

$$\operatorname{Lip}(f) = \sup\left\{\frac{d(f(x), f(y))}{d(x,y)} : x, y \in M, x \ne y\right\}.$$

Also f is injective (because of the LH inequality) and $f^{-1} \colon f(M) \to M$ is Lipschitz, with $\operatorname{Lip}(f^{-1}) \le \frac{1}{a}$. Then $\operatorname{dist}(f) = \operatorname{Lip}(f)\operatorname{Lip}(f^{-1})$.

Recall, if $T \colon X \to Y$ is a linear map between normed spaces, then T is continuous iff T is bounded ($\exists C > 0, \|Tx\| \le C\|x\|$ for all $x \in X$). The smallest C is $\|T\|$ iff T is Lipschitz, $\|T\| = \operatorname{Lip}(T)$. T is an isomorphism if T is a bijection, both T and T^{-1} are bounded. T is an *isometric embedding* or *into isomorphism* if T is an isomorphism between X and $T(X)$, iff T is bilipschitz. Then $\operatorname{dist}(T) = \|T\|\left\|T^{-1}\right\|$. T is an *isometric isomorphism embedding* if $\|Tx\| = \|x\|$ for all $x \in X$.

Notation. Write $X \hookrightarrow_C Y$ if there exists an isomorphism embedding $T\colon X \to Y$ with $\|T\| \|T^{-1}\| \leq C$. We say X *C-embeds* into Y. So $X \hookrightarrow_1 Y$ iff there exists an isometric isomorphic embedding $X \to Y$. $X \sim Y$ means X, Y are isomorphic. $X \cong Y$ means X, Y are isometrically isomorphic.

Example. (i) $\ell_p^n \hookrightarrow_1 \ell_p$.

(ii) $\ell_p \hookrightarrow_1 L_p = L_p([0,1], \lambda = \text{Leb})$. *proof:* Fix pairwise disjoint measurable sets $(A_i)_{i=1}^\infty$ each of positive measure. For $1 \leq p < \infty$, consider

$$(x_i)_{i=1}^\infty \mapsto \sum_{i=1}^\infty x_i \mathbb{1}_{A_i} \lambda(A_i)^{-1/p},$$

and for $p = \infty$, consider $(x_i)_{i=1}^\infty \mapsto \sum_{i=1}^\infty x_i \mathbb{1}_{A_i}$.

Fact. If (Ω, μ) is a measure space, $X \subset L_p(\Omega, \mu)$ separable, then $X \hookrightarrow_1 L_p$.

Notation. For a normed space X, let $B_X = \{x \in X : \|x\| \leq 1\}$ be the *closed unit ball*, and $S_X = \{x \in X : \|x\| = 1\}$, the *unit sphere* of X.

Proposition 1.1. For all $n \in \mathbb{N}$, $\ell_2^n \hookrightarrow_1 L_p$ for any $1 \leq p \leq \infty$.

Proof. **Case $1 \leq p < \infty$.** Let $B = B_{\ell_2^n}$, $\mu =$ Lebesgue measure on B, $S = S_{\ell_2^n}$. Since μ is rotation invariant, the value of $\int_B |\langle x, \omega\rangle|^p \, d\mu(\omega)$ is the same for all $x \in S$. Call this α. Define $T\colon \ell_2^n \to L_p(B, \mu)$ by $(Tx)(\omega) = \langle x, \omega \rangle \alpha^{-1/p}$. Then T is linear and $\|Tx\|_p^p = \int_B |\langle x, \omega\rangle|^p \alpha \, d\mu(\omega) = \|x\|_2^p$ for all $x \in \ell_2^n$. To finish, use the fact above to embed $L_p(B, \mu) \hookrightarrow_1 L_p$.

Case $p = \infty$. This follows from the next result and example above. □

Definition. Let X be a normed space. The *dual space* X^* of X is $X^* = \mathcal{B}(X, \mathbb{R}) = \{f\colon X \to \mathbb{R} : f \text{ linear bounded}\}$. The operator norm is $\|f\| = \sup\{|f(x)| : x \in B_X\}$. By the Hahn-Banach theorem, $\forall x \in X$, $\exists f \in S_{X^*}$ such that $f(x) = \|x\|$. So $\|x\| = \max\{g(x) : g \in S_{X^*}\}$.

Proposition 1.2. Let X be a separable normed space. Then $X \hookrightarrow_1 \ell_\infty$.

Proof. Let $(x_n)_{n=1}^\infty$ be dense in X. Then for all $n \in \mathbb{N}$, choose $f_n \in S_{X^*}$, $f_n(x_n) = \|x_n\|$ by Hahn-Banach. Define $T\colon X \to \ell_\infty$ by $Tx = (f_n(x))_{n=1}^\infty$. Given $x \in X$, $|f_n(x)| \leq \|f_n\| \|x\| = \|x\|$ for all n. So T is well-defined. T is linear and T is bounded with $\|T\| \leq 1$. For all $n \in \mathbb{N}$, $\|Tx_n\| = \|x_n\|$. So T is isometric on a dense set, so by continuity T is isometric on the whole space X. □

Remark. For any normed space X there exists a set S such that $X \hookrightarrow_1 \ell_\infty(S)$, e.g. $S = S_{X^*}$.

Corollary 1.3. (Corollary to proposition 1.1) Let M be a finite metric space. If M embeds into L_2 with distortion $\leq D$, then M embeds into L_p for all $1 \leq p \leq \infty$ with distortion $\leq D$. i.e. L_2 is the hardest thing to embed into.

Proposition 1.4. If M is an n-element subset of $L_1(\Omega, \mu)$, then $M \hookrightarrow_1 \ell_1^N$, where $N = n!$.

Proof. Let $M = \{f_1, ..., f_n\}$. [Aside: $f_i \mapsto \int_\Omega f_i \, d\mu$ is an obvious $L_1(\Omega, \mu) \to \mathbb{R}$, but

$$\left| \int f_i - \int f_j \right| \le \int |f_i - f_j| = \|f_i - f_j\|_{L^1}$$

has equality if say $f_i \le f_j$ a.e.] There exists a partition $\Omega = \bigcup_{\pi \in S_n} \Omega_\pi$ of Ω, where $\Omega_\pi \subset \{\omega \in \Omega : f_{\pi(1)}(\omega) \le ... \le f_{\pi(n)}(\omega)\}$. Here we have used the finiteness of M. [Note that we have used the subset symbol. When two fs are equal for some ω, we can arbitrarily put it in just one of the Ω_πs.] Then

$$\|f_i - f_j\|_1 = \int_\Omega |f_i - f_j| \, d\mu = \sum_{\pi \in S_n} \int_{\Omega_\pi} |f_i - f_j| \, d\mu = \sum_{\pi \in S_n} \left| \int_{\Omega_\pi} f_i - \int_{\Omega_\pi} f_j \right|.$$

Define $T\colon M \to \ell_1^N$ by $Tf_i = \left(\int_{\Omega_\pi} f_i \, d\mu \right)_{\pi \in S_n}$. So above $= \|Tf_i - Tf_j\|_1$. □

Example (More examples)**.** (i) C_4 embeds bilipschitzly into ℓ_2^2 naturally, with distortion $\sqrt{2}$. It doesn't embed isometrically. In ℓ_2, $d(x, z) = d(x, y) + d(y, z)$ iff $y \in [x, z] = \{(1 - t)x + tz : 0 \le t \le 1\}$. It follows that ℓ_2 has the *unique midpoint property*: $\forall x, z \in \ell_2$ there is at most one point y (in fact exactly 1) such that $d(x, y) = d(y, z) = \frac{1}{2} d(x, z)$. C_4 does not have this property.

(ii) Any n-element set in a Hilbert space embeds isometrically into ℓ_2^{n-1}. Cannot do better in general. See example sheet. If we relax the condition to bilipschitz, then we can do much better. In fact, $\forall \epsilon 0, \exists c > 0$ any n-element set in Hilbert space embeds into ℓ_2^m where $m = c \log n$ with distortion $< 1 + \epsilon$. See later for proof.

Observe: If M is a finite metric space, N is a metric space and $|N| \ge |M|$, then M bilipschitzly embeds into N.

Definition. Given families $(M_\alpha)_{\alpha \in A}$ and $(N_\alpha)\alpha \in A$ of metric spaces, embeddings $f_\alpha\colon M_\alpha \to N_\alpha$, $\alpha \in A$, are called *uniformly bilipschitz* if $\sup_{\alpha \in A} \operatorname{dist}(f_\alpha) < \infty$.

The sparsest cut problem.

Let $G = (V, E)$ be a connected, finite graph. We are given two functions $C\colon E \to \mathbb{R}^+ = [0, \infty)$ (*capacity*) and $D\colon V \times V \to \mathbb{R}^+$ (*demand*). A *cut* of G is a partitioning $(S, V \setminus S)$ of V. The *capacity of* $(S, V \setminus S)$ is

$$C(S, V \setminus S) = \sum_{uv \in E, u \in S, v \notin S} C(uv).$$

The demand of the cut is

$$D(S, V \setminus S) = \sum_{u \in S, v \notin S} D(uv).$$

The *sparsity* of the cut is $C(S, V \setminus S)/D(S, V \setminus S)$ whenever $D(S, V \setminus S) \neq 0$. This is NP-hard. So we look at an equivalent problem: Minimise over all cuts with nonzero demand of the following quantity

$$\frac{\sum_{uv \in E} C(uv) d_S(u, v)}{\sum_{u, v \in V} D(u, v) d_S(u, v)}$$

where d_S is the cut semimetric. Note the denominator is twice $D(S, V \setminus S)$.

Let $\varphi^*(C, D)$ be this minimum. The idea is to minimise

$$\sum_{uv \in E} C(uv) d(u, v),$$

subject to d being a semimetric and $\sum_{u,v \in V} D(u, v) d(u, v) = 1$. This is now a linear programming problem with a linear normalisation condition. The property that d is a semimetric is just constraints with inequalities. There are fast algorithms to solve this.

Let $\varphi(C, D)$ be the minimum and d_{min} be a semimetric that achieves this minimum. Clearly $\varphi(C, D) \leq \varphi^*(C, D)$.

Lemma 1.5. Let (M, d) be a finite semimetric space. Then (M, d) embeds isometrically into L_1 iff d is a non-negative linear combination of cut semimetrics.

Proof. $(\Leftarrow)$ We assume there exists cuts $(S_i, M \setminus S_i)$ for $i = 1, ..., k$ and non-negative reals α_i, $i = 1, ..., k$, such that $d = \sum_{i=1}^k \alpha_i d_S$. Let $f_i \colon M \to \mathbb{R}$ be $f_i(x) = \alpha_i \mathbb{1}_{x \in S_i}$, and $f \colon M \to \ell_1^k$, $f(x) = (f_i(x))_{i=1}^k$. Then

$$\|f(x) - f(y)\|_1 = \sum_{i=1}^k |f_i(x) - f_i(y)| = \sum_{i=1}^k \alpha_i d_{S_i}(x, y) = d(x, y).$$

$(\Rightarrow)$ By proposition 4, there exists isometric embedding $f \colon M \to \ell_1^k$, some $k \in \mathbb{N}$. Enumerate $\{f(x)_i : x \in M\}$ as $\beta_{i1} < \beta_{i2} < ... < \beta_{im_i}$. Let $S_{ij} = \{x : f(x)_i \leq \beta_{ij}\}$, for $1 \leq i \leq k$, $1 \leq j \leq m_i$. Fix $x, y \in M$, and fix $1 \leq i \leq k$. Suppose $f(x)_i = \beta_{ij_1} \leq f(y)_i = \beta_{ij_2}$. $x \in S_{ij}$ for $j \leq j_1$, $y \leq S_{ij}$ for $j \leq j_2$. If we look at the sum

$$\begin{aligned} \sum_{j=1}^{m_i - 1} (\beta_{i,j} - \beta_{i,j-1}) d_{S_{ij}}(x, y) &= \sum_{j=j_1+1}^{j_2} (\beta_{i,j} - \beta_{i,j-1}) \\ &= \beta_{i,j_2} - \beta_{i,j_1} \\ &= f(y)_i - f(x)_i = |f(x)_i - f(y)_i|. \end{aligned}$$

Sum over i:

$$\sum_{i=1}^k \sum_{j=1}^{m_i - 1} (\beta_{i,j} - \beta_{i,j-1}) d_{S_{ij}}(x, y) = \sum_{i=1}^k |f(x)_i - f(y)_i| = \|f(x) - f(y)\|_1 = d(x, y),$$

so we have written d as a sum of cut semimetrics. □

Theorem 1.6. Assume (V, d_{min}) embeds into L_1 with distortion at most K, then $K^{-1} \varphi^*(C, D) \leq \varphi(C, D) \leq \varphi^*(C, D)$.

Proof. Let $f \colon (V_1, d_{min}) \to L_1$ be an embedding with distortion at most K. Let $d(x, y) = \|f(x) - f(y)\|_1$. Since $\mathrm{dist}(f) \leq K$, there exists $a > 0$ such that $a d_{min}(x, y) \leq d(x, y) \leq K a d_{min}(x, y)$ for all $x, y \in V$. By lemma 1.5, there exists cuts $(S_i, V \setminus S_i)$, $1 \leq i \leq k$ and constants $\alpha_i \geq 0$, $i = 1, ..., k$ such that $d = \sum_{i=1}^k \alpha_i d_{S_i}$. Then

$$\varphi(C, D) = \frac{\sum_{uv \in E} C(uv) d_{min}(u, v)}{\sum_{u,v \in V} D(u, v) d_{min}(u, v)} \geq \frac{1}{K} \frac{\sum_{uv \in E} C(uv) d(u, v)}{\sum_{u,v \in V} D(u, v) d(u, v)} = \frac{1}{K} \frac{\sum_{i=1}^k \gamma_i}{\sum_{i=1}^k \delta_i},$$

where $\gamma_i = \alpha_i \sum_{uv} C(uv) d_{S_i}(u,v)$ and $\delta_i = \alpha_i \sum_{u,v \in V} D(uv) d_{S_i}(u,v)$. Let $I = \{i : \delta_i > 0\}$. The above becomes

$$\geq \frac{1}{K} \frac{\sum_{i \in I} (\gamma_i/\delta_i)\delta_i}{\sum_{i \in I} \delta_i} \geq \frac{1}{K} \min_{i \in I} \frac{\gamma_i}{\delta_i} \geq \frac{1}{K} \varphi^*(C, D).$$

□

Definition. Let $f \colon M \to N$ be a map between metric spaces. Assume there exists increasing functions $\rho_1, \rho_2 \colon \mathbb{R}^+ \to \mathbb{R}^+$ ($s \leq t \implies \rho_1(s) \leq \rho_1(t)$) such that

$$\rho_1(d(x,y)) \leq d(f(x), f(y)) \leq \rho_2(d(x,y)) \qquad \forall x, y \in M. \tag{2}$$

We say f is a *coarse embedding* if in addition to (2), $\rho_1(t) \to \infty$ as $t \to \infty$.

Example. Let $f \colon \mathbb{R} \times [0,1] \to \mathbb{R}$ by $f(x,t) = x$. This is a coarse embedding with $\rho_1(t) = \max(0, t-1)$ and $\rho_2(t) = t$.

Definition. We say f is a *uniform embedding* if in addition to (2), $\rho_2(t) \to 0$ as $t \to 0^+$ and $\rho_1(t) > 0$ for all $t > 0$. Equivalently this says f is uniformly continuous, injective; $f^{-1} \colon f(M) \to M$ is uniformly continuous.

Proposition 1.7. For all $1 < q < \infty$ there exists a map $T \colon L_1(\Omega, \mu) \to L_q(\Omega \times \mathbb{R}, \nu)$ which is simultaneously a uniform and coarse embedding. (Here $\nu = \mu \otimes \lambda$ is the product measure of μ and the Lebesgue measure λ.)

Proof. Define T as follows. For $f \in L_1(\Omega, \mu)$,

$$Tf(\omega, t) = \begin{cases} +1 & \text{if } 0 < t \leq f(\omega), \\ -1 & \text{if } f(\omega) \leq t \leq 0 \\ 0 & \text{else.} \end{cases}$$

Note that $Tf \in L_\infty(\Omega \times \mathbb{R})$. For $f, g \in L_1(\Omega, \mu)$,

$$|Tf(\omega,t) - Tg(\omega,t)| = \begin{cases} 1 & \text{if } g(\omega) \leq t \leq f(\omega), \\ 1 & \text{if } f(\omega) \leq t \leq g(\omega). \end{cases}$$

So

$$\int_\Omega \int_\mathbb{R} |Tf(\omega,t) - Tg(\omega,t)|^q \, dt \, d\mu(\omega) = \int_\Omega |f(\omega) - g(\omega)| \, d\mu(\omega) = \|f - g\|_1 .$$

So $\|Tf - Tg\|_q^q = \|f - g\|_1$. This shows that $Tf \in L_q(\Omega \times \mathbb{R})$.

If $\rho_1(t) = \rho_2(t) = t^{1/q}$, then $\rho_1(\|f-g\|_1) = \|Tf - Tg\|_q = \rho_2(\|f - g\|_1)$. And $\rho_1(t) \to \infty$ as $t \to \infty$, and $\rho_2(t) \to 0$ as $t \to 0^+$ and $\rho_1(t) > 0$ for all $t > 0$. □

Proposition 1.8. For $1 \leq p < q < \infty$ there exists $T \colon L_p(\Omega, \mu) \to L_q(\Omega \times \mathbb{R}, \nu; \mathbb{C}) = \{f \colon \Omega \times \mathbb{R} \to \mathbb{C} : f \text{ measurable}, \int_{\Omega \times \mathbb{R}} |f|^q < \infty\}$, which is simultaneously a coarse and a uniform embedding.

Lemma 1.9. For all $0 < \alpha < 2\beta$ there exists $c_{\alpha,\beta} > 0$ such that

$$f(x) := \int_\mathbb{R} \frac{(1 - \cos(tx))^\beta}{|t|^{\alpha+1}} \, dt = c_{\alpha,\beta} |x|^\alpha.$$

Proof. First check the integrand is in $L_1(\mathbb{R})$: as $t \to 0$, $(1-\cos(tx))^\beta \sim |t|^{2\beta}$, so the integrand $\sim |t|^{2\beta-\alpha-1}$, so is integrable on, say, $(-1,1)$, since $2\beta-\alpha-1 > -1$. As $|t| \to \infty$, $(1-\cos(tx))^\beta$ is bounded, so the integrand is $\sim |t|^{-\alpha-1}$, which is integrable on $\mathbb{R} \setminus (-1,1)$, since $-\alpha - 1 < -1$.

For $x > 0$,

$$f(x) = x^\alpha \int_{\mathbb{R}} \frac{(1-\cos(tx))^\beta}{|tx|^{\alpha+1}} x \, dt = x^\alpha \int_{\mathbb{R}} \frac{(1-\cos(s))^\beta}{|s|^{\alpha+1}} x \, dt = x^\alpha f(1).$$

Also, $f(0) = 0$, $f(-x) = f(x)$ for all x. So $f(x) = |x|^\alpha f(1)$ for all x. □

Proof of Proposition 1.8. [A possible attempt is $Tf(\omega,t) = \frac{(1-\cos(tf(\omega)))^{1/2}}{|t|^{(p+1)/q}}$. Then

$$\int_{\mathbb{R}} |Tf(\omega,t)|^q \, dt = \int_{\mathbb{R}} \frac{(1-\cos(tf(\omega)))^{q/2}}{|t|^{p+1}} \, dt = \|f(\omega)\|^p .$$

The problem is taking $Tf - Tg$. The clever thing is that T is exponential.]

Define

$$Tf(\omega,t) = \frac{1-e^{itf(\omega)}}{|t|^{(p+1)/q}}.$$

For $\theta \in \mathbb{R}$, $\left|1-e^{i\theta}\right| = \sqrt{(1-\cos\theta)^2 + \sin^2\theta} = \sqrt{2-2\cos\theta} = \sqrt{2}(1-\cos\theta)^{1/2}$. Then

$$\begin{aligned}
\|Tf\|_q^q &= \int_\Omega \int_{\mathbb{R}} \frac{2^{q/2}(1-\cos(tf(\omega)))^{q/2}}{|t|^{p+1}} \, dt \, d\mu(\omega) \\
&= \int_\Omega 2^{q/2} C_{p,q/2} |f(\omega)|^p \, d\mu(\omega) && \text{by Lemma 8, } \alpha = p, \beta = q/2 \\
&= 2^{q/2} C_{p,q/2} \|f\|_p^p .
\end{aligned}$$

Given $f, g \in L_p(\Omega)$,

$$\left| e^{itf(\omega} - e^{itg(\omega)} \right| = \left| 1 - e^{it(f(\omega)-g(\omega))} \right| .$$

Apply above computation with f replaced with $f-g$ to get

$$\|Tf - Tg\|_q^q = 2^{q/2} C_{p,q/2} \|f-g\|_p^p .$$

Take $\rho_1(t) = \rho_2(t) = \sqrt{2} C_{p,q/2}^{1/q} t^{p/q}$. □

Corollary 1.10. For $1 \le p < q < \infty$ there exists $T\colon L_p \to L_q$ which is a simultaneously coarse and uniform embedding.

Apply proposition 8 with $(\Omega,\mu) = ([0,1],\lambda)$ to get embedding $L_p \to L_q([0,1] \times \mathbb{R}; \mathbb{C})$. Then $L_q([0,1]\times\mathbb{R};\mathbb{C}) \hookrightarrow_2 L_q([-1,1]\times\mathbb{R})$ by $f \mapsto \tilde{f}$ where

$$\tilde{f}(s,t) = \begin{cases} \operatorname{Re} f(s,t) & s \in (0,1] \\ \operatorname{Im} f(-s,t) & s \in [-1,0). \end{cases}$$

Since $L_q([-1,1]\times\mathbb{R})$ is separable, it embeds isometrically into L_q.

Definition. Given families $(M_\alpha)_{\alpha\in A}$ of metric spaces, a family $f_\alpha\colon M_\alpha \to N_\alpha$ a family of coarse embeddings is *uniformly coarse* if there exists increasing $\rho_1, \rho_2\colon \mathbb{R}^+ \to \mathbb{R}^+$ such that $\rho_1(t) \to \infty$ as $t \to \infty$ and

$$\rho_1(d(x, y)) \leq d(f(x), f(y)) \leq \rho_2(d(x, y)) \qquad \forall x, y \in M, \forall \alpha \in A.$$

There are many connections of metric embeddings with other fields of mathematics, for example in geometry. The following two statements are non-examinable.

Theorem (Yu)**.** If M is a uniformly discrete metric space (every element is separated by at least $\delta > 0$) with bounded geometry (the number of points in any radius R is bounded by some $B(R)$)and M coarsely embeds into Hilbert space, then the coarse Baum-Connes conjecture holds for M.

Theorem (Kaspanov,Yu)**.** Same M, if M coarsely embeds into a uniformly convex Banach space then the coarse geometric Novikov conjecture hods for M.

2 Fréchet embeddings, Aharoni's theorem

Theorem 2.1 (Fréchet embedding)**.** Any metric space M embeds isometrically into $\ell_\infty(M)$. If $|M| = n < \infty$ then it isometrically embeds into ℓ_∞^{n-1}. If M is separable, then it embeds isometrically into $\ell_\infty = \ell_\infty(\mathbb{N})$.

Proof. Fix $x_0 \in M$. Define $f\colon M \to \ell_\infty(M)$ by $f(x) = d(\cdot, x) - d(\cdot, x_0)$. Then for all $y \in M$, $|f(x)(y)| = |d(y, x) - d(y, x_0)| \le d(x, x_0)$. So $f(x) \in \ell_\infty(M)$. Observe that for every $x, z \in M$, $\|f(x) - f(z)\|_\infty = \|d(\cdot, x) - d(\cdot, z)\|_\infty \le d(x, z)$ by the triangle inequality. To get the lower bound, $\|f(x) - f(z)\|_\infty \ge |f(x)(z) - f(z)(z)| = d(x, z)$.

In fact we can isometrically embed M into $\ell_\infty(M\backslash\{x_0\})$. If $M = \{x_0, ..., x_{n-1}\}$, then $M \to \ell_\infty^{n-1}$, $x \mapsto d(\cdot, x)$ works.

If M is separable, take a countable dense $S \subset M$. Then S embeds isometrically into ℓ_∞. This extends to an isometric embedding $M \to \ell_\infty$ (given $x \in M$, there exists $x_n \in S$ $x_n \to x$. Let $f(x) = \lim f(x_n)$. Since $f(x_n)$ Cauchy this limit exists. Check that this definition is independent of the choice of sequence).

Another proof: Let $f\colon M \to \ell_\infty(M)$ be an isometric embedding. Then $X = \overline{\text{span}} f(M)$ is a separable Banach space. By Proposition 1.2, $X \hookrightarrow_1 \ell_\infty$. $\square$

Definition. Let $m_\infty(n)$ be the least m such that every n-element metric space embeds isometrically into ℓ_∞^m. By Theorem 2.1, $m_\infty(n) \le n - 1$ for all $n \in \mathbb{N}$.

Aim. There exists $c > 0$, $m_\infty(n) \ge n - cn^{2/3} \log n$ for all $n \ge 2$ (due to K Ball).

Background.

(i) Ramsey Theory: $\forall t \in \mathbb{N}\ \exists n \in \mathbb{N}$ if edges of K_n are red-blue coloured, then there exists a monochormatic copy of K_t in K_n. Let $R(t)$ be the least n that works. It is easy to see that $R(t) \le 4^t$. It is also known that $R(t) \ge c^t$ for some $c > 1$. Given graphs H_1, H_2, let $R(H_1, H_2)$ be the least n s.t. whenever edges of K_n are red-blue coloured, either there exists a red copy of H_1 or there exists a blue copy of H_2 inside of K_n. So $R(t) = R(K_t, K_t)$. We can see that this exists. If $t = \max\{|H_1|, |H_2|\}$ (the *order* $|G|$ of a graph is the number of vertices), then $R(H_1, H_2) \le R(t)$.

(ii) A graph $G = (V, E)$ is *bipartite* if there exists a partition $V = V_1 \cup V_2$ s.t. $\forall x, y \in V$, $xy \in E \implies x \in V_1, y \in V_2$ or $x \in V_2, y \in V_1$. The vertices $V_{1,2}$ are called *vertex classes*. If $E = \{xy : x \in V_1, y \in V_2\}$, then G is the *complete bipartite graph.* This is denoted K_{V_1,V_2}. Denote $K_{m,n}$ = any K_{V_1,V_2} with $|V_1| = m, |V_2| = n$. Observe $K_{2,2} = C_4$.

(iii) Given a graph G, its complement $\bar{G}$ has vertex set $V(\bar{G}) = V(G)$, and $E(\bar{G})$ is the complement of $E(G)$, i.e. $xy \in E(\bar{G}) \iff xy \notin E(G)$.

Definition. For a graph G, define a metric ρ:

$$\rho(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } xy \in E \\ 2 & \text{otherwise.} \end{cases}$$

Lemma 2.2. If (G, ρ) embeds isometrically into ℓ_∞^k, then edges of $\bar{G}$ can be covered by $\le k$ complete bipartite subgraphs of $\bar{G}$.

Proof. Let $f\colon (G,\rho) \to \ell_\infty^k$ be isometric. Let $\alpha_i = \max_{x\in G} f(x)_i$, $\beta_i = \min_{x \in G} f(x)_i$, $i = 1, ..., k$. Observe $\alpha_i - \beta_i = \max x, y \in G(f(x)_i - f(y)_i) \leq \max_{x,y\in G} \|f(x) - f(y)\|_\infty \leq$ 2. Let $I = \{i = 1, ..., k : \alpha_i - \beta_i = 2\}$. Then $xy \in E(\bar{G}) \iff \exists i \in I |f(x)_i - f(y)_i| = 2 \iff \exists i \in I f(x)_i = \alpha_i, f(y)_i = \beta_i$ or vice versa.

Let $V_{i1} = \{x : f(x)_i = \alpha_i\}$ and $V_{i2} = \{x : f(x)_i = \beta_i\}$. Then $E(\bar{G}) = \bigcup_{i\in I} E(K_{V_{i1},V_{i2}})$, and $|I| \leq k$. □

Theorem 2.3. There exists $C > 0$, $\forall n \geq 2$, $m_\infty(n) \geq n - Cn^{2/3} \log n$.

Proof. We will use the following result: $\exists \alpha > 0$, $R(C_4, K_t) > \alpha(t/\log t)^{3/2}$ (Spencer uses probabilistic method). Now there exists $b > 0, \forall n$ if $t = \lceil bn^{2/3} \log n \rceil$, then $n < \alpha(t/\log t)^{3/2} < R(C_4, K_t)$. [Roughly: $n = (t/\log t)^{3/2} \implies t = n^{2/3} \log t$, so $\log t = 2/3 \log n + \log\log t$, $\log t \sim \log t - \log\log t \sim \log n$. So $t \sim n^{2/3} \log n$.] Fix $n \in \mathbb{N}$, let $t = \lceil bn^{2/3} \log n \rceil$. So $n < R(C_4, K_t)$, so there exists a red-blue colouring of K_n without red C_4 or blue K_t. Let G be the blue graph. Let $k = m_\infty(n)$. Since (G, ρ) embeds isometrically into ℓ_∞^k, by Lemma 2.2, $\bar{G}$ = red graph is covered by $\leq k$ complete bipartite subgraphs. Since $C_4 = K_{2,2} \not\subset \bar{G}$, one vertex class in each complete bipartite subgraph is of size 1. So there exists $\leq k$ vertices s.t. every edge in $\bar{G}$ is adjacent to one of them. Since $K_t \not\subset G$, it follows that $n \leq k+t-1$, so $k = m_\infty(n) \geq n-t+1 \geq n-Cn^{2/3} \log n$ for some absolute constant C. □

Remark. Since $R(t) > C^t$ for some $C > 1$, this method won't give better than $n - C \log n$ lower bound on $m_\infty(n)$.

Aim. $n - m_\infty(n) \to \infty$ as $n \to \infty$. (Pretrov, Solyanov(?), Zatitskivy(?))

Lemma 2.4 (Non-linear Hahn-Banach)**.** Let M be a metric space, $A \subset M$, $f\colon A \to \mathbb{R}$ a Lipschitz map with constant L. Then there exists a Lipschitz extension $\tilde{f}\colon M \to \mathbb{R}$ of f with constant L.

Proof. Fix $x_0 \in M \setminus A$. Define $\tilde{f}\colon A \cup \{x_0\} \to \mathbb{R}$ by

$$\tilde{f}(x) = \begin{cases} f(x) & x \in A \\ \alpha & x = x_0. \end{cases}$$

We need to choose the right $\alpha \in \mathbb{R}$. Want

$$|\alpha - f(x)| \leq Ld(x_0, x) \qquad \forall x \in A,$$

i.e.,

$$f(y) - Ld(y, x_0) \leq \alpha \leq f(x) + Ld(x, x_0) \qquad \forall x, y \in A.$$

Such α exists if

$$f(y) - Ld(y, x_0) \leq f(x) + Ld(x, x_0) \qquad \forall x, y \in A \qquad (*).$$

Indeed, then take

$$\alpha = \sup_{y\in A}\{f(y) - Ld(y, x_0)\}.$$

To see (*),

$$f(y) - f(x) \leq Ld(x, y) \leq Ld(x, x_0) + Ld(x_0, y) \qquad \forall x, y \in A.$$

If $M \setminus A$ is finite or countable, then apply above recursively to get an extension. In general, use Zorn's Lemma to get a maximal extension $(\tilde{M}, \tilde{f})$. By above, $\tilde{M} = M$. □

Proposition 2.5. If A is a subset of a finite metric space M, and there exists an isometric embedding $f\colon A \to \ell_\infty^{|A|-k}$, then there exists an isometric embedding $g\colon M \to \ell_\infty^{|M|-k}$.

Proof. Let $f_i(x) = f(x)_i$, $1 \le i \le |A| - k$. Then each f_i is 1-Lipschitz so by Lemma 2.4, there exists a 1-Lipschitz extension $g_i\colon M \to \mathbb{R}$. Enumerate $M \setminus A$ as y_i, $|A| - k + 1 \le i \le |M| - k$ and let $g_i(x) = d(x, y_i)$, $x \in M$. Then $g\colon M \to \ell_\infty^{|M|-k}$, $g(x) = (g_i(x))_{i=1}^{|M|-k}$ is an isometric embedding. □

Background.

(i) Some more Ramsey Theory: For $s \ge 2$, $n \in \mathbb{N}$, $K_n^{(s)} = \{A \subset [n] : |A| = s\}$, $[n] = \{1, ..., n\}$. e.g. $K_n^{(2)} = K_n$. Then Ramsey says $\forall s, \forall t, \exists n$ if $K_n^{(s)}$ is red-blue coloured, then there exists a monochromatic copy of $K_t^{(s)}$, i.e. $\exists A \subset [n]$, $|A| = t$ s.t. $A^{(s)} = \{B \subset A : |B| = s\}$ is monochromatic.

Also $\forall s, \forall t, \forall c, \exists n$ if $K_n^{(s)}$ is c-coloured then $\exists$ monochromatic copy of $K_t^{(s)}$.

(ii) Recall that a *tree* T is a connected, acyclic graph. Equivalently, $\forall x, y \in T$, $\exists$ unique path x to y. If $\mathrm{diam}(T) = \max x, y \in T d(x, y) \le 4$, then there exists $c \in T$ $\forall x$ $d(c, x) \le 2$. Call this c a *centre* of T. Vertices in $\Gamma(c) = \{a \in T : ac \in E\}$ are the *main vertices*. Every other vertex is connected to a unique main vertex.

(iii) *Oriented graph.* An *orientation* of a graph G is an assignment of direction for each edge: if $e = xy \in E(S)$, there are two choices $\overrightarrow{xy}$ or $\overrightarrow{yx}$. This is called *alternating* if $\forall x$ either $\forall y \in \Gamma(x)$ we have $\overrightarrow{xy}$ (x is a *source*) or $\forall y \in \Gamma(x)$ we have $\overrightarrow{yx}$ (x is a *sink*). [The name comes from an alternating path, because once we make a choice on one edge, all the other edges are alternating in direction.] A connected graph has 0 or 2 alternating orientations. It has 0 iff it has an odd cycle, i.e. not bipartite. A tree has exactly two alternating orientations.

(iv) A metric space $\{x_1, ..., x_n\}$ is *generic* if the $\binom{n}{2}$ distances $d(x_i, x_j)$, $1 \le i < j \le n$ are linearly independent over $\mathbb{Q}$.

Theorem 2.6. For every $k \in \mathbb{N}$, there exists $N \in \mathbb{N}$ for all $n \ge N$, $m_\infty(n) \le n - k$.

Proof. Step 1. We can restrict to generic metric spaces. *Proof.* Let $M = \{x_1, ..., x_n\}$ with metric d be an arbitrary metric space. For $j \in \mathbb{N}$, we can pick $\alpha_{rs} \in (\frac{1}{2j}, \frac{1}{j})$, $1 \le r < s \le n$ s.t. $d_j(x_r, x_s) = d(x_r, x_s) + \alpha_{rs}$ defines a generic metric.

Suppose $\forall j, \exists$ isometric embedding $f_j\colon (\{x_1, ..., x_n\}, d_j) \to \ell_\infty^m$ for some m. WLOG $\mathrm{im}(f_j)$ is bounded independent of j. By compactness, after passing to a subsequence, we have $f(x_r) = \lim_{j\to\infty} f_j(x_r)$ exists $\forall r$. Then $f\colon (M, d) \to \ell_\infty^m$ is an isometric embedding.

From now on, M is an n-element generic metric space and the elements of M are real numbers.

Step 2. Assume $f\colon M \to \mathbb{R}$ is 1-Lipschitz. We define a graph $G(f)$ with vertex set M and $xy \in E \iff |f(x) - f(y)| = d(x, y)$. We will orient an edge xy s.t. for $\overrightarrow{xy}$ we have $f(x) - f(y) = d(x, y)$.

Example. $f(x) = d(x, a)$, then this is a star with centre a and every edge pointing to it. For $x \neq y$ in $M \setminus \{a\}$, $f(x) - f(y) < d(x, y)$.

We have functions $f_1, ..., f_m \colon M \to \mathbb{R}$, $f \colon M \to \ell_\infty^m$ given by $f(x) = (f_i(x))_{i=1}^m$. Then f is an isometric embedding $\iff$ the f_i are 1-Lipschitz, $\forall x \neq y, \exists i, xy \in E(G(f_i))$. So M embeds isometrically into ℓ_∞^m $\iff$ the edges of the complete graph on M can be covered by at most m *Lipschitz graphs.*

Step 3. Let T be a tree on M with $\operatorname{diam}(T) \leq 4$. For fixed $x_0 \in T$, $\alpha \in \mathbb{R}$, alternating orientation of T, consider the unique function $f \colon M \to \mathbb{R}$ where $f(x_0) = \alpha$, $f(x) - f(y) = d(x, y)$ for all $\overrightarrow{xy} \in E(T)$. Then f is 1-Lipschitz $\iff$ for every path $wxyz$ in T, $d(w, x) + d(y, z) < d(x, y) + d(w, z)$. [We only need $\Leftarrow$ direction.]

Proof. Given $x, y \in T$, we need $|f(x) - f(y)| \leq d(x, y)$. If $d_T(x, y) = 0$ or 1, then this is true [d_T is the graph distance on T.]

If we have a path xzy, then $|f(x) - f(y)| = |f(x) - f(z) + f(z) - f(y)| = |d(x, z) - d(z, y)| < d(x, y)$ [here we use orientation.]

If we have a path $xwzy$ then $|f(x) - f(y)| = |f(x) - f(w) + f(w) - f(z) + f(z) - f(y)| = |d(x, w) - d(w, z) + d(z, y)|$. If this $= d(x, w) - d(w, z) + d(z, y)$ then $< d(x, y)$ by assumption. If this $= -d(x, w) + d(w, z) - d(z, y)$ then by the triangle inequality this is $< d(x, z) - d(z, y) < d(x, y)$.

If we have a path $xuwzy$ then $|f(x) - f(y)| = |d(x, u) - d(u, w) + d(w, z) - d(z, y)|$. WLOG this $= d(x, y) - d(u, w) + d(w, z) - d(z, y)$ because we have an even number of terms. By the assumption, this $< d(x, z) - d(z, y)$ and by the triangle inequality, this $< d(x, y)$. Thus we have proved step 3.

A tree T on M is *admissible* if it has diam ≤ 4 and satisfies the assumption in step 3.

Step 4. Given distinct points $c, a_1, ..., a_m$ in M, there exists a unique admissible tree on M with centre c and main vertices $a_1, ..., a_m$. Denote this by $T(c; a_1, ..., a_m)$.

Proof. Given $x \in M \setminus \{c, a_1, ..., a_m\}$, x can be joined to main vertex a $\iff$ for every main vertex $b \neq a$ we have $d(x, a) + d(c, b) < d(a, c) + d(x, b)$, i.e.

$$d(x, a) - d(a, c) < d(x, b) - d(c, b).$$

So a is uniquely determined.

Step 5. We colour $M^{(4)}$ using as colours elements of S_3 as follows: given $w < x < y < z$ in M, let $R_1 = d(w, x) + d(y, z)$, $R_2 = d(w, y) + d(x, z)$, $R_3 = d(w, z) + d(x, y)$. We give $wxyz$ colour ijk if $R_i > R_j > R_k$. This is a 6-colouring of $M^{(4)}$.

Main Claim. $\forall k \in \mathbb{N}, \forall c \in S_3, \exists t_c \in \mathbb{N}$, every (any) monochromatic metric space of size t_c and colour c can be covered by $\leq t_c - k$ admissible trees.

From main claim, let $t = \max_{c \in S_3} t_c$. By Ramsey $\exists N$ s.t. if $K_N^{(4)}$ is 6-coloured, then there exists a monochromatic $K_t^{(4)}$. So given $n \geq N$, an n-element metric space M, there exists a colour $c \in S_3$ and $A \subset M$, $|A| = t_c$ s.t. A is monochromatic. By main claim, the complete graph on A can be covered by $|A| - k$ admissible trees, so by step 2, A embeds isometrically into $\ell_\infty^{|A|-k}$. By Proposition 2.5, M embeds isometrically into ℓ_∞^{M-k}. So done. It remains to check the main claim.

Recall the *Main Claim*: $\forall c \in S_3, \forall k, \exists t$ every metric space M with $|M| = t$ and colour c can be coloured by $t - k$ admissible trees [edges of the complete graph of M].

Proof of Main Claim:

Case 1, $c = 213$. Then there does not exist M of colour c of size ≥ 5 ($t = 5$ will do). To see this, assume the contrary and aim for a contradiction. Fix $u < w < x < y < z$ in M. Then

$$d(u,w) + d(x,y) > d(u,y) + d(w,x);$$
$$d(w,x) + d(y,z) > d(w,z) + d(x,y);$$
$$d(u,y) + d(w,z) > d(u,w) + d(y,z).$$

Adding these gives $0 > 0$, a contradiction.

Case 2, $c = 312$. Same; just replace $>$ with $<$.

Case 3, $c = 132$. *Mini claim:* Assume if $|M| = n$, colour 132, then all but m edges of K_M can be covered by s admissible trees. Then if $|M| = n+2$ of colour 132, then all but $m-1$ edges of K_M can be covered by $s+2$ admissible trees.

Proof of mini claim. $|M| = n$, colour 132 and we have s trees that cover all but m edges. Let $ab, a < b$ be one of these edges. Let $|M'| = n+2$, colour 132. WLOG $M' = M \cup \{a', b'\}$, where $a < a' < b' < b$ and $M \cap ((a, a'] \cup [b', b)) = \emptyset$. Extend the s trees to the whole of M'. By step 4, add $T(a; a', b), T(b; a', b')$. Every $x \in M' \setminus \{a, a', b\}$ is joined to a' in $T(a; a', b)$ Every $x \in M' \setminus \{b, a', b'\}$ is joined to b' in $T(b; a', b')$. □

Apply mini claim: start with $|M| = k$ and $s = 0$, $m = \binom{k}{2}$. Apply mini claim m times to get M' with $t = |M'| = k + 2\binom{k}{2} = k^2$, $s = 2\binom{k}{2} = t - k$, $m = 0$.

Case 4, $c = 123$. We prove Main Claim by induction on k. For $k = 1$, $t = 1$ will do. I have 0 edges so 0 trees will do. Let $k \geq 1$ and assume t works for k. For $k+1$, we prove that $2t+3$ works. Take $M = \{-1, 0, 1, 2, ..., t+1, t+2, ..., 2t+1\}$. Consider $T(0; -1, 2), T(1; 0, 2), T(t+1+i; i, i+1), 1 \leq i \leq t$. This covers all edges except perhaps edges between vertices in $\{t+2, ..., 2t+1\}$. These can be covered by $t - k$ trees by the induction hypothesis. So we need $2 + t + t - k = 2t + 2 - k = |M| - (k+1)$.

Case 5, $c = 231$. We show $t = 2k$ works for k. Take $M = \{-k, ..., -1, 1, ..., k\}$, take trees $T(-i; -k, -k+1, ..., -i-1, 1, ..., k)$, $1 \leq i \leq k$. This works [a bit fiddly and uninteresting].

Case 6, $c = 321$. $t = 4k+1$ works. $M = \{0, 1, ..., 4k\}$, take trees $T(0; i, 4k+1-i)$, $1 \leq i \leq 2k$, $T(i; 2k+i, 2k+i+1, ..., 4k+1, i)$ $1 \leq i \leq k$. So the number of tree is $3k = |M| - (k+1)$.

□

Remark. $m_\infty(n)$ = least m s.t. every n-element subset of some $L_\infty(\Omega, \mu)$ embeds isometrically into ℓ_∞^m.

We define for $1 \leq p < \infty$, $m_p(n)$ = least m s.t. every n-element subset of some $L_p(\Omega, \mu)$ embeds isometrically into ℓ_p^m.

Remark. Note $m_1(n) \leq n!$ (Proposition 1.4), $m_2(n) = n-1$ (Example Sheet).

Theorem 2.7. For all $1 \leq p < \infty$ and for all $n \geq 2$, $m_p(n) \leq \binom{n}{2}$.

Remark. For $1 \leq p < 2$, this is essentially best possible. [Example sheet: $m_p(2n+1) \geq n$.]

Lemma 2.8 (Carathéodory's Theorem)**.** Given $L \subset \mathbb{R}^N$, then $\operatorname{conv} L = \{\sum_{i=0}^N t_i x_i : x_i \in L, t_i \geq 0, \forall i, \sum_{i=0}^N t_i = 1\}$. It follows that $\overline{\operatorname{conv}} L = \operatorname{conv} L$ if L is compact.

Proof. Given $x \in \operatorname{conv} L$, we write $x = \sum_{i=1}^m t_i x_i$. WLOG $m > N+1, t_i > 0, \forall i$. Then $x_1, ..., x_m$ are affinely dependent – this means $x_1 - x_2, x_1 - x_3, ..., x_1 - x_m$ are linearly dependent. There exists $\lambda_1, ..., \lambda_m$ not all zero with $\sum \lambda_i = 0$, $\sum \lambda_i x_i = 0$. For any $s \in \mathbb{R}$, $\sum (t_i - s\lambda_i) = 1$, $\sum (t_i - s\lambda_i) x_i = x$. For $s > 0$, $t_i - s\lambda_i \geq 0$ if $\lambda_i \leq 0$. So we take $s = \min\{t_i/\lambda_i : \lambda_i > 0\}$ $(\exists i, \lambda_i > 0)$. Now $t_i - s\lambda_i \geq 0, \forall i$ and $\exists i, t_i - s\lambda_i = 0$. □

Proof of Theorem 2.7. Fix $n \geq 2$. Given an n-tuple $M = (x_1, ..., x_n)$ in some $L_p(\Omega, \mu)$, let $\theta_M = (\|x_i - x_j\|_p^p)_{1 \leq i < j \leq n} \in \mathbb{R}^N$ where $N = \binom{n}{2}$. Let $C = \{\theta_M : M$ is an n-tuple in some $L_p(\Omega, \mu)\}$.

C is a cone: $\theta \in C, t > 0 \implies t\theta \in C$. Suppose $M = (x_1, ..., x_n)$ is an n-tuple in $L_p(\Omega, \mu)$, $M' = (y_1, ..., y_n)$ in $L_p(\Omega', \mu')$. Then consider $z_i = (x_i, y_i) \in L_p(\Omega \amalg \Omega')$. Then

$$\|z_i - z_j\|_p^p = (\theta_M)_{ij} + (\theta_{M'})_{ij} \qquad \forall 1 \leq i < j \leq n.$$

So $\theta_M + \theta_{M'} \in C$.

Let

$$K = C \cap \left\{ \theta \in \mathbb{R}^N : \sum_{1 \leq i < j \leq n} \theta_{ij} = 1 \right\}.$$

Say $\theta \in C$ is *linear* if there exists $(t_1, ..., t_n) \in \mathbb{R}^n$ s.t. $\theta_{ij} = |t_i - t_j|^p$. Let

$$\begin{aligned} L &= \{\theta \in K : \theta \text{ is linear}\} \\ &= \left\{ (|t_i - t_j|^p)_{1 \leq i < j \leq n} : t_1, ..., t_n \in \mathbb{R}, \sum_{1 \leq i < j \leq n} |t_i - t_j|^p = 1. \right\} \end{aligned}$$

Note L is compact. K is convex so $\operatorname{conv} L \subset K$.

Given $\theta \in K$, say $\theta = (\|x_i - x_j\|_p^p)_{1 \leq i < j \leq n}$, where $x_1, ..., x_n \in L_p(\Omega, \mu)$. Can approximate x_i with simple function y_i s.t. $\varphi = (\|y_i - y_j\|_p^p) \in K$. So we have a measurable partition $\Omega = \bigcup_{r=1}^R A_r$ of Ω s.t. $y_i|_{A_r}$ is constant $\forall i, r$. Let $\varphi_r = (\|y_i|_{A_r} - y_j|_{A_r}\|_p^p)_{1 \leq i < j \leq n}$. Then φ_r is linear and $\varphi = \sum_{r=1}^R \varphi_r$. Let $\alpha_r = \sum_{1 \leq i < j \leq n} (\varphi_r)_{ij}$. Then $\sum_{r=1}^R \alpha_r = 1$. So $\varphi = \sum_{r=1}^R \alpha_r (\varphi_r / \alpha_r) \in \operatorname{conv} L$. This shows $K \subset \overline{\operatorname{conv}} L$. By Lemma 2.8, $K = \operatorname{conv} L$, and every $\theta \in C$ is a sum $\theta = \sum_{r=1}^N \theta_r$, where θ_r is linear for all r. Note $\{\theta : \sum \theta_{ij} = 1\}$ is $(N-1)$-dimensional. For each r, there exists $t_{ri} \in \mathbb{R}$ with $\theta_r = (|t_{r,i} - t_{r,j}|^p)_{1 \leq i < j \leq n}$. If $\theta = \theta_M$, $M = (x_1, ..., x_n) \in L_p(\Omega, \mu)^n$, define $f \colon M \to \ell_p^N$ by using these as coordinates: $f(x_i) = (t_{r,i})_{r=1}^R$. Then one line to check that this works. For $1 \le i < j \le n$,

$$\|f(x_i) - f(x_j)\|_p^p = \sum_r |t_{r,i} - t_{r,j}|^p = \sum_r (\theta_r)_{ij} = \theta_{ij} = \|x_i - x_j\|_p^p .$$

□

Theorem 2.9 (Aharoni's Theorem)**.** For any $\epsilon > 0$, any separable metric space embeds into c_0 with distortion $\le 3 + \epsilon$.

Motivation. Given Banach spaces X, Y, if X bilipschitzly embeds into Y, must X be isomorphically embed into Y? Yes, if Y is separable and there exists a Banach Space W such that $Y \sim W^*$. Theorem 9 shows that in general, the answer is no.

Notation. (i) In a metric space M, for $x \in M$ and $\delta > 0$ let $B_\delta(x) = \{y \in M : d(y,x) \le \delta\}$. $A \subset M$ is *δ-dense* in M if $\forall x \in M, d(x, A) < \delta$.

(ii) Given a set S, let $c_0(S) = \{f \in \ell_\infty(S) : \forall \epsilon > 0 \{s \in S : |f(s)| > \epsilon\}$ is finite$\}$. So $c_0 = c_0(\mathbb{N}) \cong c_0(S)$ for S countably infinite.

Idea. We will have a countable set S and a subset $M_S \subset M$ and we use maps $f \colon M \to c_0(S), f(x) = (d(x, M_S))_{s \in S}$. Fix $\delta > 1$, for $x \ne y$ in M, $\delta^n \le d(x,y) \le \delta^{n+1}$ for some $n \in \mathbb{Z}$. We will have $c \in M$ (a centre). One of x or y, say x, has $d(c,x) > \delta^n/2$. We will partition $M \setminus B_{\delta^n/2}(c)$.

Lemma 2.10. Let M be a separable metric space, $\lambda > 2, a > 0, N \subset M$. Then there exists subsets $M_1, M_2, \ldots$ of N such that

(i) $\forall x \in N, \exists i, d(x, M_i) < a$;

(ii) $\forall x \in M$, the set $\{i : d(x, M_i) < (\lambda - 1)a\}$ is finite;

(iii) $\forall i$, $\operatorname{diam}(M_i) \le 2\lambda a$.

Proof. WLOG $a = 1$ (just replace the distance d by d/a). M is separable, hence so is N, so there exists a countable sets

$$Z \subset N, \text{ which is 1-dense in } N,$$
$$Y \subset M, \text{ which is 1-dense in } M.$$

WLOG $Z \subset Y$ (replace Y by $Z \cup Y$). Enumerate Y as $y_1, y_2, y_3, \ldots$. Let $M_i = (B_\lambda(y_i) \cap Z) \setminus \bigcup_{j<i} M_j$. Then $\forall i, M_i \subset Z \subset N$, and $\forall i, M_i \subset B_\lambda(y_i)$. So $\operatorname{diam}(M_i) \le 2\lambda$. This shows (iii).

Given $x \in N$, there exists i such that $y_i \in Z$ and $d(x, y_i) < 1$. Then $y_i \in B_\lambda(y_i) \cap Z \subset \bigcup_{j=1}^i M_j$. So there exists $j \le i$ such that $d(x, M_j) < 1$. This shows (i).

Given $x \in M$, there exists i_0 such that $d(x, y_{i_0}) < 1$. If $d(x, M_i) < \lambda - 1$, then $d(y_{i_0}, M_i) < \lambda$ by the triangle inequality. For $i > i_0$, $y \in M_i$. Since $y_{i_0} \in \bigcup_{j \le i_0} M_j$ and $M_i \cap \bigcup_{j \le i_0} M_j = \emptyset$, we have $d(y_{i_0}, y) \ge \lambda$ so $d(y_{i_0}, M_i) \ge \lambda$. So $\{i : d(x, M_i) < \lambda - 1\}$ has at most i_0 elements. This shows (ii). □

Proof of Theorem 9, Assonad. Given separable metric space M and $\epsilon > 0$, choose $\lambda > 2$, $\eta > 0$ such that $\frac{3\lambda}{\lambda-2}(1+\eta) < 3+\epsilon$ [first choose λ so that $\frac{3\lambda}{\lambda-2} < 3+\epsilon$]. For $k \in \mathbb{Z}$, let $a_k = (1+\eta)^{-k}$. Fix a centre $c \in M$, let $M_k = M \setminus B_{3\lambda a_k/2}(c)$. Apply Lemma 10 to M, $N = M_k$, $a = a_k$ to get subsets $M_{k,1}, M_{k,2}, \ldots$ satisfying (i),(ii),(iii) in Lemma 10 with $M_{k,i}$ in place of M_i.

Let $S = \{(k,i) : k \in \mathbb{Z}, i = 1, 2, \ldots\}$. For $x \in M$, let $f_{k,i}(x) = [(\lambda-1)a_k - d(x, M_{k,i})] \vee 0$. Let $f(x) = (f_{k,i}(x))_{(k,i)\in S}$.

We first prove that $f(x) \in c_0(S)$. Since $(\lambda-1)a_k \to 0$ as $k \to \infty$, enough to show that for any $s \in \mathbb{Z}$, $T = \{(k,i) : f_{k,i}(x) \geq (\lambda-1)a_s\}$ is finite. For $k > s$, we have $f_{k,i}(x) \leq (\lambda-1)a_k < (\lambda-1)a_s$ so $(k,i) \notin T$ for all $k > s$ and for all i. For the other direction, since $a_k \to \infty$ as $k \to -\infty$, $\exists r < s$ s.t. $d(x,c) < (\frac{\lambda}{2}+1)a_r$. For $k < r$, $d(x,c) < (\frac{\lambda}{2}+1)a_k$, so $\forall i$,

$$d(x, M_{k,i}) \geq d(x, M \setminus B_{3\lambda a_k/2}(c)) \geq \frac{3\lambda a_k}{2} - d(x,y) > (\lambda-1)a_k,$$

so $\forall k < r$, $\forall i$, $f_{k,i}(x) = 0$, so $(k,i) \notin T$. Finally, by Lemma 10, for each k, $\{i : f_{k,i}(x) > 0\} = \{i : d(x, M_{k,i}) < (\lambda-1)a_k\}$ is finite. So $T \subset \bigcup_{k=1}^{s}\{i : f_{k,i}(x) > 0\}$ is finite.

Now we have $f\colon M \to c_0(S)$. This is clearly 1-Lipschitz. For the lower bound, fix $x \neq y$ in M and choose k such that

$$3\lambda a_k < d(x,y) \leq 3\lambda a_k(1+\eta).$$

By the triangle inequality, both x and y cannot belong to $B_{3\lambda a_k/2}(c)$, so WLOG $x \in M_k$. By Lemma 10(i), there exists i such that $d(x, M_{k,i}) < a_k$. So $f_{k,i}(x) \geq (\lambda-1)a_k - a_k = (\lambda-2)a_k$.

Pick $w \in M_{k,i}$, $d(x,w) < a_k$. For any $z \in M_{k,i}$ we have

$$d(y,z) \geq d(y,x) - d(x,w) - d(w,z) > 3\lambda a_k - a_k - \operatorname{diam}(M_{k,i}) \geq (\lambda-1)a_k,$$

because $\operatorname{diam}(M_{k,i}) \leq 2\lambda a_k$. So $d(y, M_{k,i}) \geq (\lambda-1)a_k$ and $f_{k,i}(y) = 0$. So

$$\begin{aligned}\|f(x) - f(y)\|_\infty &\geq |f_{k,i}(x) - f_{k,i}(y)| \\ &\geq (\lambda-2)a_k \\ &= \frac{3\lambda a_k(1+\eta)}{3\lambda(1+\eta)}(\lambda-2) \\ &> \frac{d(x,y)}{3+\epsilon}.\end{aligned}$$

□

Remark. Here we are embedding into $c_0^+(S) = \{f\colon S \to \mathbb{R}^+ : f \in c_0(S)\}$. Pelant showed that

$$\sup_M \inf_{f\colon M \to c_0^+} \operatorname{dist}(f) = 3,$$

where the supremum is over all separable metric space M and the infimum is over all bilipschitz embeddings f.

Kalton and Lancien showed that

$$\sup_M \inf_{f\colon M \to c_0} \operatorname{dist}(f) = 2.$$

3 Bourgain's Embedding Theorem

For metric spaces X, Y, let

$$c_Y(X) = \inf\{\operatorname{dist}(f) : f\colon X \to Y \text{ a bilipschitz embedding}\}.$$

If $Y = L_p$, we write $c_p(X) = c_{L_p}(X)$, the *L_p-distortion* of X. $c_2(X)$ is called the *Euclidean distortion* of X. By Proposition 1.1, $c_p(X) \le c_2(X)$ for any finite X.

Theorem 3.1 (Dvoretzky's Theorem)**.** $\forall n \in \mathbb{N}, \forall \epsilon > 0, \exists N = N(n, \epsilon)$, s.t. every Banach space Y with $\dim Y \ge N$ contains a $(1+\epsilon)$-isomorphic copy of ℓ_2^n.

Remark.

(i) $N \le \exp(Cn/\epsilon^2)$ for some absolute constant C.

(ii) $c_Y(X) \le c_2(X)$ for every finite metric space and every infinite dimensional Banach space Y.

Aim. $c_2(X) \le C \log |X|$ for every finite X (Bourgain's embedding theorem).

From now on we fix a metric space X with $|X| = n$. Let $\mathcal{P}_X$ be the set of all partitions of X [pairwise disjoint non-empty subsets of X whose union is X]. For $P \in \mathcal{P}_X$, the elements of P are called *clusters*. For $x \in X$, we let $P(x)$ be the unique cluster to which x belongs. A *stochastic decomposition of* X is a probability measure Ψ on $\mathcal{P}_X$. Given $\Delta > 0$, $\epsilon\colon X \to (0, 1]$, we say Ψ is an *(ϵ, Δ)-padded decomposition* if

(i) $\forall P \in \mathcal{P}_X$ if $\Psi(P) > 0$ then $\forall C \in P$, $\operatorname{diam}(C) < \Delta$ [clusters can't be too big];

(ii) $\forall x \in X, \Psi(d(x, X \setminus P(x)) \ge \epsilon(x)\Delta) \ge \frac{1}{2}$.

Write $\operatorname{supp}(\Psi) = \{P \in \mathcal{P}_X : \Psi(P) > 0\}$, the *support* of Ψ.

Lemma 3.2. Let Ψ be an (ϵ, Δ)-padded decomposition of X, and let $1 \le q < \infty$. Then there exists 1-Lipschitz map $f\colon X \to \ell_q$ s.t.

(i) $\|f(x)\|_q \le \Delta, \forall x \in X$ (technical condition);

(ii) $\|f(x) - f(y)\|_q \ge C\epsilon(x)d(x, y), \forall x, y$ such that $d(x, y) \in [\Delta, 2\Delta)$, where C is an absolute constant (I think $C = \frac{1}{16}$) (lower Lipschitz condition).

Definition. For Banach spaces $X_1, X_2, \ldots$, for $1 \le q < \infty$ define $\left(\bigoplus_{i\ge1} X_i\right)_q$ to be the space of sequences $(x_i)_{i\ge1}$ s.t. $\sum_{i\ge1} \|x_i\|^q < \infty$. This is a Banach space with norm

$$\|(x_i)\| = \left(\sum_{i\ge1} \|x_i\|^q\right)^{1/q}.$$

Can also define $\left(\bigoplus_{i\ge1} X_i\right)_\infty$; $\|(x_i)\| = \sup_{i\ge1} \|x_i\|$. This has subspace $\left(\bigoplus_{i\ge1} X_i\right)_{c_0}$ of sequences $(x_i)_{i\ge1}$ such that $\|x_i\| \to 0$.

If $X_i = \ell_q$ for all i, then $\left(\bigoplus_{i\ge1} X_i\right)_q \cong \ell_q$.

Proof of Lemma 2. Fix $P \in \operatorname{supp}(\Psi)$. Let $C_1, C_2, ..., C_{m(P)}$ be the clusters of P. Let $U_1, ..., U_{2^{m(P)}}$ be all possible unions of the C_j. Fix $1 \leq j \leq 2^{m(P)}$ and define $f_{P,j} \colon X \to \mathbb{R}$ by

$$f_{P,j}(x) = \begin{cases} d(x, X \setminus P(x)) \wedge \Delta & \text{if } x \in U_j; \\ 0 & \text{otherwise.} \end{cases}$$

[Here $\wedge$ denotes the minimum.] We have $f_{P,j}(x) \leq \Delta$ for all $x \in X$.

Fix $x, y \in X$. If $P(x) \neq P(y)$ then

$$0 \leq f_{P,j}(x), f_{P,j}(y) \leq d(x,y).$$

If $P(x) = P(y)$, then either $x, y \in U_j$, in which case

$$f_{P,j}(z) = d(z, X \setminus P(x)) \wedge \Delta, \qquad z = x, y,$$

or $x, y \notin U_j$ in which case $f_{P,j}(x) = f_{P,j}(y) = 0$. In all cases $|f_{P,j}(x) - f_{P,j}(y)| \leq d(x,y)$. So $f_{P,j}$ is 1-Lipschitz.

Do this for each j, and define $f_P \colon X \to \ell_q^{2^{m(P)}}$ by

$$f_P(x) = \left(2^{-m(P)/q} f_{P,j}(x)\right)_{j=1}^{2^{m(P)}}.$$

So for all x,

$$\|f_P(x)\|_q = \left(\sum_{j=1}^{2^{m(P)}} 2^{-m(P)} f_{P,j}(x)^q\right)^{1/q} \leq \Delta$$

and for all x, y,

$$\|f_P(x) - f_P(y)\|_q = \left(\sum_{j=1}^{2^{m(P)}} 2^{-m(P)} |f_{P,j}(x) - f_{P,j}(y)|^q\right)^{1/q} \leq d(x,y).$$

So f_P is 1-Lipschitz.

Finally define

$$f \colon X \to \left(\bigoplus_{P \in \operatorname{supp}(\Psi)} \ell_q^{2^{m(P)}}\right)_{\ell_q} \hookrightarrow_{\cong} \ell_q,$$

by

$$f(x) = \left(\Psi(P)^{1/q} f_P(x)\right)_{P \in \operatorname{supp}(\Psi)}.$$

For all $x \in X$,

$$\|f(x)\|_q = \left(\sum_P \Psi(P) \|f_P(x)\|^q\right)^{1/q} \leq \Delta.$$

Similarly, f is 1-Lipschitz.

Fix x, y such that $d(x, y) \in [\Delta, 2\Delta)$. Let

$$E = \{P \in \operatorname{supp}(\Psi) : d(x, X \setminus P(x)) \geq \epsilon(x)\Delta\}.$$

Fix $P \in E$. If $x \in U_j$ and $y \notin U_j$ then$|f_{P,j}(x) - f_{P,j}(y)| \geq d(x, X \setminus P(x)) \geq \epsilon(x)\Delta$.

For $\frac{1}{4}$ of values of j we have $x \in U_j$, $y \notin U_j$ (note $P(x) \neq P(y)$, since $\forall C \in P, \operatorname{diam}(C) < \Delta \leq d(x, y)$). So

$$\|f_P(x) - f_P(y)\|_q \geq \left(\sum_{j, x \in U_j, y \notin U_j} 2^{-m(P)} |f_{P,j}(x) - f_{P,j}(y)|^q \right)^{1/q} \geq \epsilon(x)\Delta 4^{-1/q}.$$

Finally,

$$\|f(x) - f(y)\| \geq \left(\sum_{P \in E} \Psi(P) \|f_P(x) - f_P(y)\|^q \right)^{1/q} \geq \epsilon(x)\Delta 4^{-1/q} \Psi(E),$$

and this is

$$\geq \frac{\epsilon(x)\Delta}{4^{1/q} 2} \geq \frac{\epsilon(x)}{4^{1/q} 4} d(x, y) \geq \frac{1}{16} \epsilon(x) d(x, y).$$

□

Definition. Define the set of *relevant scales* to be

$$S(X) = \{\ell \in \mathbb{Z} : \exists x, y \in X, d(x, y) \in [2^\ell, 2^{\ell+1})\},$$

and $R(X) = |S(X)|$.

Example. If X is a connected graph with the graph distance, then $R(X) \leq \lceil \log_2 n \rceil$.

Definition. A map $f\colon X \to Y$, given $K, \tau > 0$, is called a *scaled-τ embedding with deficiency K* if f is 1-Lipschitz and $d(f(x), f(y)) \geq K^{-1} d(x, y)$ for all x, y such that $d(x, y) \in [\tau, 2\tau)$.

Proposition 3.3. Given $K > 0$, $1 \leq q < \infty$, assume $\forall \ell \in S(X), \exists f_\ell\colon X \to \ell_q$ a scale-2^ℓ embedding with deficiency K. Then $C_q(X) \leq K R(X)^{1/q}$.

Proof. Define $f\colon X \to \left(\bigoplus_{\ell \in S(X)} \ell_q\right) \cong \ell_q$ by $f(x) = (f_\ell(x))_{\ell \in S(X)}$. For all x, y, $\|f(x) - f(y)\| = \left(\sum_{\ell \in S(X)} \|f_\ell(x) - f_\ell(y)\|^q\right)^{1/q} \leq R(X)^{1/q} d(x, y)$. So f is $R(X)^{1/q}$-Lipschitz. Given $x \neq y$, there exists $\ell \in S(X)$ s.t. $d(x, y) \in [2^\ell, 2^{\ell+1})$. Then

$$\|f(x) - f(y)\| \geq \|f_\ell(x) - f_\ell(y)\| \geq \frac{1}{K} d(x, y).$$

So $c_q(X) \leq \operatorname{dist}(f) \leq K R(X)^{1/q}$. □

Corollary 3.4. If $\forall \ell \in S(X)$ there exists an $(\epsilon, 2^\ell)$-padded decomposition of X with $\epsilon(x) \geq \frac{1}{K}$ for all x, then $c_q(X) \leq C K R(X)^{1/q}$ $(1 \leq q < \infty)$.

Proof. Lemma 2 + Proposition 3. □

Remark. Actually, $c_q(X) \leq CKR(X)^{1/2 \wedge 1/q}$, because $c_q(X) \leq c_2(X)$.

Theorem 3.5 (Existence of a decomposition)**.** For every $\ell \in \mathbb{Z}, \exists(\epsilon, 2^\ell)$-padded decomposition of X with

$$\epsilon(x) = \left[16 + 16 \log\left(\frac{|B_{2^\ell}(x)|}{|B_{2^{\ell-3}}(x)|}\right)\right]^{-1}.$$

Remark. Note $\epsilon(x) \geq C\frac{1}{\log n}$, so by Corollary 4, $c_2(X) \leq C(\log n)\sqrt{R(X)}$.

Proof of Theorem 5. Fix $\ell \in \mathbb{Z}$ and set $\Delta = 2^\ell$. Fix an ordering $<$ on X. Consider a pair (π, α) where $\pi \in S_n$ (the symmetry group of X) and $\alpha \in (\frac{1}{4}, \frac{1}{2})$ and π, α are chosen uniformly at random and independently. To (π, α) there corresponds an element $P \in \mathcal{P}_X$ with clusters

$$C_y = B_{\alpha\Delta}(y) \setminus \bigcup_{z:\pi(z)<\pi(y)} B_{\alpha\Delta}(z), \qquad y \in X.$$

We throw away the empty clusters. This gives a random partition, so we have a stochastic decomposition.

Now we check this gives us an (ϵ, Δ)-padded decomposition. Note that $\operatorname{diam}(C_y) \leq 2\alpha\Delta < \Delta$ for ally $y \in X$. Now fix $x \in X$, $t \leq \frac{\Delta}{8}$. Let B (B for Bad) be the event that $d(x, X \setminus P(x)) \leq t$, i.e. $B_t(x) \not\subset P(x)$. The aim is to show that $\mathbb{P}(B) \leq \frac{1}{2}$ for $t = \epsilon(x)\Delta$. Then we would be done.

Note that B occurs $\iff$ $B_t(x) \not\subset C_y$ for all y. Assume $y \in X$ and $B_t(x) \cap C_y \neq \emptyset$. Then $B_t(x) \cap B_{\alpha\Delta}(y) \neq \emptyset$. So $d(x, y) \leq \alpha\Delta + t < \frac{\Delta}{2} + \frac{\Delta}{8} < \Delta$ by the triangle inequality. So $y \in B_\Delta(x)$. Let $b = |B_\Delta(x)|$ and $y_1 (= x), y_2, ..., y_b$ be the elements of $B_\Delta(x)$ in order of increasing distance to x.

Let $y \in X$ such that this necessary condition holds: $d(x, y) \leq \alpha\Delta + t$ and $\pi(y)$ is minimal in $<$. So $B_t(x)$ is disjoint from $\bigcup_{z:\pi(z)<\pi(y)} C_z = \bigcup_{z:\pi(z)<\pi(y)} B_{\alpha\Delta}(z)$ (by minimality). So $B_t(x) \subset C_y \iff B_t(x) \subset B_{\alpha\Delta(y)}$.

Now if B happens, then $B_t(x) \not\subset B_{\alpha\Delta}(y)$ and hence

$$d(x, y) > \alpha\Delta - t \geq \frac{\Delta}{4} - \frac{\Delta}{8} = \frac{\Delta}{8}.$$

Let $a = |B_{\Delta/8}(x)|$. Then $B_{\Delta/8}(x) = \{y_1, ..., y_a\}$. So the y above is y_k for some k with $a < k \leq b$.

So we proved that $B \subset \bigcup_{k=a+1}^b E_k$ where E_k is the event that $d(x, y_k) \leq \alpha\Delta + t$ with $\pi(y_k)$ is $<$-minimal with this property, and $d(x, y_k) > \alpha\Delta - t$.

Let $I_k = [d(x, y_k) - t, d(x, y_k) + t)$. Then $E_k \implies \alpha\Delta \in I_k$.

So $\mathbb{P}(B) \leq \sum_{k=a+1}^b \mathbb{P}(E_k) = \sum_{k=a+1}^b \mathbb{P}(E_k | \alpha\Delta \in I_k)\mathbb{P}(\alpha\Delta \in I_k)$. If $\alpha\Delta \in I_k$ then $d(x, y_j) \leq d(x, y_k) \leq \alpha\Delta + t$ for $1 \leq j \leq k$.

If in addition E_k occurs, we must have $\pi(y_k) < \pi(y_j)$ for all $j < k$. So

$$\begin{aligned}
\mathbb{P}(B) &\leq \sum_{k=a+1}^b \mathbb{P}(\pi(y_k) < \pi(y_j), \forall j < k | \alpha\Delta \in I_k)\mathbb{P}(\alpha\Delta \in I_k) \\
&= \sum_{k=a+1}^b \mathbb{P}(\pi(y_k) < \pi(y_j), \forall j < k)\mathbb{P}(\alpha\Delta \in I_k) \qquad \text{by independence of } \alpha, \pi \\
&\leq \sum_{k=a+1}^b \frac{1}{k}\frac{8t}{\Delta} \leq \frac{8t}{\Delta} \log\frac{b}{a} \leq \frac{1}{2} \qquad \text{if } t = \epsilon(x)\Delta.
\end{aligned}$$

So we have our (ϵ, Δ)-padded decomposition as desired. $\square$

Notation. For functions a, b on a set S and values in $\mathbb{R}^+$, $a \lesssim b$ means $\exists$ absolute constant C such that $a(s) \leq Cb(s)$ for all $s \in S$.

Theorem 3.6 (Gluing Lemma)**.** Let $1 \leq q < \infty, K > 0$. Assume $\forall \ell \in \mathbb{Z}, \exists$ a scale-2^ℓ embedding $f_\ell \colon X \to \ell_q$ of deficiency K and with $\|f_\ell(x)\| \leq 2^\ell$ for all $x \in X$. Then $c_q(X) \lesssim K^{1-1/q}(\log n)^{1/q}$.

Let's see how the Gluing Lemma implies Bourgain's Embedding Theorem.

Corollary 3.7 (Bourgain's Embedding Theorem)**.** $c_2(X) \lesssim \log n$.

Proof. By Theorem 5, there exists $(\epsilon, 2^\ell)$-padded decomposition for $X, \forall \ell \mathbb{Z}$ where $\epsilon(x) \geq C \frac{1}{\log n}$. By Lemma 2, for all $\ell \in \mathbb{Z}$ $\exists$ scale-2^ℓ embedding $f_\ell \colon X \to \ell_2$ with deficiency $K \leq C \log n$ and $\|f_\ell(x)\| \leq 2^\ell$ for all $x \in X$. By Theorem 6, $c_2(X) \leq C(\log n)^{1-1/2}(\log n)^{1/2} = C \log n$. $\square$

Now we will prove the Gluing Lemma. But first we need some notation.

Notation. For $x, y \in X, \ell \in \mathbb{Z}$, let

$$\gamma_\ell(x, y) = \begin{cases} x & \text{if } |B_{2^\ell}(x)| \geq |B_{2^\ell}(y)| \\ y & \text{otherwise.} \end{cases}$$

To prove the Gluing Lemma, we need two further lemmas.

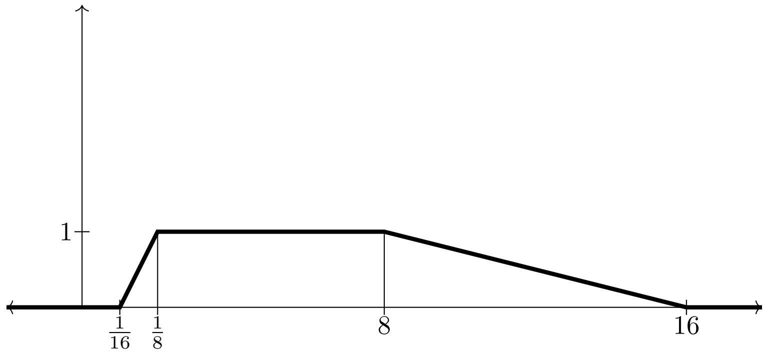
Lemma 3.8. Assume $\forall \ell \in \mathbb{Z}$ there exists 1-Lipschitz $h_\ell \colon X \to \ell_q$ $(1 \leq q < \infty)$ s.t. $\|h_\ell(x)\| \leq 2^\ell$ for all $x \in X$. Then there exists $H \colon X \to \ell_q$ s.t.

(i) $\mathrm{Lip}(H) \lesssim (\log n)^{1/q}$;

(ii) $\forall x, y \in X, \forall \ell \in \mathbb{Z}$ if $d(x, y) \in [2^\ell, 2^{\ell+1})$, then

$$\|H(x) - H(y)\| \geq \left(\log_2 \frac{|B_{2^{\ell+1}}(\gamma_{\ell-3}(x, y))|}{|B_{2^{\ell-3}}(\gamma_{\ell-3}(x, y))|} \right)^{1/q} \|h_\ell(x) - h_\ell(y)\|.$$

Proof. Let $\rho \colon \mathbb{R} \to \mathbb{R}^+$ be the function that is 0 on $(-\infty, \frac{1}{16}]$ then piecewise linear connecting $(\frac{1}{8}, 1)$, $(8, 1)$ and $(16, 0)$ and then 0 on $[16, +\infty)$. Note that $\mathrm{Lip}(\rho) \leq 16$.



Fix $t \in \{0, 1, 2, ..., \lceil \log_2 n \rceil - 1\}$. For $x \in X$ let

$$R(x, t) = \sup\{R : |B_R(x)| \leq 2^t\}.$$

This is 1-Lipschitz in x: given $x, y \in X$, if $|B_R(x)| \leq 2^t$, then $|B_{R-d(x,y)}(y)| \leq 2^t$ and so $R(y, t) \geq R - d(x, y)$. Take sup over R, $R(y, t) \geq R(x, t) - d(x, y)$.

Define

$$H_t \colon X \to \left(\bigoplus_{\ell \in \mathbb{Z}} \ell_q \right)_q \cong \ell_q$$

by

$$H_t(x) = \left(\rho\left(\frac{R(x, t)}{2^\ell} \right) h_\ell(x) \right)_{\ell \in \mathbb{Z}}.$$

Well-defined: Fix $x \in X$. Then $\rho\left(\frac{R(x,t)}{2^\ell}\right) = 0$ if $2^{\ell-4} \geq R(x, t)$ or $R(x, t) \geq 2^{\ell+4}$. Choose $m \in \mathbb{Z}$ s.t. $2^m \leq R(x, t) < 2^{m+1}$. Then $\rho\left(\frac{R(x,t)}{2^\ell}\right) = 0$ provided $2^{\ell-4} \geq 2^{m+1}$ or $2^m \geq 2^{\ell+4}$, so if $\ell \geq m + 5$ or $\ell \leq m - 4$. So $H_t(x)$ has ≤ 8 non-zero coordinates. So it is in ℓ_q.

Next we show H_t is Lipschitz with $\mathrm{Lip}(H_t) \leq 16 \times 17$. Note

$$\begin{aligned}
&\left\| \rho\left(\frac{R(x, t)}{2^\ell} \right) h_\ell(x) - \rho\left(\frac{R(y, t)}{2^\ell} \right) h_\ell(y) \right\| \\
&\leq \left| \rho\left(\frac{R(x, t)}{2^\ell} \right) - \rho\left(\frac{R(y, t)}{2^\ell} \right) \right| \|h_\ell(x)\| + \rho\left(\frac{R(y, t)}{2^\ell} \right) \|h_\ell(y) - h_\ell(x)\| \\
&\leq 16 \frac{1}{2^\ell} d(x, y) 2^\ell + d(x, y) \\
&= 17 d(x, y).
\end{aligned}$$

Since both $H_t(x), H_t(y)$ have ≤ 8 nonzero coordinates, we are done.

Now define

$$H \colon X \to \left(\bigoplus_{t=0}^{\lceil \log_2(n) \rceil - 1} \right)_q \cong \ell_q$$

by $H(x) = (H_t(x))_{t=0}^{\lceil \log_2 n \rceil - 1}$. It's clear that $\mathrm{Lip}(H) \lesssim (\log n)^{1/q}$. This proves (i).

To show (ii), fix $x, y \in X$, choose ℓ s.t. $d(x, y) \in [2^\ell, 2^{\ell+1})$. Then

$$\|H_t(x) - H_t(y)\| \geq \|h_\ell(x) - h_\ell(y)\| \qquad (*)$$

provided $\rho\left(\frac{R(x,t)}{2^\ell}\right) = \rho\left(\frac{R(y,t)}{2^\ell}\right) = 1$ which holds if $R(x, t), R(y, t) \in [2^{\ell-3}, 2^{\ell+3}]$.

This will follow if $|B_{2^{\ell-3}}(x)| \leq 2^t$, $|B_{2^{\ell+3}}(x)| > 2^t$ (same for y). So $(*)$ holds for all t such that

$$2^t \in [|B_{2^{\ell-3}}(x)|, |B_{2^{\ell+3}}(x)|) \cap [|B_{2^{\ell-3}}(y)|, |B_{2^{\ell+3}}(y)|).$$

WLOG $\gamma_{\ell-3}(x, y) = x$. Since $d(x, y) < 2^{\ell+1}$, $B_{2^{\ell+1}}(x) \subset B_{2^{\ell+3}}(y)$.

So $(*)$ holds if

$$2^t \in [|B_{2^{\ell-3}}(x)|, |B_{2^{\ell+1}}(x)|).$$

So

$$\|H(x) - H(y)\| = \left(\sum_t \|H_t(x) - H_t(y)\|\right)^{1/q}$$
$$\geq \left(\log_2 \frac{|B_{2^{\ell+1}}(x)|}{|B_{2^{\ell-3}}(x)|}\right)^{1/q} \|h_\ell(x) - h_\ell(y)\|.$$

□

Lemma 3.9. Let $1 \leq q < \infty$. Then there exists $H\colon X \to \ell_q$ such that

(i) $\mathrm{Lip}(H) \lesssim (\log n)^{1/q}$;

(ii) $\forall x, y \in X$, $\forall \ell \in \mathbb{Z}$, if $d(x,y) \in [2^\ell, 2^{\ell+1})$ and

$$\log_2 \left(\frac{|B_{2^{\ell-1}}(x)|}{|B_{2^{\ell-2}}(x)|}\right) < 1,$$

then $\|H(x) - H(y)\| \gtrsim d(x,y)$.

Proof. Fix $t \in \{1, 2, ..., \lceil \log_2 n \rceil\}$. Let W be a random subset of X where each $x \in X$ is placed in W independently at random with probability 2^{-t}. Let $\mathbb{P}_t$ be the resulting probability measure on $\mathcal{P}(X)$, the power set of X. So $\mathbb{P}_t(W) = 2^{-t|W|}(1-2^{-t})^{n-|W|}$ for any $W \subset X$. Note that $L_q(\mathcal{P}(X), \mathbb{P}_t) \cong \ell_q^{2^n}$ by

$$g \leftrightarrow \left(\mathbb{P}_t(W)^{1/q} g(W)\right)_{W \in \mathcal{P}(X)}.$$

Note

$$\|g\|_q^q = \int_{\mathcal{P}(X)} |g(W)|^q \, d\mathbb{P}_t(W)$$
$$= \sum_W \mathbb{P}_t(W) |g(W)|^q$$
$$= \left\| (\mathbb{P}_t(W)^{1/q} g(W))_W \right\|_q^q.$$

Define $H_t\colon X \to L_q(\mathcal{P}(X), \mathbb{P}_t) \cong \ell_q^{2^n}$ by $H_t(x) = (d(x, W))_W$. Then for all $x, y \in X$,

$$\|H_t(x) - H_t(y)\| = \left(\int_{\mathcal{P}(X)} |d(x,W) - d(y,W)|^q \, d\mathbb{P}_t(W)\right)^{1/q}$$
$$\leq d(x,y),$$

so H_t is 1-Lipschitz.

Define $H\colon X \to \left(\bigoplus_{t=1}^{\lceil \log_2 n \rceil} \ell_q^{2^n}\right)_q \hookrightarrow_{\cong} \ell_q$ by $H(x) = (H_t(x))_{t=1}^{\lceil \log_2 n \rceil}$. Then $\mathrm{Lip}(H) \lesssim (\log n)^{1/q}$. This shows (i).

To see (ii), fix $x, y \in X$, $\ell \in \mathbb{Z}$ such that $d(x,y) \in [2^\ell, 2^{\ell+1})$ and

$$\log_2 \left(\frac{|B_{2^{\ell-1}}(x)|}{|B_{2^{\ell-2}}(x)|}\right) < 1.$$

Fix $s \in \{1, 2, ..., \lceil \log_2 n \rceil\}$ such that $2^{s-1} \leq |B_{2^{\ell-1}}(x)| \leq 2^s$. Note $2^s \geq |B_{2^{\ell-2}}(x)| \geq 2^{s-2}$. Consider 4 events:

$$\begin{aligned} E_x &= \{W : d(x, W) \leq 2^{\ell-2}\} = \{W : W \cap B_{2^{\ell-2}}(x) \neq \emptyset\}, \\ F_x &= \{W : d(x, W) > 2^{\ell-1}\} = \{W : W \cap B_{2^{\ell-1}}(x) = \emptyset\}, \\ E_y &= \{W : d(y, W) \leq \frac{3}{2} 2^{\ell-2}\} = \{W : W \cap B_{\frac{3}{2}2^{\ell-2}}(y) \neq \emptyset\}, \\ F_y &= \mathcal{P}(X) \setminus E_y = \{W : W \cap B_{\frac{3}{2}2^{\ell-2}}(y) = \emptyset\}. \end{aligned}$$

Since $d(x, y) \geq 2^\ell$, $B_{2^{\ell-1}}(x) \cap B_{\frac{3}{2}2^{\ell-2}}(y) = \emptyset$, and hence any of E_x, F_x is independent of E_y, F_y.

Now we calculate the probabilities.

$$\begin{aligned} \mathbb{P}_s(E_x) &= 1 - (1 - 2^{-s})^{|B_{2^{\ell-2}}(x)|} \geq 1 - (1 - 2^{-s})^{2^{s-2}} \geq 1 - e^{-1/4} > 0, \\ \mathbb{P}_s(F_x) &= 1 - (1 - 2^{-s})^{|B_{2^{\ell-1}}(x)|} \geq 1 - (1 - 2^{-s})^{2^s} \geq (1 - \frac{1}{2})^2 = \frac{1}{4} > 0. \end{aligned}$$

So

$$\begin{aligned} \|H(x) - H(y)\| &\geq \|H_s(x) - H_s(y)\| \\ &= \left(\int_{\mathcal{P}(X)} |d(x, W) - d(y, W)|^q \, d\mathbb{P}_s(W) \right)^{1/q} \\ &\geq \left(\int_{E_x \cap F_y} + \int_{E_y \cap F_x} (\cdots) \right)^{1/q} \\ &\gtrsim (2^{(\ell-3)q} \mathbb{P}_s(F_y) + 2^{(\ell-3)q} \mathbb{P}_s(E_y))^{1/q} \\ &\gtrsim 2^{\ell+1} \geq d(x, y), \end{aligned}$$

as required. [Here we have used independence.] □

Proof of Theorem 6. Apply Lemma 8 with $h_\ell = f_\ell$ to get H, which we will call $F\colon X \to \ell_q$ such that $\mathrm{Lip}(F) \lesssim (\log n)^{1/q}$, and $\forall x, y \in X, \ell \in \mathbb{Z}$ if $d(x, y) \in [2^\ell, 2^{\ell+1})$, then

$$\|F(x) - F(y)\| \geq \left(\log_2 \frac{|B_{2^{\ell+1}}(\gamma_{\ell-3}(x, y))|}{|B_{2^{\ell-3}}(\gamma_{\ell-3}(x, y))|} \right)^{1/q} \|f_\ell(x) - f_\ell(y)\|.$$

Remember $\|f_\ell(x) - f_\ell(y)\| \geq \frac{1}{K} d(x, y)$.

From Theorem 5 and Lemma 2, we get $\forall \ell \in \mathbb{Z}$ a 1-Lipschitz $g_\ell\colon X \to \ell_q$ such that $\|g_\ell(x)\| \leq 2^\ell$ for all x and $\forall x, y \in X$, if $d(x, y) \in [2^\ell, 2^{\ell+1})$, then

$$\|g_\ell(x) - g_\ell(y)\| \gtrsim \left[16 + 16 \log \left(\frac{|B_{2^\ell}(x)|}{|B_{2^{\ell-3}}(x)|} \right) \right]^{-1} d(x, y).$$

Apply Lemma 8 with $h_\ell = g_\ell$ to get H which we call G here such that (i) and (ii) of Lemma 8 hold.

Let H be the function from Lemma 9. Define $\Phi\colon X \to (\ell_q \oplus \ell_q \oplus \ell_q)_q \cong \ell_q$ where $\Phi(x) = (F(x), G(x), H(x))$. Clearly we have $\mathrm{Lip}(\Phi) \lesssim (\log n)^{1/q}$. Fix

$x, y \in X$ with $d(x,y) \in [2^\ell, 2^{\ell+1})$ for some $\ell \in \mathbb{Z}$. Let $A = \log_2\left(\frac{|B_{2^{\ell+1}}(x)|}{|B_{2^{\ell-3}}(x)|}\right)$ and assume $\gamma_{\ell-3}(x,y) = x$. If $A < 1$ then by Lemma 9, $\|H(x) - H(y)\| \gtrsim d(x,y)$. If $A \geq 1$ then $\|F(x) - F(y)\| \geq A^{1/q}\frac{1}{K}d(x,y)$.

$$\|G(x) - G(y)\| \gtrsim \frac{A^{1/q}}{1+A}d(x,y).$$

Consider $A \geq K$ and $A \leq K$ to get $K^{-1+1/q}d(x,y)$ lower bound. So $\operatorname{dist}(\Phi) \lesssim K^{1-1/q}(\log n)^{1/q}$. □

4 Lower Bounds on Distortions, Poincaré Inequalities

In Section 3, we proved that $c_2(X) \lesssim \log|X|$ for any finite metric space X. Is this best possible? One might think that $c_2(X) \lesssim \sqrt{\log|X|}$.

Definition. For normed spaces X, Y we define the *Banach-Mazur distance*

$$d(X,Y) = \inf\{\|T\| \left\|T^{-1}\right\| : T\colon X \to Y \text{ is an onto isomorphism}\}.$$

[By convention $\inf \emptyset = \infty$.]

Note $1 \le \left\|T \circ T^{-1}\right\| \le \|T\| \left\|T^{-1}\right\|$, so $1 \le d(X,Y)$. Also $d(X,Z) \le d(X,Y) \times d(Y,Z)$ for all X, Y, Z. [If $T\colon X \to Y, S\colon Y \to Z$, then $\|ST\| \le \|S\| \|T\|$.] If $X \cong Y$ then $d(X,Y) = 1$. The converse is false in general.
Aside. Let $\mathcal{M}_n$ be the class of all n-dimensional normed spaces (we identify spaces that are isometrically isomorphic). On $\mathcal{M}_n$, $\log d$ is a metric and $\mathcal{M}_n$ is compact – the *Banach-Mazur compaction*.

Theorem (John's Lemma)**.** For any n-dimensional normed space X, $d(X, \ell_2^n) \le \sqrt{n}$.

Remark. (i) For all X, Y n-dimensional normed spaces, $d(X,Y) \le n$. [$\exists c > 0, \forall n, \operatorname{diam}(\mathcal{M}_n) \ge cn$ (Gluskin)].

(ii) For a general finite metric space X, the analogue of dimension, is $\log|X|$. This is to do with entropy. By analogy with John's Lemma, one might hope $c_2(X) \lesssim \sqrt{\log|X|}$.

Proof of John's Lemma. We can think of X as $\mathbb{R}^n$ with some norm $\|\cdot\|$. Let $K = B_X = \{x \in X : \|x\| \le 1\}$. This is a symmetric, convex body. [*Symmetric* means $\forall x \in K$, $-x \in K$, i.e. $K = -K$. *Body* means compact with nonempty interior.] Conversely, if K is a symmetric convex body, then $K = B_X$ where $X = (\mathbb{R}^n, \|\cdot\|)$ and $\|x\| = \inf\{t > 0 : x \in tK\}$. An *ellipsoid* is a subset $E \subset \mathbb{R}^n$ such that $E = T(B_{\ell_2^n})$ where $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is a linear bijection. Then $n^{-1/2}E \subset K \subset E \iff d(X, \ell_2^n) \le \sqrt{n}$ (first inequality is saying $\|T\| \le \sqrt{n}$, second inequality is saying $\left\|T^{-1}\right\| \le 1$, $T\colon \ell_2^n \to X$.) John's Lemma is equivalent to: for every symmetric convex body $K \subset \mathbb{R}^n$, there exists an ellipsoid, $n^{-1/2}E \subset K \subset E$.

By compactness, there exists an ellipsoid E of minimal volume such that $K \subset E$. We will show $n^{-1/2}E \subset K$. By applying a linear bijection, WLOG $E = B_{\ell_2^n}$ [by replacing K with $T^{-1}(K)$]. Assume for contradiction that $n^{-1/2}E \not\subset K$. Then there exists $z \in \partial K = S_X$ such that $\|z\|_2 < \frac{1}{\sqrt{n}}$. By Hahn-Banach, there exists a linear functional $f\colon \mathbb{R}^n \to \mathbb{R}$ such that $f(z) = 1$ and $|f(x)| \le 1$ for all $x \in K$. Let $H = \{x : f(x) + 1\}$. Then $z \in H$ and K is between H and $-H$. After applying a rotation, WLOG $H = \{x \in \mathbb{R}^n : x_1 = \frac{1}{c}\}$ for some $c > \sqrt{n}$ (as H contains a point with $\|\cdot\| < \frac{1}{\sqrt{n}}$). We still have $K \subset E = B_{\ell_2^n}$ and $K \subset \{x : |x_1| \le \frac{1}{c}\}$. Let $a > b > 0$, $E_{a,b} = \{x : a^2x_1^2 + \sum_{i=2}^n b^2x_i^2 \le 1\}$ which is the image of $B_{\ell_2^n}$ under the map with matrix diagonal $(\frac{1}{a}, \frac{1}{b}, \dots, \frac{1}{b})$. We have $\operatorname{vol}(E_{a,b}) = \frac{1}{ab^{n-1}} \operatorname{vol}(E)$. For $x \in K$, $a^2x_1^2 + \sum_{i=2}^n b^2x_i^2 = (a^2 - b^2)x_1^2 + \sum_{i=1}^n b^2x_i^2 \le \frac{a^2-b^2}{c^2} + b^2$ (using $K \subset E$). Need a, b such that $\frac{a^2-b^2}{c^2} + b^2 \le 1$ and

$ab^{n-1} > 1$. Then we would be done because $\text{vol}(E_{a,b}) < \text{vol}(E)$ and $K \subset E_{a,b}$ which contradicts the minimality of $\text{vol}(E)$.

For a given $0 < a < c$, set $b = \sqrt{\frac{c^2-a^2}{c^2-1}}$. Then $\frac{a^2-b^2}{c^2} + b^2 = 1$. Let $f(a) = ab^{n-1} = a\left(\frac{c^2-a^2}{c^2-1}\right)^{\frac{n-1}{2}}$. Then $f(1) = 1$,

$$\begin{aligned} f'(a) &= \left(\frac{c^2-a^2}{c^2-1}\right)^{\frac{n-1}{2}} + a\frac{n-1}{2}\frac{-2a}{c^2-1}\left(\frac{c^2-a^2}{c^2-1}\right)^{\frac{n-1}{2}} \\ &= \left(\frac{c^2-a^2}{c^2-1}\right)^{\frac{n-1}{2}-1}\left(\frac{c^2-a^2}{c^2-1} - \frac{(n-1)a^2}{c^2-1}\right) \\ &= \left(\frac{c^2-a^2}{c^2-1}\right)^{\frac{n-1}{2}-1}\left(\frac{c^2-na^2}{c^2-1}\right). \end{aligned}$$

Since $c^2 > n$, $f'(1) > 0$, there exists $a > 1$ such that $f(a) > f(1) = 1$. □

Definition. Let X, Y be metric spaces. A *Poincaré inequality* for functions $f\colon X \to Y$ is one of the form

$$\sum_{u,v\in X} a_{u,v}\Psi(d(f(u), f(v))) \geq \sum_{u,v\in X} b_{u,v}\Psi(d(f(u), f(v))), \qquad (*)$$

where a, b are $X \times X$ matrices, i.e. functions $a, b\colon X \times X \to \mathbb{R}^+$ of finite support, and Ψ is an increasing function $\mathbb{R}^+ \to \mathbb{R}^+$.

Define the *Poincaré ratio*

$$P_{a,b,\Psi}(X) = \frac{\sum_{u,v} b_{u,v}\Psi(d(u,v))}{\sum_{u,v} a_{u,v}\Psi(d(u,v))}, \qquad \text{whenever this is defined.}$$

Proposition. Let $1 \leq p < \infty$, $\Psi(t) = t^p$. Assume X, Y are metric spaces satisfying for some a, b the Poincaré inequality $(*)$ above for all functions $f\colon X \to Y$. Then $c_Y(X) \geq P_{a,b,\Psi}(X)^{1/p}$.

Proof. Let $f\colon X \to Y$ be a bilipschitz embedding [if there isn't any, then $c_Y(X) = \infty$]. Then

$$1 \geq \frac{\sum_{u,v} b_{u,v} d(f(u), f(v))^p}{\sum_{u,v} a_{u,v} d(f(u), f(v))^p} \geq \frac{1}{\text{dist}(f)^p}\frac{\sum_{u,v} b_{u,v} d(u,v)^p}{\sum_{u,v} a_{u,v} d(u,v)^p}$$

where the first inequality is by $(*)$. Hence $\text{dist}(f)^p \geq P_{a,b,\Psi}(X)^p$. Taking inf over all f gives the result. □

Example. In ℓ_2,

$$\|x_1 - x_3\|^2 + \|x_2 - x_4\|^2 \leq \|x_1 - x_2\|^2 + \|x_2 - x_3\|^2 + \|x_3 - x_4\|^2 + \|x_4 - x_1\|^2,$$

for all $x_1, x_2, x_3, x_4 \in \ell_2$. This is a Poincaré inequality for functions $C_4 \to \ell_2$. Hence by the proposition above, $c_2(C_4) \geq \sqrt{\frac{2^2+2^2}{4}} = \sqrt{2}$. This can be achieved by the obvious embedding. So $c_2(C_4) = \sqrt{2}$.

To show that there is always a Poincaré inequality that gets arbitrarily close to the distortion, we need Hahn-Banach separation theorems (see Section 4).

Hahn-Banach Separation Theorems

To study Poincaré inequalities, we need to use the Hahn-Banach Separation Theorems. This section is a digression from Metric Embeddings.

Let X be a real vector space. A functional $p\colon X \to \mathbb{R}$ is *positive homogeneous* if $p(tx) = tp(x), \forall t \geq 0, \forall x \in X$, and *subadditive* if $p(x+y) \leq p(x)+p(y), \forall x, y \in X$. For example, a seminorm or a norm on X.

Theorem 4.1. Let X, p be as above. Let Y be a subspace of X, $g\colon Y \to \mathbb{R}$ a linear map such that $g(y) \leq p(y), \forall y \in Y$. Then there exists a linear map $f\colon X \to \mathbb{R}$ such that $f|_Y = g$ and $f(x) \leq p(x)$ for all $x \in X$.

Proof. (This is similar to proof of Lemma 2.4). Let $P = \{(Z, h) : Z \leq X, h\colon Z \to \mathbb{R} \text{ linear}, Y \subset Z, h|_Y = g, h(z) \leq p(z), \forall z \in Z\}$. This is a poset with $(Z_1, h_1) \leq (Z_2, h_2) \iff Z_1 \subset Z_2$ and $h_2|_{Z_1} = h_1$. Note that $(Y, g) \in P$ so $P \neq \emptyset$. Given a chain $C = \{(Z_i, h_i) : i \in I\}$ in P (so C is linearly ordered) with $C \neq \emptyset$, then $Z = \bigcup_{i \in I} Z_i$ and $h\colon Z \to \mathbb{R}$ is defined by $h_{Z_i} = h_i, i \in I$ gives an upper bound (Z, h) for C. By Zorn's Lemma, P has a maximal element (W, k). We show that $W = X$, then we're done by taking $f = k$. Assume not. Fix $x_0 \in X \setminus W$ and let $W_1 = W + \mathbb{R}x_0$. Fix $\alpha \in \mathbb{R}$ and define $k_1\colon W_1 \to \mathbb{R}$ by $k_1(w + \lambda x_0) = k(w) + \lambda\alpha$ for $w \in W, \lambda \in \mathbb{R}$. We need α so that $k_1(w + \lambda x_0) \leq p(w + \lambda x_0$ for all $w \in W$ and $\lambda \in \mathbb{R}$. Then $(W, k) \lneqq (W_1, k_1)$, contradicting maximality of (W, k).

Since k_1 is linear and p is homogeneous, enough to get

$$k_1(w + x_0) \leq p(w + x_0), \qquad k_1(w - x_0) \leq p(w - x_0) \qquad \forall w \in W.$$

So we need

$$k(w) + \alpha \leq p(w + x_0), \qquad k(w) - \alpha \leq p(w - x_0) \qquad \forall w \in W.$$

So we need

$$k(z) - p(z - x_0) \leq \alpha \leq p(w + x_0) - k(w) \qquad \forall w, z \in W.$$

We need $k(z) - p(z - x_0) \leq p(w + x_0) - k(w), \forall w, z \in W$. Then $\alpha = \inf_{w \in W}(p(w + x_0) - k(w))$ will do. But $k(z) + k(w) = k(z + w) \leq p(z + w) = p(z - x_0 + w + x_0) \leq p(z - x_0) + p(w + x_0), \forall w, z \in W$. ☐

Corollary 4.2. Let X be a real normed space.

(i) If Y is a subspace and $g \in Y^*$ then there exists $f \in X^*$ s.t. $f|_Y = g$ and $\|f\| = \|g\|$. [Hahn-Banach Extension Theorem]

(ii) Given $x_0 \in X$, $x_0 \neq 0$, there exists $f \in S_{X^*}$ such that $f(x) = \|x_0\|$. [Norming functional for x_0]

Proof. (i) Let $p(x) = \|g\| \|x\|$ for $x \in X$. Then p is a seminorm. We have $g(y) \leq p(y)$ for all $y \in Y$. By Theorem 2, there exists a linear $f\colon X \to \mathbb{R}$ such that $f|_Y = g$ and $f(x) \leq \|g\| \|x\|$ for all $x \in X$. Apply this to $-x$ to get $-f(x) = f(-x) \leq \|g\| \|x\|$. So $|f(x)| \leq \|g\| \|x\|$ for all $x \in X$. So $f \in X^*$ and $\|f\| \leq \|g\|$. Since $f|_Y = g$, $\|f\| = \|g\|$.

(ii) Define $g\colon Y := \mathbb{R}x_0 \to \mathbb{R}$ by $g(\lambda x_0) = \lambda \|x_0\|$ for $\lambda \in \mathbb{R}$. Then $g \in Y^*$ and $\|g\| = 1$. So by (i), there exists $f \in S_{X^*}$ such that $f|_Y = g$, and so $f(x_0) = \|x_0\|$.

☐

Remark 4.3. If Z is a complex vector space, let $Z_{\mathbb{R}}$ be Z viewed as a real vector space. Then for a complex normed space X, the map $(X^*)_{\mathbb{R}} \to (X_{\mathbb{R}})^*$, $f \mapsto \operatorname{Re} f$ is an isometric isomorphism. Thus (i) follows in the complex case.

Given a normed space X and a convex subset C of X with $0 \in \operatorname{Int} C$, the *Minkowski functional* of C is $\mu_C \colon X \to \mathbb{R}$ defined by

$$\mu_C(x) = \inf\{t > 0 : x \in tC\}.$$

This is well-defined: given $x \in X$, $\frac{x}{n} \to 0 \in \operatorname{Int} C$, so $\exists n, \frac{x}{n} \in C$, i.e. $x \in nC$. Example: If $C = B_X$, then $\mu_C = \|\cdot\|$ as $x \in tB_X \iff \|x\| \le t$.

Lemma 4.4. Let X, C be as above. Then μ_C is positive homogeneous and subadditive. Moreover,

$$\{x \in X : \mu_C(x) < 1\} \subset C \subset \{x \in X : \mu_C(x) \le 1\},$$

with equality in the first inclusion if C is open, and with equality in the second inclusion if C is closed.

Proof. For positive homogeneity, e need $\mu_C(tx) = t\mu_C(x)$ for all $t \ge 0$ and $x \in X$. For $t = 0$, we need $\mu_C(0) = 0$. This is true since $x \in tC$ for all $t > 0$. If $t > 0$, then for any $s > 0$, $tx \in sC \iff x \in \frac{s}{t}C$, so $\mu_C(tx) = t\mu_C(x)$.

For subadditivity, fix $x, y \in X$ and let $s > \mu_C(x), t > \mu_C(y)$. Then by definition, there exists s', $\mu_C(x) \le s' < s$ such that $x \in s'C$. Then $\frac{x}{s} = \frac{s'}{s}\frac{x}{s'} + (1 - \frac{s'}{s})0 \in C$, since C is convex. So $x \in sC$. Also $y \in tC$. Thus $\frac{x+y}{s+t} = \frac{s}{s+t}\frac{x}{s} + \frac{t}{s+t}\frac{y}{t} \in C$. This shows $\mu_C(x+y) \le s + t$. Taking inf over all s, t we get subadditivity.

If $1 > \mu_C(x)$, then by above $x \in C$, showing the first inclusion. If $x \in C$, then $\mu_C(x) \le 1$ by definition, showing the second inclusion. Assume C is open. If $x \in C$, then since $(1 + \frac{1}{n})x \to x$ and C is open, then there exists n with $(1 + \frac{1}{n})x \in C$, i.e. $x \in \frac{n}{n+1}C$, so $\mu_C(x) \le \frac{n}{n+1} < 1$. Now assume C is closed and $\mu_C(x) \le 1$. Then $\mu_C(\frac{n}{n+1}x) \le \frac{n}{n+1} < 1$ so $\frac{n}{n+1}x \in C$ for all $n \in \mathbb{N}$. Since $\frac{n}{n+1}x \to x$ and C is closed, $x \in C$. □

Theorem 4.5. Let X be a real normed space, and let C be an open convex subset of X with $0 \in C$. For $x_0 \in X \setminus C$, there exists $f \in X^*$ such that $f(x) < f(x_0)$ for all $x \in C$. (Note that $f \neq 0$.)

Proof. Define $Y = \mathbb{R}x_0$ and $g \colon Y \to \mathbb{R}$ by $g(\lambda x_0) = \lambda\mu_C(x_0)$. Then g is linear and for $\lambda \ge 0$, $g(\lambda x_0) = \lambda\mu_C(x_0) = \mu_C(\lambda x_0)$, and for $\lambda < 0$, $g(\lambda x_0) = \lambda\mu_C(x_0) \le 0 \le \mu_C(\lambda x_0)$. By Lemma 4 and Theorem 2, there exists a linear map $f \colon X \to \mathbb{R}$ such that $f|_Y = g$ and $f(x) \le \mu_C(x)$ for all $x \in X$. Since $x_0 \notin C$, $\mu_C(x_0) \ge 1$. So for all $x \in C$, $f(x) \le \mu_C(x) < 1 \le \mu_C(x_0) = f(x_0)$ [here we used C is open]. Since $0 \in C$, C open, $\exists \delta > 0$ such that $\delta B_X \subset C$. So $f(x) \le 1$ on δB_X, but this is symmetric, so $|f(x)| \le 1$. So $f \in X^*$. □

Remark. If Lemma 4, if C is symmetric, then μ_C is a seminorm. If in addition, C is bounded, then μ_C is a norm [we used this in the proof of Theorem 1].

Corollary 4.6 (The Hahn-Banach Separation Theorems)**.** Let A, B be non-empty, disjoint convex sets in a normed space X.

(i) If A is open, then there exists $f \in X^*$ and $\alpha \in \mathbb{R}$ such that $f(x) < \alpha \leq f(y)$ for all $x \in A$, for all $y \in B$.

(ii) If A is compact, and B is closed, then $\exists f \in X^*$ and $\alpha \in \mathbb{R}$ such that $\sup_A f < \alpha < \inf_B f$.

Remark. In both cases, the hyperplane $\{x \in X : f(x) = \alpha\}$ separates A and B.

Proof. (i) Fix $a_0 \in A, b_0 \in B$. Let $C = A - B - a_0 + b_0$, $x_0 = -(a_0 - b_0)$. Then C is convex and open, $0 \in C$ and $x_0 \notin C$ since $A \cap B = \emptyset$. By Theorem 5, $\exists f \in X^*$ such that $f(x) < f(x_0)$ for all $x \in C$. So $f(x - y + x_0) < f(x_0)$ for all $x \in A, y \in B$, i.e., $f(x) < f(y)$ for all $x \in A, y \in B$. Let $\alpha = \inf_B f$. So certainly we have $f(y) \geq \alpha$ for all $y \in B$. Also, $f(x) \leq \alpha$ for all $x \in A$. Since $f \neq 0$, we can fix $u \in X$ such that $f(u) > 0$. For $x \in A$, since A is open, $\exists n \in \mathbb{N}$ such that $x + \frac{1}{n}u \in A$. Then $f(x) < f(x + \frac{1}{n}u) \leq \alpha$.

(ii) For $x \in A$, $d(x, B) > 0$ since B is closed and $x \notin B$. Since A is compact, $\delta = \inf_{x \in A} d(x, B) > 0$. Then $A' = \{x \in X : d(x, A) < \delta\}$ is an open, convex set with $A' \cap B = \emptyset$. [If $d(x, A), d(y, A) < \delta$ then $\exists u, v \in A$, $\|x - u\|, \|y - v\| < \delta$ and then $\forall t \in (0, 1)$,

$$\|((1-t)x + ty) - ((1-t)u + tv)\| < \delta,$$

$(((1-t)u + tv) \in A)$, so $(1-t)x + ty \in A'$]. By (i), $\exists f \in X^*, \exists \beta \in \mathbb{R}$ such that $f(x) < \beta \leq f(y)$ for all $x \in A', y \in B$. As A is compact, $\sup_A f < \beta \leq \inf_B f$.

□

Poincaré Inequalities

Now we can show that Poincaré inequalities are worth studying because they get arbitrarily close to the distortion of f.

Theorem 4.7. Let $1 \leq p < \infty$ and X be a finite metric space. Then

$$c_p(X) = \sup \left(P_{a,b,t^p}(X)\right)^{1/p},$$

where the sup is over all non-negative, non-trivial $X \times X$ matrices a, b for which the Poincaré inequality

$$\sum_{u,v \in X} a_{u,v} \|f(u) - f(v)\|_p^p \geq \sum_{u,v \in X} b_{u,v} \|f(u) - f(v)\|_p^p \qquad (*)$$

holds.

Proof. From Proposition 2, $c_p(X) \geq \sup \left(P_{a,b,t^p}(X)\right)^{1/p}$. Taking $a_{u,v} = b_{u,v} = 1$ for all u, v, $(*)$ holds, and $P_{a,b,t^p}(X) = 1$, so if $c_p(X) = 1$ then we are done.

Now assume $1 < C < c_p(X)$. Let $X = \{x_1, ..., x_n\}$. Let

$$B = \left\{ \left(\|f(x_i) - f(x_j)\|_p^p \right)_{1 \leq i < j \leq n} : f \colon X \to L_p \right\} \subset \mathbb{R}^N,$$

where $N = \binom{n}{2}$. From proof of Theorem 2.7, we know B is a cone, and in particular, B is convex. Also $B \neq \emptyset$ because $0 \in B$. Let

$$A = \left\{ (\theta_{ij})_{1 \le i < j \le n} \in \mathbb{R}^N : \exists r > 0, r d(x_i, x_j)^p < \theta_{ij} < r C^p d(x_i, x_j)^p, \forall i, j \right\}.$$

Then A is open, convex and non-empty since $C > 1$. Since $C < c_p(X)$, we have $A \cap B = \emptyset$. By Corollary 6, there exists a linear map $c\colon \mathbb{R}^N \to \mathbb{R}$ and $\alpha \in \mathbb{R}$ such that $c(\theta) < \alpha \le c(\varphi)$ for all $\theta \in A, \varphi \in B$. We have $c = (c_{ij})_{1 \le i < j \le n}$ where $c(\theta) = \sum_{1 \le i < j \le n} c_{ij} \theta_{ij}$. Since $0 \in B$, $\alpha \le 0$. By continuity, $c(\theta) \le \alpha$ for all $\theta \in \bar{A}$, and $0 \in \bar{A}$, so $0 \le \alpha$. Hence $\alpha = 0$. So $c(\theta) \le 0 \le c(\varphi)$ for all $\theta \in \bar{A}, \varphi \in B$. Let $a_{ij} = \max(c_{ij}, 0)$, $b_{ij} = \max(-c_{ij}, 0)$. So $c_{ij} = a_{ij} - b_{ij}$. We have

$$\sum c_{ij} \left\| f(x_i) - f(x_j) \right\|_p^p \ge 0,$$

for all $f\colon X \to L_p$, i.e.

$$\sum_{1 \le i < j \le n} a_{ij} \left\| f(x_i) - f(x_j) \right\|_p^p \ge \sum_{1 \le i < j \le n} b_{ij} \left\| f(x_i) - f(x_j) \right\|_p^p,$$

for all $f\colon X \to L_p$.

Let

$$\theta_{ij} = \begin{cases} C^p d(x_i, x_j)^p & \text{if } c_{ij} \ge 0, \\ d(x_i, x_j)^p & \text{if } c_{ij} < 0. \end{cases}$$

Then $\theta = (\theta_{ij}) \in \bar{A}$, so

$$0 \ge c(\theta) = \sum_{ij} a_{ij} C^p d(x_i, x_j)^p - \sum_{ij} b_{ij} d(x_i, x_j)^p.$$

Thus $P_{a,b,t^p}(X) \ge C^p$. □

Hamming Cube

Recall $H_n = \{0,1\}^n$, which is a graph: $x = (x_i), y = (y_i)$ are joined by an edge $\iff$ $x_i \neq y_i$ for exactly one value of i. So H_n is a metric space with the graph distance d:

$$d(x, y) = \sum_{i=1}^n |x_i - y_i|.$$

So H_n is isometrically a subset of ℓ_1^n.

H_n is also a probability space with the uniform distribution μ: $\mu(\{x\}) = 2^{-n}$.

We think of $\{0,1\}$ as the field $\mathbb{F}_2$. Then H_n is the n-dimensional vector space $\mathbb{F}_2^n$ over $\mathbb{F}_2$. So in particular, H_n is an abelian group.

Notation. Let $(e_i)_{i=1}^n$ is the standard basis of $H_n = \mathbb{F}_2^n$. For $j = 1, ..., n$, let $r_j\colon H_n \to \mathbb{R}$, $r_j(x) = (-1)^{x_j}$. This is the jth *Rademacher function*. Note that $r_1, ..., r_n$ are iid random variables on (H_n, μ) with $\{\pm 1\}$-valued $\texttt{Rademacher}(\frac{1}{2})$ distribution. For $A \subset \{1, ..., n\}$, we define $w_A\colon H_n \to \mathbb{R}$, $w_A = \prod_{j \in A} r_j$. These are the *Walsh functions*. These are the characters of H_n, i.e. abelian group homomorphisms $H_n \to \mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$.

Lemma 4.8. The Walsh functions form an orthonormal basis of $L_2(H_n, \mu)$.

Proof. We have $r_j^2 = 1$, so for $A, B \subset \{1, ..., n\}$, $w_A w_B = \prod_{j \in A} r_j \prod_{k \in B} r_k = \prod_{j \in A \triangle B} r_j = w_{A \triangle B}$. So if $A = B$, $\langle w_A, w_A \rangle = \int_{H_n} w_\emptyset \, d\mu = 1$. If $A \neq B$, by independence, $\langle w_A, w_B \rangle = \int_{H_n} w_{A \triangle B} \, d\mu = \prod_{j \in A \triangle B} \int_{H_n} r_j \, d\mu = 0$. Alternatively, shifting is a measure-preserving transformation. Fix $j \in A \triangle B$, $\int_{H_n} w_{A \triangle B}(x) \, d\mu(x) = \int_{H_n} w_{A \triangle B}(x + e_j) \, d\mu(x) = -\int_{H_n} w_{A \triangle B}(x) \, d\mu(x)$. We're done as $\dim L_2(H_n, \mu) = 2^n$. □

Definition. For $f \colon H_n \to \mathbb{R}$, we let $\hat{f}_A = \langle f, w_A \rangle = \int_{H_n} f w_A \, d\mu$ for $A \subset \{1, ..., n\}$. These are the *Fourier coefficients* of f with respect to this orthonormal basis. More generally, for a Banach space X and $f \colon H_n \to X$, we define $\hat{f}_A = \int_{H_n} f(x) w_A(x) \, d\mu(x)$, $A \subset \{1, ..., n\}$. Normally this would involve the Bochner integral, but here everything is finite, so this is just a summation.

Lemma 4.9. (a) For any $f \in L_2(H_n, \mu)$ we have

$$f(x) = \sum_{A \subset \{1, ..., n\}} \hat{f}_A w_A(x), \qquad x \in H_n,$$

$$\int_{H_n} |f(x)|^2 \, d\mu(x) = \sum_{A \subset \{1, ..., n\}} |\hat{f}_A|^2, \qquad \text{Parseval's identity}$$

(b) If X is a Banach space, then for all $f \colon H_n \to X$ we have

$$f(x) = \sum_{A \subset [n]} \hat{f}_A w_A(x), \qquad x \in H_n.$$

If in addition X is a Hilbert space, then

$$\int_{H_n} \|f(x)\|^2 \, d\mu(x) = \sum_{A \subset [n]} \left\| \hat{f}_A \right\|^2 \qquad \text{Parseval's identity.}$$

Proof. (a) Follows from Lemma 8. (b) Fix $\varphi \in X^*$. Then

$$\varphi(\hat{f}_A) = \int_{H_n} \varphi(f(x)) w_A(x) \, d\mu(x) = \widehat{\varphi \circ f}_A \qquad \forall A \subset [n].$$

So for any $x \in H_n$, we have, by (a),

$$\varphi(f(x)) = \sum_A \widehat{\varphi \circ f}_A w_A(x) = \varphi(\sum_A \hat{f}_A w_A(x)).$$

This holds for all $\varphi \in X^*$, so by Hahn-Banach, $f(x) = \sum_A \hat{f}_A w_A(x)$. True for all $x \in H_n$.

If X is a Hilbert space, then WLOG $\dim X < \infty$. Fix an orthonormal basis $v_1, ..., v_k$ of X. Then for $1 \leq j \leq k$, let $f_j(x) = \langle f(x), v_j \rangle$. By above, taking $\varphi(u) = \langle u, v_j \rangle$, $\hat{f_j}_A = \langle \hat{f}_A, v_j \rangle$ Then by Parseval in X, in $L_2(H_n, \mu)$, and in X respectively,

$$\begin{aligned} \int_{H_n} \|f(x)\|^2 \, d\mu(x) &= \int_{H_n} \sum_{j=1}^k |f_j(x)|^2 \, d\mu(x) = \sum_{j=1}^k \sum_A |\hat{f_j}_A|^2 \\ &= \sum_A \sum_j |\langle \hat{f}_A, v_j \rangle|^2 = \sum_A \|\hat{f}_A\|^2. \end{aligned}$$

□

Definition. For each $1 \leq j \leq n$, we define a *difference operator* ∂_j as follows. For a Banach space X and $f\colon H_n \to X$, let $\partial_j f\colon H_n \to X$ be defined as

$$(\partial_j f)(x) = \frac{f(x+e_j) - f(x)}{2}.$$

Lemma 4.10. (i) For $1 \leq j \leq n$, $A \subset [n]$,

$$\partial_j w_A(x) = \begin{cases} -w_A(x) & j \in A \\ 0 & j \notin A. \end{cases}$$

(ii) For a Banach space X and $f\colon H_n \to X$,

$$\widehat{\partial_j f}_A = \begin{cases} -\hat{f}_A & j \in A \\ 0 & j \notin A. \end{cases}$$

(iii) If X is a Hilbert space, then for $f\colon H_n \to X$,

$$\sum_{j=1}^n \int_{H_n} \|\partial_j f(x)\|^2 \, d\mu(x) = \sum_A |A| \|\hat{f}_n\|^2.$$

Proof. (i) We have

$$r_i(x+e_j) = \begin{cases} -r_i(x) & j = i \\ r_i(x) & j \neq i. \end{cases}$$

So

$$w_A(x+e_j) = \prod_{i \in A} r_i(x+e_j) = \begin{cases} -w_A(x) & j \in A \\ w_A(x) & j \notin A. \end{cases}$$

Hence result follows.

(ii) This is integration by parts:

$$\begin{aligned}
(\widehat{\partial_j f})_A &= \int_{H_n} (\partial_j f)(x) w_A(x) \, d\mu(x) \\
&= \frac{1}{2} \int_{H_n} f(x+e_j) w_A(x) \, d\mu(x) - \frac{1}{2} \int_{H_n} f(x) w_A(x) \, d\mu(x) \\
&= \frac{1}{2} \int_{H_n} f(x) w_A(x+e_j) \, d\mu(x) - \frac{1}{2} \int_{H_n} f(x) w_A(x) \, d\mu(x) \\
&= \int_{H_n} f(x) (\partial_j w_A)(x) \, d\mu(x) \\
&= \begin{cases} -\hat{f}_A & j \in A \\ 0 & j \notin A. \end{cases}
\end{aligned}$$

(iii) We use Parseval:

$$\begin{aligned}
\sum_{j=1}^n \int_{H_n} \|\partial_j f(x)\|^2 \, d\mu(x) &= \sum_{j=1}^n \sum_A \|(\widehat{\partial_j f})_A\|^2 \\
&= \sum_A \sum_j \|(\widehat{\partial_j f})_A\|^2 \\
&= \sum_A |A| \|\hat{f}_A\|^2,
\end{aligned}$$

as required. □

Theorem 4.11 (Poincaré inequality for L_2-valued functions on H_n)**.** Let $e = e_1 + e_2 + ... + e_n = (1, 1, ..., 1)$. Then for all $f\colon H_n \to L_2$, we have

$$\int_{H_n} \|f(x+e) - f(x)\|^2 \, d\mu(x) \leq 4 \sum_{j=1}^{n} \int_{H_n} \|(\partial_j f)(x)\|^2 \, d\mu(x).$$

Proof. For $A \subset [n]$, $w_A(x+e) = (-1)^{|A|} w_A(x)$. So

$$\begin{aligned}
&\int_{H_n} \|f(x+e) - f(x)\|^2 \, d\mu(x) \\
&\qquad = \int_{H_n} \left\| \sum_A \hat{f}_A w_A(x+e) - \sum_A \hat{f}_A w_A(x) \right\|^2 d\mu(x) \quad \text{(by Lemma 9)} \\
&\qquad = 4 \int_{H_n} \left\| \sum_{A:|A| \text{ odd}} \hat{f}_A w_A(x) \right\|^2 d\mu(x) \\
&\qquad = 4 \sum_{A:|A| \text{ odd}} \|\hat{f}_A\|^2 \quad \text{(by Lemma 9)} \\
&\qquad \leq 4 \sum_{|A|} |A| \|\hat{f}_A\|^2 \\
&\qquad = 4 \sum_{j=1}^{n} \int_{H_n} \|(\partial_j f)(x)\|^2 \, d\mu(x).
\end{aligned}$$

□

Corollary 4.12. $c_2(H_n) = \sqrt{n}$.

Remark. $|H_n| = 2^n$, so $c_2(H_n) = \sqrt{\log |H_n|}$. Compare with the upper bound $c_2(H_n) \lesssim \log |H_n|$ in Bourgain's embedding theorem.

Proof of Corollary 12. $H_n \subset \ell_2^n$ in the obvious way which gives $c_2(H_n) \leq \sqrt{n}$. By Proposition 2, a lower bound on $c_2(H_n)$ is obtained from the Poincaré ratio

$$\frac{\int_{H_n} d(x+e, x)^2 \, d\mu(x)}{4 \sum_{j=1}^{n} \int_{H_n} \frac{d(x+e_j, x)^2}{4} \, d\mu(x)} = \frac{n^2}{n} = n,$$

so $c_2(H_n) \geq \sqrt{n}$. □

From now on, think of H_n as the n-dimensional vector space $\mathbb{F}_2^n$ over $\mathbb{F}_2$.

Theorem 4.13. For every $f\colon \mathbb{F}_2^n \to L_2$ we have the Poincaré inequality:

$$\int_{\mathbb{F}_2^n \times \mathbb{F}_2^n} \|f(x) - f(y)\|_{L_2}^2 \, d\mu(x) \, d\mu(y) \leq \frac{2}{\max\{|A| : A \neq \emptyset, \hat{f}_A \neq 0\}} \sum_{j=1}^{n} \int_{\mathbb{F}_2^n} \|\partial_j f(x)\|^2 \, d\mu(x).$$

Proof. Without loss of generality, after replacing f with $f - \hat{f}_\emptyset w_\emptyset$, can assume $\hat{f}_\emptyset = 0$ (recall $w_\emptyset = 1$). Then by Parseval,

$$\begin{aligned} LHS &= \int_{\mathbb{F}_2^n \times \mathbb{F}_2^n} \|f(x)\|^2 + \|f(y)\|^2 - 2\langle f(x), f(y)\rangle \, d\mu(x)\, d\mu(y) \\ &= 2\sum_A \|\hat{f}_A\|^2 - 2\int_{\mathbb{F}_2^n} \left\langle \int_{\mathbb{F}_2^n} f(x)\, d\mu(x), f(y) \right\rangle d\mu(y) \\ &= 2\sum_A \|\hat{f}_A\|^2 - 2\int_{\mathbb{F}_2^n} \langle \hat{f}_\emptyset, f(y)\rangle \, d\mu(y) \\ &= 2\sum_A \|\hat{f}_A\|^2. \end{aligned}$$

By Lemma 10,

$$\sum_{j=1}^n \int_{\mathbb{F}_2^n} \|\partial_j f(x)\|^2 \, d\mu(x) = \sum_A |A| \|\hat{f}_A\|^2 \geq \min\{|A| : A \neq \emptyset, \hat{f}_A \neq 0\} \sum_A \|\hat{f}_A\|^2.$$

□

Definition. A *linear code* of $\mathbb{F}_2^n$ is a subspace C of $\mathbb{F}_2^n$. We let $d(C) = \min\{d(x, 0) : x \in C, x \neq 0\} = d(0, C \setminus \{0\})$. For $x = (x_i), y = (y_i)$ in $\mathbb{F}_2^n$, let $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ (operations in $\mathbb{F}_2^n$). This is a symmetric bilinear form, but $\langle x, x\rangle = 0$ does not imply $x = 0$. For a subset $S \subset \mathbb{F}_2^n$, let $S^\perp = \{x \in \mathbb{F}_2^n : \langle x, s\rangle = 0, \forall s \in S\}$.

Linear Codes

Lemma 4.14. For a linear code C, $\dim C + \dim C^\perp = n$ and $C^{\perp\perp} = C$.

Proof. Let $m = \dim C$ and $v_1, ..., v_m$ be a basis of C. Define $\theta \colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ by $\theta(x) = (\langle x, v_i\rangle)_{i=1}^m$. Then $\ker\theta = C^\perp$ and so $n = \dim C^\perp + \dim \operatorname{im}\theta$. We need θ to be onto. For $1 \leq j \leq m$, let $f \colon \mathbb{F}_2^n \to \mathbb{F}_2$ be a linear map such that $f(v_i) = \delta_{ij}$ (Kronecker delta). Let $y_j = f(e_j)$ for $1 \leq j \leq n$ and $y = (y_j)$. Then $f(x) = \sum_{j=1}^n x_j f(e_j) = \langle x, y\rangle$, so $\theta(y) = (f(v_j))_{j=1}^n = i$th standard basis vector of $\mathbb{F}_2^m$. So $n = \dim C^\perp + m = \dim C^\perp + \dim C$. For the final part: from definition, $C \subset C^{\perp\perp}$, and $\dim C^{\perp\perp} = n - \dim C^\perp = \dim C$, so $C = C^{\perp\perp}$. □

Lemma 4.15. There exists $\delta \in (0, \frac{1}{2}), \exists N \in \mathbb{N}, \forall n \geq N, (m+1)\binom{n}{m} \leq 2^{n/8}$ where $m = \lfloor \delta n \rfloor$.

Proof. First choose $\delta \in (0, \frac{1}{2})$ such that $\delta(2 + \log(\frac{2}{\delta})) < \frac{\log 2}{8}$. Then choose $N \in \mathbb{N}$ such that $\lfloor \delta n \rfloor \geq \frac{\delta n}{2}, \forall n \geq N$. Let $n \geq N$ and $m = \lfloor \delta n \rfloor$. If $m = 0$ then we are done, so assume $m \geq 1$. Then $\binom{n}{m} = \frac{n(n-1)(n-2)\dots(n-m+1)}{m!} \leq \frac{n^m}{m!}$. For the denominator use $\log(m!) = \sum_{j=1}^m \log(j) \geq \int_1^m \log x \, dx = [x\log x - x]_1^m = m\log m - m + 1 \geq m\log m - m$. So $\binom{n}{m} \leq (\frac{en}{m})^m$ and $(m+1)\binom{n}{m} \leq (m+1)(\frac{en}{m})^m$.

Now

$$\begin{aligned}\log\left((m+1)\binom{n}{m}\right) &\le \log(m+1) + m(1+\log(n/m)) \\ &\le m(2+\log(n/m)) && (\log x \le x-1, \forall x>0) \\ &\le \delta n(2+\log(2/\delta)) && (\frac{\delta n}{2} \le m = \lfloor \delta n \rfloor \le \delta n) \\ &\le \frac{\log 2}{8} n.\end{aligned}$$

Thus $(m+1)\binom{n}{m} \le 2^{n/8}$. □

Lemma 4.16. *$\exists \alpha > 0, \forall n \in \mathbb{N}, \exists$ linear code C in $\mathbb{F}_2^n$ with $\dim C \ge \frac{n}{4}$ and $d(C) \ge \alpha n$.*

Proof. Let δ, N be as in Lemma 15. If $1 \le n \le N$, choose any C with $\dim C \ge \frac{n}{4}$. Then $d(C) \ge 1 \ge \frac{1}{N} n$. Now let $n > N$. We show there exists a linear code C in $\mathbb{F}_2^n$ such that $\dim C \ge \frac{n}{4}$ and $d(C) \ge \delta n$. So $\alpha = \min(\frac{1}{N}, \delta)$ will do.

We choose C greedily. Assume that for some k, $1 \le k < \frac{n}{4}$ we have a linear code C_k with $\dim C_k = k$ and $d(C_k) \ge \delta n$. For $k = 1$ this holds. We seek a suitable $x \in \mathbb{F}_2^n \setminus C_k$ such that putting $C_{k+1} = \operatorname{span}(C_k \cup \{x\}) = C_k \cup (C_k + x)$, we have $d(C_{k+1}) \ge \delta n$. Once we find such x, we continue inductively. Taking $C = C_{\lceil n/4 \rceil}$ will complete the proof.

We estimate from above the number of unsuitable vectors x. For $v \in C_k$,

$$\begin{aligned}|\{x : d(v+x, 0) < \delta n\}| &= |\{x : d(x,0) < \delta n\}| \\ &= \sum_{0 \le \ell < \delta n} \binom{n}{\ell} \\ &\le (m+1)\binom{n}{m},\end{aligned}$$

where $m = \lfloor \delta n \rfloor$. Note in the range $0 \le \ell \le \frac{n}{2}$, $\binom{n}{\ell}$ is increasing, and $\delta < \frac{1}{2}$. It follows that

$$\begin{aligned}|\{x \in \mathbb{F}_2^n : \exists v \in C_k, d(x+v,0) < \delta n\}| &= \left| \bigcup_{v \in C_k} \{x \in \mathbb{F}_2^n : d(x+v,0) < \delta n\} \right| \\ &\le 2^k (m+1)\binom{n}{m}.\end{aligned}$$

If $2^k(m+1)\binom{n}{m} < 2^n - 2^k$ then there is a suitable x, i.e. we need $(m+1)\binom{n}{m} < 2^{n-k} - 1$. Now $2^{n-k} - 1 > 2^{3n/4} - 1 \ge 2^{n/8}$, so we are done by choice of δ, N. □

From now on, C will be an arbitrary linear code in $\mathbb{F}_2^n$. Let $q \colon \mathbb{F}_2^n \to \mathbb{F}_2^n / C^\perp$ be the quotient map. Let $\tilde{\mu}$ be the image measure induced by μ and q: $\tilde{\mu}(E) = \mu(q^{-1}(E))$. Let ρ be the quotient metric on $\mathbb{F}_2^n / C^\perp$: $\rho(q(x), q(y)) = d(x + C^\perp, y + C^\perp) = d(x-y, C^\perp) = \min_{v \in C^\perp} d(x-y, v)$.

Lemma 4.17. *For every $h \colon \mathbb{F}_2^n / C^\perp \to L^1$ and for every $A \subset [n]$ with $A \ne \emptyset$ and $|A| < d(C)$ we have $\hat{f}_A = 0$ where $f = h \circ q$.*

Proof. Let $v = \sum_{i \in A} e_i$. Then $v \neq 0$ since $A \neq \emptyset$ and $d(v, 0) = |A| < d(C)$¿ So $v \notin C = C^{\perp\perp}$ (Lemma 14). So $\exists w \in C^\perp$ such that $\langle v, w\rangle \neq 0$, i.e. $\langle v, w\rangle = 1$. Now

$$\begin{aligned}
\hat{f}_A &= \int_{\mathbb{F}_2^n} f(x) w_A(x)\, d\mu(x) \\
&= \int_{\mathbb{F}_2^n} f(x+w) w_A(x+w)\, d\mu(x) \\
&\qquad\qquad \text{(translation invariance of } \mu) \\
&= \int_{\mathbb{F}_2^n} f(x) \prod_{j \in A} r_j(x+w)\, d\mu(x) \\
&\qquad\qquad (w \in C^\perp \text{ so } f(x+w) = hq(x+w) = hq(x) = f(x)) \\
&= \int_{\mathbb{F}_2^n} f(x) \prod_{j \in A} (-1)^{w_j} r_j(x)\, d\mu(x) \\
&= \int_{\mathbb{F}_2^n} f(x) (-1)^{\langle v, w\rangle} w_A(x)\, d\mu(x) \\
&= -\hat{f}_A.
\end{aligned}$$

Hence $\hat{f}_A = 0$. □

Theorem 4.18 (Poincaré inequality for L_1-valued functions on $\mathbb{F}_2^n/C^\perp$)**.** For every $h\colon \mathbb{F}_2^n/C^\perp \to L_1$ we have

$$\int_{(\mathbb{F}_2^n/C^\perp)^2} \|h(u) - h(v)\|_{L_1}\, d\tilde{\mu}(u)\, d\tilde{\mu}(v) \leq \frac{1}{d(C)} \sum_{j=1}^n \int_{\mathbb{F}_2^n/C^\perp} \|\partial_j h(u)\|_{L_1}\, d\tilde{\mu}(u) \qquad (*)$$

where

$$\partial_j h(u) = \frac{h(u + q(e_j)) - h(u)}{2},$$

and $u \in \mathbb{F}_2^n/C^\perp$.

Proof. Let $f = h \circ q$. Then $(*)$ is equivalent to

$$\int_{\mathbb{F}_2^n \times \mathbb{F}_2^n} \|f(x) - f(y)\|_{L_1}\, d\mu(x)\, d\mu(y) \leq \frac{1}{d(C)} \sum_{j=1}^n \int_{\mathbb{F}_2^n} \|\partial_j f(x)\|_{L_1}\, d\mu(x).$$

From (proof of) Proposition 1.7, there exists a map $T\colon L_1 \to L_2$ such that

$\|Ta - Tb\|_{L_2} = \|a-b\|_{L_1}^{1/2}$. Now

$$\begin{aligned}
&\int\limits_{\mathbb{F}_2^n \times \mathbb{F}_2^n} \|f(x) - f(y)\|_{L_1} \, d\mu(x)\, d\mu(y) \\
&\qquad = \int_{\mathbb{F}_2^n \times \mathbb{F}_2^n} \|Tf(x) - Tf(y)\|_{L_2}^2 \, d\mu(x)\, d\mu(y) \\
&\qquad \le \frac{2}{\min\{|A| : A \neq \emptyset, \hat{f}_A \neq 0\}} \sum_{j=1}^n \int_{\mathbb{F}_2^n} \|\partial_j Tf(x)\|_{L_2}^2 \, d\mu(x) \quad \text{(Theorem 13)} \\
&\qquad \le \frac{2}{d(C)} \sum_{j=1}^n \int_{\mathbb{F}_2^n} \|\partial_j Tf(x)\|_{L_2}^2 \, d\mu(x) \quad \text{(Lemma 17)} \\
&\qquad = \frac{1}{d(C)} \sum_{j=1}^n \int_{\mathbb{F}_2^n} \|\partial_j f(x)\|_{L_1}^2 \, d\mu(x)
\end{aligned}$$

since $\|\partial_j Tf(x)\|_{L_2}^2 = \frac{\|Tf(x+e_j) - Tf(x)\|_{L_2}^2}{4} = \frac{\|f(x+e_j)-f(x)\|_{L_1}}{4} = \frac{1}{2}\|\partial_j f(x)\|_{L_1}$. □

Lemma 4.19. $\exists \beta > 0, \forall n \in \mathbb{N}$, if $\dim C \ge \frac{n}{4}$ then $\forall x \in \mathbb{F}_2^n$,

$$\mu(\{y : \rho(qx, qy) \ge \beta n\}) \ge \frac{1}{2}.$$

Proof. Let n, δ be as in Lemma 15. WLOG $N \ge 8$. WLOG $x = 0$. For $1 \le n \le N$, $\mu(\{y : \rho(qy, 0) \ge \frac{n}{N}\}) = \mu(\mathbb{F}_2^n \setminus C^\perp) = \frac{2^n - |C^\perp|}{2^n}$. From Lemma 14, $\dim C^\perp = n - \dim C \le n-1$, so $\frac{2^n - |C^\perp|}{2^n} \ge \frac{2^n - 2^{n-1}}{2^n} = \frac{1}{2}$. Now let $n > N$. For $v \in C^\perp$, consider

$$|\{y : d(v,y) < \delta n\}| \le \sum_{0 \le \ell < \delta n} \binom{n}{\ell} \le (m+1)\binom{n}{m},$$

where $m = \lfloor \delta n \rfloor$. So

$$\begin{aligned}
|\{y : \exists v \in C^\perp, d(y,v) < \delta n\}| &= |\{y : \rho(qy, 0) < \delta n\}| \\
&\le 2^{\dim C^\perp} (m+1) \binom{n}{m} \\
&\le 2^{3n/4} 2^{n/8} \le \frac{1}{2} 2^n.
\end{aligned}$$

(Here we use $n > N \ge 8$). So $\mu(\{y : \rho(qy, 0) \ge \delta n\}) \ge \frac{1}{2}$. So $\beta = \min(\delta, \frac{1}{N})$ works. □

Theorem 4.20. $\exists \eta > 0, \exists$ sequence (X_n) of metric spaces such that $|X_n| \to \infty$ and $c_1(X_n) \ge \eta \log |X_n|$.

Remark. Recall $c_2(X) \ge c_1(X)$ for any finite metric space. So Theorem 20 says that the upper bound in Bourgain's Embedding Theorem is best possible up to constant.

Proof. By Lemma 16, for every n there exists a linear code C in $\mathbb{F}_2^n$ with $\dim C \geq \frac{n}{4}$ and $d(C) \geq \alpha n$. Let $X_n = \mathbb{F}_2^n / C^\perp$ with the quotient metric ρ. By Lemma 14, $|X_n| = 2^{n - \dim C^\perp} = 2^{\dim C} \geq 2^{n/4} \to \infty$. By Proposition 2, a lower bound on $c_1(X_n)$ is given by the Poincaré ratio corresponding to the inequality in Theorem 18. Thus

$$\begin{aligned} c_1(X_n) &\geq \frac{\int_{X_n \times X_n} \rho(u,v)\, d\tilde{\mu}(u)\, d\tilde{\mu}(v)}{\frac{1}{d(C)} \sum_{j=1}^n \int_{X_n} \frac{\rho(u + q(e_j), u)}{2}\, d\tilde{\mu}(u)} \\ &= \frac{\int_{\mathbb{F}_2^n \times \mathbb{F}_2^n} \rho(q(x), q(y))\, d\mu(x)\, d\mu(y)}{\frac{1}{2d(C)} \sum_{j=1}^n \int_{\mathbb{F}_2^n} \rho(q(x + e_j), q(x))\, d\mu(x)}. \end{aligned}$$

It's clear that the denominator $\leq \frac{n}{2d(C)} \leq \frac{n}{2\alpha n} = \frac{1}{2\alpha}$. By Lemma 19, for each $x \in \mathbb{F}_2^n$, $\int_{\mathbb{F}_2^n} \rho(q(x), q(y))\, d\mu(y) \geq \frac{\beta n}{2}$. Hence the numerator is at least $\frac{\beta n}{2}$. Thus $c_1(X_n) \geq \frac{\beta n}{2} / \frac{1}{2\alpha} = \alpha\beta n \geq \alpha\beta \log_2 |X_n|$. □

5 Dimension Reduction

Theorem 5.1 (Johnson-Lindenstrauss Lemma)**.** There exists a constant $C > 0$ such that $\forall k, n \in \mathbb{N}, \forall \epsilon \in (0, 1)$, if $k \geq C\epsilon^{-2} \log n$ then any n-element subset of ℓ_2 embeds into ℓ_2^k with distortion at most $\frac{1+\epsilon}{1-\epsilon}$.

Remark. In the 90's there was a sudden explosion of citation for this result, because the computer scientists realised there are many applications in compress sensing etc. For applications, see Matousek's lecture notes.

Idea. We will take a random linear map $T\colon \ell_2^n \to \ell_2^k$ and show that for each $x \in \ell_2^n$, we have $(1-\epsilon)\|x\|_2 \leq \|Tx\|_2 \leq (1+\epsilon)\|x\|_2$ with high probability. It follows that, given $x_1, \ldots, x_n \in \ell_2^n$, we have

$$(1-\epsilon)\|x_i - x_j\|_2 \leq \|Tx_i - Tx_j\|_2 \leq (1+\epsilon)\|x_i - x_j\|_2$$

with positive probability. In particular, there exists a suitable map of $\{x_1, \ldots, x_n\}$ to ℓ_2^k.

Lemma 5.2. Let $k, n \in \mathbb{N}, \epsilon(0, 1)$. Define $T\colon \ell_2^n \to \ell_2^k$ by the $k \times n$ matrix $(\frac{1}{\sqrt{k}} Z_{ij})_{ij}$ where the Z_{ij} $(1 \leq i \leq k, 1 \leq j \leq n)$ are iid random variables with $Z_{ij} \sim N(0, 1)$. Then there exists a constant $c > 0$ (independent of k, ϵ) such that for each $x \in \ell_2^n$, we have

$$\mathbb{P}\Big((1-\epsilon)\|x\|_2 \leq \|Tx\|_2 \leq (1+\epsilon)\|x\|_2\Big) \geq 1 - 2e^{-ck\epsilon^2}.$$

Proof of Theorem 1. We choose $C > 0$ sufficiently large so that if $k, n \in \mathbb{N}, \epsilon \in (0, 1)$ satisfy $k \geq C\epsilon^{-2} \log n$, then $1 - 2e^{-ck\epsilon^2} \geq 1 - \frac{1}{n^2}$. Clearly, C depends only on c. Now let $T\colon \ell_2^n \to \ell_2^k$ be as in Lemma 2. Then for each $x \in \ell_2^n$,

$$\mathbb{P}\Big((1-\epsilon)\|x\|_2 \leq \|Tx\|_2 \leq (1+\epsilon)\|x\|_2\Big) \geq 1 - \frac{1}{n^2}.$$

So given $x_1, \ldots, x_n \in \ell_2$, WLOG $x_1, \ldots, x_n \in \ell_2^n$ and

$$\mathbb{P}\Big(\forall i, j \quad (1-\epsilon)\|x_i - x_j\|_2 \leq \|Tx_i - Tx_j\|_2 \leq (1+\epsilon)\|x_i - x_j\|_2\Big) \geq 1 - \binom{n}{2}\frac{1}{n^2} > 0.$$

So there exists a linear map T that has $\frac{1+\epsilon}{1-\epsilon}$-distortion on $\{x_1, \ldots, x_n\}$. $\square$

Recall that if $Z \sim N(0, 1)$ then Z has probability density function (pdf) $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. If $Z_1, \ldots, Z_n$ are iid $\sim N(0, 1)$ and $x \in \ell_2^n$ with $\|x\| = \sqrt{\sum_{i=1}^n x_i^2} = 1$, then $\sum_{i=1}^n x_i Z_i \sim N(0, 1)$.

Lemma 5.3 (Tail Estimates)**.** Let X be a random variable with $\mathbb{E}X = 0$. Assume that for some $C > 0, u_0 > 0$ we have $\mathbb{E}e^{uX} \leq e^{Cu^2}$ for $0 \leq u \leq u_0$. Then $\mathbb{P}(X > t) \leq e^{-t^2/4C}$ for $0 \leq t \leq 2Cu_0$.

Proof. For any $u \geq 0$,

$$\begin{aligned} \mathbb{P}(X > t) = \mathbb{P}(e^{uX} > e^{ut}) &\leq e^{-ut}\mathbb{E}e^{uX} && \text{(Markov's inequality)} \\ &\leq e^{-ut + Cu^2} && \text{(provided } 0 \leq u \leq u_0\text{)} \end{aligned}$$

If $0 \leq t \leq 2Cu_0$, then we can take $u = t/2C$ to obtain

$$\mathbb{P}(X > t) \leq e^{-\frac{t^2}{2C} + \frac{t^2}{4C}} = e^{-\frac{t^2}{4C}}.$$

□

Lemma 5.4. Assume $Z \sim N(0,1)$. Then there exists absolute constant $C, u_0 > 0$ such that $\mathbb{E}e^{u(Z^2-1)} \leq e^{Cu^2}$ and $\mathbb{E}e^{u(1-Z^2)} \leq e^{Cu^2}$ for $0 \leq u \leq u_0$.

Proof. This is straightforward computation.

$$\begin{aligned}
\mathbb{E}e^{u(1-Z^2)} &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{u(1-x^2)} e^{-x^2/2}\, dx \\
&= e^u \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}(2u+1)x^2}\, dx \\
&= \frac{e^u}{\sqrt{2u+1}} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-y^2}\, dy && (\text{put } y = \sqrt{2u+1}x) \\
&= \frac{e^u}{\sqrt{2u+1}} \\
&= e^{u - \frac{1}{2}\log(2u+1)} \\
&= e^{u^2 + O(u^3)}
\end{aligned}$$

using $\log(1+x) = -\sum_{n=1}^\infty \frac{(-x)^n}{n}$. A similar computation shows $\mathbb{E}e^{u(Z^2-1)} \leq e^{u^2+O(u^3)}$. □

Proof of Lemma 2. Fix $x \in \ell_2^n$. WLOG assume $\|x\|_2 = 1$. Then

$$(Tx)_i = \frac{1}{\sqrt{k}} \sum_{j=1}^n x_j Z_{ij}, \quad 1 \leq i \leq k.$$

Let $Z_i = \sum_{j=1}^n x_j Z_{ij}$. Then $Z_1, \ldots, Z_n$ are iid with $Z_i \sim N(0,1)$. Then

$$\mathbb{E}\|Tx\|^2 = \sum \mathbb{E}|(Tx)_i|^2 = \frac{1}{k}\sum_{i=1}^k \mathbb{E}(Z_i^2) = 1.$$

Let $W = \frac{1}{\sqrt{k}} \sum_{i=1}^k (Z_i^2 - 1)$. Then $\mathbb{E}W = 0$ and $\operatorname{var} W = 1$. Fix C, u_0 as given by Lemma 4 and WLOG $2cu_0 \geq 1$. Then

$$\begin{aligned}
\mathbb{E}e^{uW} &= \prod_{i=1}^k e^{\frac{u}{\sqrt{k}}(Z_i^2-1)} && (\text{by independence}) \\
&\leq \prod_{i=1}^k e^{Cu^2/k} && (\text{Lemma 4}) \\
&= e^{Cu^2} && (\text{if } 0 \leq u \leq \sqrt{k}u_0).
\end{aligned}$$

Similarly $\mathbb{E}(e^{-uW}) = \prod_{i=1}^{k} e^{u/sqrtk(1-Z_i^2)} \leq e^{Cu^2}$. So $\mathbb{P}(W > t) \leq e^{-t^2/4C}$, $\mathbb{P}(W < -t) \leq e^{-t^2/4C}$ for $0 \leq t \leq 2Cu_0\sqrt{k}$ (note $2Cu_0 \geq 1$). So

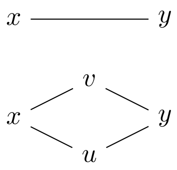
$$\begin{aligned}\mathbb{P}\Big((1-\epsilon)\|x\|_2 \leq \|Tx\|_2 \leq &(1+\epsilon)\|x\|_2\Big) \\ &= \mathbb{P}\Big((1-\epsilon)^2 \leq \|Tx\|_2^2 \leq (1+\epsilon)^2\Big) \\ &\geq \mathbb{P}\Big(1-\epsilon \leq \frac{1}{k}\sum_{i=1}^{k} Z_i^2 \leq 1+\epsilon\Big) \\ &= \mathbb{P}\Big(1-\epsilon \leq \frac{1}{\sqrt{k}} W + 1 \leq 1+\epsilon\Big) \\ &= \mathbb{P}(-\epsilon\sqrt{k} \leq W \leq \epsilon\sqrt{k}) \\ &\geq 1 - 2e^{-\epsilon^2 k/4C}.\end{aligned}$$

□

Aim. Our aim is to prove that dimension reduction as in JL Lemma does not work in ℓ_1.

Theorem 5.5. For all $n \in \mathbb{N}$ there exists a subset X of ℓ_1 of size $|X| = N \geq n$ such that if X embeds into ℓ_1^k with distortion $\leq D$, then $k \geq n^{\frac{1}{32D^2}}$.

We introduce the *diamond graphs* D_n, $n = 0, 1, 2, \dots$: D_0 consists of 2 vertices joined by an edge. D_{n+1} is obtained from D_n by replacing every edge xy in D_n with new vertices u, v and edges xv, vy, xu, uy. Note $D_0 = K_2, D_1 = C_4$.



Let $E_n = E(D_n), V_n = V(D_n)$. Then $|E_n| = 4^n, |V_n| = 2 + 2(1 + 4 + \cdots + 4^{n-1}) = \frac{2}{3}(4^n + 2)$. Observe that $|V_n| \leq 4^n$ for all $n \geq 1$.

Let $d_n = d_{D_n}$. For every $n \geq m \geq 0, \forall x, y \in D_m$, $d_n(x, y) = 2^{n-m} d_m(x, y)$.

We define sets A_n for $n \geq 1$ of "non-edges" as follows: For $n \geq 1$, D_n consists of copies of $D_1 = C_4$ of the form $xyuv$ where $xy \in E_{n-1}$ and $u, v \in V_n \setminus V_{n-1}$. Let A_n consist of all pairs $\{u, v\}$.

Let's label the vertices as follows. $D_0 = \ell r$ for left and right, $D_1 = \ell brt$ where b for bottom and t for top. Write $D_n(\ell r)$ for D_n. $D_{n+1}(\ell r)$ consists of 4 copies of D_n: $D_n(t\ell), D_n(tr), D_n(b\ell), D_n(br)$. If e, f are two of the edges $t\ell, tr, b\ell, br$, then $V(D_n(e)) \cap V(D_n(f)) = e \cap f$.

Remark. $d_n(\ell, r) = 2^n$ for all $n \geq 0$, and $d_n(t, b) = 2^n$ for all $n \geq 1$. For every $x \in D_n$, $d_n(\ell, x) + d_n(x, r) = 2^n$.

Lemma 5.6. For all $n \geq 0$, D_n embeds into $\ell_1^{2^n}$ with distortion ≤ 2.

Proof. Let $f_0 \colon D_0 \to H_k \subset \ell_1^k$ be such that $f_0(\ell), f_0(r)$ are neighbours in H_k. So f_0 is isometric (e.g. $k = 1 = 2^0$, $f_0(\ell) = (0), f_0(r) = (1)$). Assume $f_n \colon D_n \to$

$H_{k2^n} \subset \ell^{k2^n}$ has been defined. Then we define $f_{n+1}: D_{n+1} \to H_{k2^{n+1}} \subset \ell^{k2^{n+1}}$ as follows: let $x \in D_n$ we let $f_{n+1}(x) = (f_n(x), f_n(x))$. If $xy \in E_n$ and u, v are the corresponding new vertices in D_{n+1}, we let $f_{n+1}(u) = (f_n(x), f_n(y))$, $f_{n+1}(v) = (f_n(y), f_n(x))$.

Observe that for $x, y \in D_n$, $\|f_{n+1}(x) - f_{n+1}(y)\|_1 = 2\|f_n(x) - f_n(y)\|_1$. So $\forall n \geq m \geq 0, \forall x, y \in D_m$, $\|f_n(x) - f_n(y)\|_1 = 2^{n-m}\|f_m(x) - f_m(y)\|_1$.

First show that $\forall n \geq 0, \forall xy \in E_n$, $\|f_n(x) - f_n(y)\|_1 = d_n(x, y) = 1$. Proof by induction on n: $n = 0$ (and $n = 1$) is clear. Now assume $n \geq 1$. An edge in D_n is of the form xu, where $\exists xy \in E_{n-1}$ and u, v are the corresponding new vertices in D_n. Now $\|f_n(x) - f_n(y)\| = \|(f_{n-1}(x), f_{n-1}(x)) - (f_{n-1}(x), f_{n-1}(y))\|_1 = \|f_{n-1}(x) - f_{n-1}(y)\|_1 = 1$ by induction hypothesis. It follows that f_n is 1-Lipschitz for all $n \geq 0$. To see this, given $x, y \in D_n$, there exists a path $x = x_0, x_1, x_2, \ldots, x_m = y$ in D_n with $m = d_n(x, y)$. Then $\|f_n(x) - f_n(y)\|_1 \leq \sum_{i=1}^m \|f_n(x_i) - f_n(x_{i-1})\|_1 = m = d_n(x, y)$.

Claim. $\forall n \geq 0, \forall x, y \in D_n$, $\|f_n(x) - f_n(y)\|_1 \geq \frac{1}{2} d_n(x, y)$.

Note that $\forall n \geq m \geq 0$, if $xy \in E_m$, then $\|f_n(x) - f_n(y)\|_1 = 2^{n-m}\|f_m(x) - f_m(y)\|_1 = 2^{n-m} = 2^{n-m} d_m(x, y) = d_n(x, y)$. In fact, it is enough if $\|f_m(x) - f_m(y)\|_1 = d_m(x, y)$.

This claim is proved by induction on n. Note that f_0, f_1 are isometric. Assume $n \geq 2$ and the claim holds for $n - 1$. Fix $x, y \in D_n$. Recall that D_n consists of 4 copies of D_{n-1}. We have 3 cases.

Case 1: x, y in the same copy, WLOG $x, y \in D_{n-1}(t\ell)$. Define $g_0: D_0(t\ell) \to H_{2k}$, $g_0(u) = f_1(u)$. Then define $g_m: D_m \to H_{2^m k}$ inductively starting with g_0 in the same way as f_m is defined from f_0. Then by easy induction, $g_{n-1} = f_n|_{D_{n-1}(t\ell)}$. By induction hypothesis, $\|f_n(x) - f_n(y)\|_1 = \|g_{n-1}(x) - g_{n-1}(y)\|_1 \geq \frac{1}{2} d_{D_{n-1}(t\ell)}(x, y) \geq \frac{1}{2} d_{D_n}(x, y)$. [In fact, the last inequality is an equality, because the four copies of D_{n-1} only meet at ℓ, b, r or t.]

Case 2: x, y are in neighbouring copies, WLOG $x \in D_{n-1}(t\ell), y \in D_{n-1}(tr)$. Now $\|f_n(x) - f_n(y)\| \geq \|f_n(\ell) - f_n(r)\|_1 - \|f_n(\ell) - f_n(x)\|_1 - \|f_n(y) - f_n(r)\|_1 = 2^{n-1}\|f_1(\ell) - f_1(r)\| - d_n(x, \ell) - d_n(y, r) = 2^n - d_n(x, \ell) - d_n(y, r) = \left(2^{n-1} - d_{D_{n-1}(t\ell)}(x, \ell)\right) + \left(2^{n-1} - d_{D_{n-1}(tr)}(y, r)\right) = d_n(x, t) + d_n(t, y) = d_n(x, y)$.

Case 3: x, y are in opposite copies, WLOG $x \in D_{n-1}(t\ell), y \in D_{n-1}(br)$. Then

$$d_n(x, y) = \left(d_n(x, \ell) + 2^{n-1} + d_n(b, y)\right) \wedge \left(d_n(x, t) + 2^{n-1} + d_n(r, y)\right) \leq 2^n,$$

since $d_n(x, \ell) + d_n(b, y) + d_n(x, t) + d_n(r, y) = 2^n$. Assume WLOG $d_n(x, t) + d_n(y, b) \leq d_n(x, \ell) + d_n(y, r)$. So $d_n(x, t) + d_n(y, b) \leq 2^{n-1}$. Then by the triangle inequality and the fact that f_n is 1-Lipschitz,

$$\begin{aligned} \|f_n(x) - f_n(y)\|_1 &\geq \|f_n(t) - f_n(b)\|_1 - \|f_n(x) - f_n(t)\|_1 - \|f_n(y) - f_n(b)\|_1 \\ &\geq 2^n - d_n(x, t) - d_n(y, b) \geq 2^{n-1} \geq \frac{1}{2} d_n(x, y). \end{aligned}$$

□

Recall that for all $x_1, x_2, x_3, x_4 \in \ell_2$ we have

$$\begin{aligned}\|x_1 - x_3\|_2^2 + \|x_2 - x_4\|_2^2 \leq \|x_1 - x_2\|_2^2 + \|x_2 - x_3\|_2^2 \\ + \|x_3 - x_4\|_2^2 + \|x_4 - x_1\|_2^2,\end{aligned}$$

also called the *Short Diagonal Lemma*.

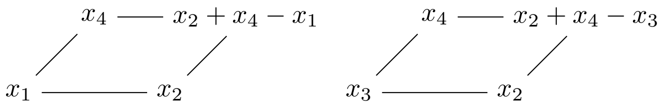
Lemma 5.7 (Short diagonal Lemma in L_p)**.** Let $1 < p < 2$. Then $\forall x_1, x_2, x_3, x_4 \in L_p$, we have

$$\begin{aligned}\|x_1 - x_3\|_p^2 + (p-1)\|x_2 - x_4\|_p^2 \leq \|x_1 - x_2\|_p^2 + \|x_2 - x_3\|_p^2 \\ + \|x_3 - x_4\|_p^2 + \|x_4 - x_1\|_p^2,\end{aligned}$$

Proof. WLOG $x_1, x_2, x_3, x_4 \in \ell_p^k$ for some k ($k = 6$ will do by Theorem 2.7). Lemma 7 can be deduced from the following:

$$\|x\|_p^2 + (p-1)\|y\|_p^2 \leq \frac{\|x+y\|_p^2 + \|x-y\|_p^2}{2} \qquad \forall x, y \in \ell_p^k. \qquad (*)$$

To see this, consider two parallelograms:



For the first parallelogram, set $x = x_2 + x_4 - 2x_1, y = x_4 - x_2$. For the second parallelogram, set $x = x_2 + x_4 - 2x_3, y = x_4 - x_2$. Apply $(*)$ for both parallelograms:

$$\begin{aligned}\|x_2 + x_4 - 2x_1\|_p^2 + (p-1)\|x_2 - x_4\|_p^2 \leq 2\|x_4 - x_1\|_p^2 + 2\|x_2 - x_1\|_p^2, \\ \|x_2 + x_4 - 2x_3\|_p^2 + (p-1)\|x_2 - x_4\|_p^2 \leq 2\|x_4 - x_3\|_p^2 + 2\|x_2 - x_3\|_p^2.\end{aligned}$$

We take average of these 2 inequalities and use convexity of $z \mapsto \|z\|_p^2$ to get

$$\begin{aligned}
&\|x_1 - x_3\|_p^2 + (p-1)\|x_2 - x_4\|_p^2 \\
&\quad = \left\|\frac{x_2 + x_4 - 2x_3}{2} + \frac{2x_1 - x_2 - x_4}{2}\right\|_p^2 + (p-1)\|x_2 - x_4\|_p^2 \\
&\quad \leq \frac{\|x_2 + x_4 - 2x_3\|_p^2 + \|x_2 + x_4 - 2x_1\|_p^2}{2} + (p-1)\|x_2 - x_4\|_p^2 \\
&\quad \leq \|x_1 - x_2\|_p^2 + \|x_2 - x_3\|_p^2 \\
&\qquad\qquad + \|x_3 - x_4\|_p^2 + \|x_4 - x_1\|_p^2,
\end{aligned}$$

as required.

To prove $(*)$, use the fact that for $a, b \geq 0$, $\left(\frac{a^q + b^q}{2}\right)^{1/q}$ is increasing in $q \in [1, \infty)$. So $(*)$ follows from

$$\|x\|_p^2 + (p-1)\|y\|_p^2 \leq \left(\frac{\|x+y\|_p^p + \|x-y\|_p^p}{2}\right)^{2/p}.$$

Define

$$\begin{aligned} L(t) &= \|x\|_p^2 + (p-1)\|y\|_p^2 t^2, \\ R(t) &= \left(\frac{\|x+ty\|_p^p + \|x-ty\|_p^p}{2}\right)^{2/p} = H(t)^{2/p}, \\ H(t) &= \frac{1}{2}\sum_{i=1}^k \left(|x_i + ty_i|^p + |x_i - ty_i|^p\right), \qquad t \in \mathbb{R}. \end{aligned}$$

We need that $L(1) \le R(1)$. We have $L(0) = R(0) = \|x\|_p^2$. From now we assume $x \neq 0, y \neq 0$. Next we differentiate.

$$\begin{aligned} L'(t) &= 2(p-1)\|y\|_p^2 t \\ R'(t) &= \frac{2}{p} H(t)^{\frac{2}{p}-1} H'(t) \\ H'(t) &= \frac{p}{2}\sum_{i=1}^k \big(|x_i + ty_i|^{p-1}\operatorname{sgn}(x_i + ty_i) y_i - |x_i - ty_i|^{p-1}\operatorname{sgn}(x_i - ty_i) y_i\big). \end{aligned}$$

Note that $L'(0) = R'(0) = 0$. Differentiate again:

$$L''(t) = 2(p-1)\|y\|_p^2.$$

Let $I = [k] \setminus \{i \in [k] : x_i = y_i = 0\}$, where $[k] = \{1, \ldots, k\}$. Note $I \neq \emptyset$ as $x, y \neq 0$. For $i \in I$, there is ≤ 1 value of t such that $x_i + ty_i = 0$. So there exists dissection $0 = t_0 < t_1 < \cdots < t_m = 1$ of $[0,1]$ such that $x_i + ty_i \neq 0, \forall i \in I, \forall t \in \bigcup_{j=1}^m (t_{j-1}, t_j)$. For such t, we have

$$\begin{aligned} R''(t) &= \frac{2}{p}\left(\frac{2}{p} - 1\right) H(t)^{\frac{2}{p}-2}(H'(t))^2 + \frac{2}{p} H(t)^{\frac{2}{p}-1} H''(t) \\ &\ge \frac{2}{p} H(t)^{\frac{2}{p}-1} H''(t) \\ &= \frac{2}{p} H(t)^{\frac{2}{p}-1} \frac{p}{2}(p-1)\sum_{i\in I}\big(|x_i + ty_i|^{p-2} y_i^2 + |x_i - ty_i|^{p-2} y_i^2\big). \end{aligned}$$

We now use *reverse Hölder's inequality*: suppose $0 < r < 1$ and $\frac{1}{r} + \frac{1}{s} = 1$, so $s = \frac{r}{r-1} < 0$. Given $a_i, b_i \in \mathbb{R}$, $b_i \neq 0$, we have

$$\begin{aligned} \left(\sum_{i\in I} |a_i|^r\right)^{1/r} &= \left(\sum_{i\in I} |a_i b_i|^r |b_i|^{-r}\right)^{1/r} \qquad \left(\text{take } p = \frac{1}{r}, q = \frac{1}{1-r}\right) \\ &\le \left(\sum_{i\in I} |a_i b_i|\right)\left(\sum_{i\in I} |b_i|^s\right)^{-1/s}, \end{aligned}$$

so

$$\left(\sum_{i\in I} |a_i|^r\right)^{1/r}\left(\sum_{i\in I} |b_i|^s\right)^{1/s} \le \sum_{i\in I} |a_i b_i|.$$

Apply this with $b_i = |x_i \pm ty_i|^{p-2}$, $a_i = y_i^2$, $r = \frac{p}{2}$, $s = \frac{p}{p-2}$, we have

$$\begin{aligned}
R''(t) &\geq H(t)^{\frac{2}{p}-1}(p-1)\left(\sum_{i\in I}|y_i|^p\right)^{2/p}\left(\left(\sum_{i\in I}|x_i+ty_i|^p\right)^{\frac{p-2}{p}}+\left(\sum_{i\in I}|x_i-ty_i|^p\right)^{p}\right. \\
&\geq H(t)^{\frac{2}{p}-1}(p-1)\|y\|_p^2\cdot 2\left(\frac{\|x+ty\|_p^{p-2}+\|x-ty\|_p^{p-2}}{2}\right) \\
&\geq H(t)^{\frac{2}{p}-1}(p-1)2\|y\|_p^2\left(\frac{\|x+ty\|_p^{p}+\|x-ty\|_p^{p}}{2}\right)^{\frac{p-2}{p}} \qquad (r\mapsto r^{\frac{p-2}{p}}\text{ convex}) \\
&= 2(p-1)\|y\|_p^2 = L''(t).
\end{aligned}$$

So for each $1 \leq j \leq m$, $(R-L)'' \geq 0$ on (t_{j-1}, t_j), so $(R-L)'$ is increasing on $[t_{j-1}, t_j]$. So $(R-L)'$ is increasing on $[0,1]$ and hence $(R-L)' \geq 0$ on $[0,1]$. So $R-L$ is increasing on $[0,1]$ and hence $R(1)-L(1) \geq 0$. □

Corollary 5.8. For $1 < p < 2, n \in \mathbb{N}$, $c_p(D_2) \geq \sqrt{1+(p-1)n}$.

Proof. D_n consists of copies of $D_1 = xuyv$, where $xy \in E_{n-1}, uv \in V_n \setminus V_{n-1}$. Apply Lemma 7 for a function $f\colon D_n \to L_p$:

$$\begin{aligned}
\|f(x)-f(u)\|_p^2 + \|f(u)-f(y)\|_p^2 + \|f(y)-f(v)\|_p^2 + \|f(v)-f(x)\|_p^2& \\
\geq \|f(x)-f(y)\|_p^2 + (p-1)\|f(u)-f(v)\|_p^2&.
\end{aligned}$$

Sum over all copies of D_1 in D_n:

$$\begin{aligned}
\sum_{xy\in E_n}\|f(x)-f(y)\|_p^2 &\geq \sum_{xy\in E_{n-1}}\|f(x)-f(y)\|_p^2 + (p-1)\sum_{xy\in A_n}\|f(x)-f(y)\|_p^2 \\
&\geq \cdots \\
&\geq \|f(\ell)-f(r)\|_p^2 + (p-1)\sum_{xy\in A_1\cup\cdots\cup A_n}\|f(x)-f(y)\|_p^2.
\end{aligned}$$

We bound $c_p(D_n)$ from below using the corresponding Poincaré ratio. For $xy \in A_k$, $d_n(x,y) = 2^{n-k}d_k(x,y) = 2^{n-k+1}$ and $|A_k| = 4^{k-1}$. So $d_n(\ell, r)^2 + (p-1)\sum_{k=1}^n 4^{k-1}4^{n-k+1} = 4^n(1+(p-1)n)$. So $c_p(D_n) \geq \left(\frac{4^n(1+(p-1)n)}{4^n}\right)^{1/2} = \sqrt{1+(p-1)n}$. □

Lemma 5.9. Given $k \geq 2$, the identity $i_p\colon \ell_1^k \to \ell_p^k$ where $p = 1 + \frac{1}{\log_2 k}$ has distortion at most 2.

Proof. For $x = (x_i)_{i=1}^k \in \mathbb{R}^k$, by Hölder, $\|x\|_p \leq \|x\|_1 = \sum_{i=1}^k |x_i| \leq k^{1-1/p}\|x\|_p$. Now $k^{1-1/p} = k^{\frac{1/\log_2 k}{1+1/\log_2 k}} = k^{\frac{1}{\log_2 k+1}} = 2^{\frac{\log_2 k}{\log_2 k+1}} \leq 2$. □

Proof of Theorem 5. Let $n \in \mathbb{N}$. By Theorem 6, there exists an embedding $f\colon D_n \to \ell_1$ of distortion at most 2. Set $X = f(D_n)$. So $|X| = |D_n| \leq 4^n$. Assume $g\colon X \to \ell_1^k$ has distortion at most D. Then $i_p g f\colon X \to \ell_p^k$, $p = 1 + \frac{1}{\log_2 k}$ has distortion $\leq 4D$ (Lemma 9). By Corollary 8, $4D \geq \sqrt{1+(p-1)n}$, and $16D^2 \geq \frac{n}{\log_2 k} \geq \frac{\log_2|X|}{2\log_2 k}$. So $\log_2 k \geq \frac{\log_2|X|}{32D^2}$ and hence $k \geq |X|^{\frac{1}{32D^2}}$. □

6 Ribe Programme

Definition. Given Banach spaces X, Y, we say X is *finitely representable* in Y if $\forall E \subset X, \dim E < \infty$, $\forall \lambda > 1$, $\exists F \subset Y$ such that $d(E,F) < \lambda$, i.e. there exists a linear bijection $T\colon E \to F$ such that $\|T\|\|T^{-1}\| < \lambda$.

Example.

(i) Every X is finitely representable in c_0.

(ii) ℓ_2 is finitely representable in every ∞-dimensional X [Dvoretzky].

Definition. X is *crudely finitely representable* in Y if $\exists \lambda > 1$, $\forall E \subset X$, $\dim E < \infty$, $\exists F \subset Y$, s.t. $d(E,F) < \lambda$.

Definition. A *local property* (or *local isomorphic property*) of a Banach space is one that depends only on its finite-dimensional subspaces.

Definition. For $1 \leq p \leq 2$, we say X has *type* p if $\exists C > 0$, $\forall n \in \mathbb{N}$, $\forall x_1, \ldots, x_n \in X$, $\mathbb{E}\|\sum_{i=1}^n \epsilon_i x_i\| \leq C\left(\sum_{i=1}^n \|x_i\|^p\right)^{1/p}$. Here, $\epsilon_1, \ldots, \epsilon_n$ are $\{\pm 1\}$-valued independent $\mathtt{Rademacher}(\frac{1}{2})$ random variables.

For $2 \leq q \leq \infty$, we say X has *cotype* q if $\exists C > 0$, $\forall n \in \mathbb{N}$, $\forall x_1, \ldots, x_n \in X$, $\mathbb{E}\|\sum_{i=1}^n \epsilon_i x_i\| \geq \frac{1}{C}\left(\sum_{i=1}^n \|x_i\|^q\right)^{1/q}$. For $q = \infty$, RHS $= \frac{1}{C}\max_{1\leq i \leq n}\|x_i\|$.

Example. Every X has type 1, cotype ∞; ℓ_2 has type 2 and cotype 2 with $C = 1$.

If X is crudely finitely representable in Y and Y has some local property, then so does X.

Theorem 6.1 (Ribe's Theorem)**.** If Banach spaces X, Y are uniformly homeomorphic then X is crudely finitely representable in Y and vice versa.

Proof. Omitted. □

Remark. Local properties depend only on the metric structure of the Banach space, not the linear structure.

Aim. Aim for the Ribe programme:

(i) Find metric characterisations of local properties of Banach spaces.

(ii) Find metric analogues of local properties of Banach spaces.

Our aim is to find a metric characterisation of *super-reflexivity*.

Definition. Recall that given a Banach space X, there is an isometric isomorphism $X \longrightarrow X^{**}$ $x \longmapsto \hat{x}$, where $\hat{x}(f) = f(x)$. Easy to check $\hat{x} \in X^{**}$ and $\|\hat{x}\| \leq \|x\|$. By Hahn-Banach, we have $\|\hat{x}\| = \|x\|$. It's then clear that $x \longmapsto \hat{x}$ is linear. So the image of X in X^{**} is a closed subspace of X^{**}, which we will always identify with X. Say X is reflexive if $X = X^{**}$.

Warning. There exists Banach space J such that J is isometrically isomorphic to J^{**} but J^{**}/J has dimension 1.

Definition. We say X is *super-reflexive* if every Y finitely representable in X is reflexive. So super-reflexive $\implies$ reflexive.

Example. Let $X = \left(\bigoplus_{n\in\mathbb{N}} \ell_1^n\right)_{\ell_2} = \{(x_n) : x_n \in \ell_1^n \,\forall n, \sum \|x_n\|^2 < \infty\}$. X is reflexive, but ℓ_1 is finitely representable in X (see example sheet), so X is not super-reflexive.

We recall the following for a Banach space X:

(i) The *weak topology* on X is defined as follows: $U \subset X$ is *w-open* if $\forall x \in U$, $\exists n \in \mathbb{N}$, $\exists f_1, \dots, f_n \in X^*$, $\exists \epsilon > 0$ such that $\{y : |f_i(y - x)| < \epsilon, \forall i\} \subset U$. Note $|f_i(y-x)| < \epsilon$ can be written as $f_i(x) - \epsilon < f_i(y) < f_i(x) + \epsilon$. So this is a cylindrical set with finite codimension. This is the weakest topology on X for which every $f \in X^*$ is continuous.

(ii) A convex subset C of X is $\|\cdot\|$-closed $\iff$ w-closed.

Proof. ($\Leftarrow$) is clear. ($\Rightarrow$) if $x \notin C$, then by Hahn-Banach separation ($\{x\}$ compact convex, C closed convex), there exists $f \in X^*$ such that $\sup_C f < f(x)$. So $\{y : f(y) > \sup_C f\}$ is a weak neighbourhood of x disjoint from C. □

(iii) The *w^*-topology* on X^* is defined as follows: $U \subset X^*$ is *w^*-open* $\iff$ $\forall f \in U$, $\exists n \in \mathbb{N}$, $x_1, \dots, x_n \in X$, $\epsilon > 0$ such that $\{g \in X^* : |(g - f)(x_i)| < \epsilon, \forall i\} \subset U$. This is the weakest topology on X^* for which every $x \in X \subset X^{**}$ is continuous. So w^*-topology $\subset$ w-topology on X^*.

(iv) **Banach-Alaoglu Theorem**: $B_{X^*} = \{f \in X^* : \|f\| \le 1\}$ is w^*-compact.

Proof. Define

$$(B_{X^*}, w^*) \xrightarrow{\varphi} \prod_{x\in X} \{\lambda \in \mathbb{R} : |\lambda| \le \|x\|\} \ ,$$

with $\varphi(f) = (f(x))_{x\in X}$ where the codomain is equipped with the product topology, which is compact by Tychonov. It's clear that φ is a homeomorphism of B_{X^*} onto $\varphi(B_{X^*})$. Then $\varphi(B_{X^*}) = \bigcap_{x,y\in X, a,b\in\mathbb{R}} \{(\lambda_x)_{x\in X} : \lambda_{ax+by} - a\lambda_x - b\lambda_y = 0\}$, which is closed, hence compact. □

(v) **Goldstine's Theorem**: $\overline{B_X}^{w^*} = B_{X^{**}}$ in X^{**}.

(vi) X is reflexive $\iff$ (B_X, w) is compact.

Proof. ($\Rightarrow$): We have $X = X^{**}$, so $(X, w) = (X^{**}, w^*)$ so $(B_X, w) = (B_{X^{**}}, w^*)$ which is compact by Banach-Alaoglu.

($\Leftarrow$): The restriction of the w^*-topology of X^{**} to X is the w-topology. So B_X is w^*-compact in X^{**}. So B_X is w^*-closed and hence $B_{X^{**}} = \overline{B_X}^{w^*} = B_X$ and hence $X^{**} = X$. □

Lemma 6.2 (Local reflexivity)**.** Let X be a Banach space, $E \subset X^*$ with $\dim E < \infty$ and let $\varphi \in X^{**}$ and let $M > \|\varphi\|$. Then $\exists x \in X$ such that $\|x\| < M$ and $\hat{x}|_E = \varphi|E$.

Remark. We can now prove Goldstine: $\overline{B_X}^{w^*} = B_{X^{**}}$. Since $B_X \subset B_{X^{**}}$ and $B_{X^{**}}$ is w^*-closed, it follows that $\overline{B_X}^{w^*} \subset B_{X^{**}}$. Fix $\psi \in B_{X^{**}}$ and a w^*-neighbourhood U of ψ. Then $\exists n \in \mathbb{N}, f_1, \ldots, f_n \in X^*, \exists \epsilon > 0$ such that $\{\chi \in X^{**} : |(\chi - \psi)(f_i)| < \epsilon, \forall i\} \subset U$. Fix $\delta > 0$ to be determined. By Lemma 2, $\exists x \in X, \|x\| < 1 + \delta$, and $f_i(x) = \psi(f_i)$ for all i. If $\|x\| \leq 1$, then $x \in B_X \cap U$, so done. Assume $\exists x\| > 1$. Then

$$\left|\frac{\hat{x}}{\|x\|}(f_i) - \psi(f_i)\right| = \left|\frac{f_i(x)}{\|x\|} - f_i(x)\right| = \frac{|f_i(x)|}{\|x\|}\,|1 - \|x\|| \leq \delta\|f_i\|, \quad \forall i.$$

We can choose $\delta > 0$ such that $\delta\|f_i\| < \epsilon$ for all i, and then $\frac{x}{\|x\|} \in B_X \cap U$.

Proof of Lemma 2. Fix a basis $f_1, \ldots, f_n$ of E. Define $T\colon X \longrightarrow \mathbb{R}^n$ by $Tx = (f_i(x))_{i=1}^n$ and let $C = \{Tx : \|x\| < M\}$. We need $(\varphi(f_i))_{i=1}^n \in C$. Then we will be done. T is a bounded linear map and C is convex. We show that T is onto: if not, then there exists $a = (a_1, \ldots, a_n) \in \mathbb{R}^n \setminus \{0\}$ such that $\sum_{i=1}^n a_i f_i(x) = 0$ for all x, i.e., $\sum_{i=1}^n a_i f_i = 0$, but this is a contradiction. By the Open Mapping Theorem, C is an open set. Let's assume that $(\varphi(f_i))_{i=1}^n \notin C$. By Hahn-Banach separation, $\exists a = (a_1, \ldots, a_n) \neq 0$ such that $\sum_{i=1}^n a_i f_i(x) < \sum_{i=1}^n a_i \varphi(f_i)$ for all $x \in X, \|x\| < M$. Hence $\|\sum_{i=1}^n a_i f_i\| M \leq \varphi(\sum_{i=1}^n a_i f_i) \leq \|\varphi\| \|\sum_{i=1}^n a_i f_i\|$. Since $\sum_{i=1}^n a_i f_i \neq 0$, we get $M \leq \|\varphi\|$, a contradiction. ☐

Theorem 6.3. Let X be a Banach space. Then the following are equivalent:

(i) X is non-reflexive;

(ii) $\forall \theta \in (0,1)$, $\exists (x_i)_{i=1}^\infty$ in B_X, $(f_i)_{i=1}^\infty$ in B_{X^*}, such that

$$f_i(x_j) = \begin{cases} \theta & \text{if } i \leq j \\ 0 & \text{if } i > j; \end{cases}$$

(iii) $\exists \theta \in (0,1)$, the above holds;

(iv) $\forall \theta \in (0,1)$, $\exists (x_i)$ in B_X such that $\forall n \in \mathbb{N}$,

$$d(\operatorname{conv}\{x_1, \ldots, x_n\}, \operatorname{conv}\{x_{n+1}, x_{n+2}, \ldots\}) \geq \theta.$$

(v) $\exists \theta \in (0,1)$, such that the above holds.

Proof. (i) $\implies$ (ii): Since X is a proper closed subspace of X^{**}, $\exists T \in X^{***}$ such that $\|T\| = 1$, $T|_X = 0$ (by Hahn-Banach). Fix $\theta \in (0,1)$ and choose $\varphi \in X^{**}$, $\|\varphi\| < 1$, $T(\varphi) > \theta$. Let $\lambda = T(\varphi)$. Then $\theta < \lambda = T(\varphi) \leq \|T\|\|\varphi\| = \|\varphi\| < 1$, i.e. $\theta < \lambda < 1$.

Since $\|\varphi\| > \theta$, there exists $f_1 \in B_{X^*}$ such that $\varphi(f_1) = \theta$. Then $\theta = \varphi(f_1) \leq \|\varphi\|\|f_1\| < \|f_1\|$, and hence $\exists x_1 \in B_X$ such that $f_1(x_1) = \theta$.

Assume now that for some $n \geq 1$ we have found sequences $(x_i)_{i=1}^n$ in B_X and $(f_i)_{i=1}^n$ in B_{X^*} such that

$$f_i(x_j) = \begin{cases} \theta & \text{if } 1 \leq i \leq j \leq n \\ 0 & \text{if } 1 \leq j < i \leq n, \end{cases}$$

and $\varphi(f_i) = \theta$ for $1 \le i \le n$. Since $T(x_i) = 0$ for $1 \le i \le n$ and $T(\varphi) = \lambda$ and $\|T\| = 1 < \frac{\lambda}{\theta}$, by Lemma 2, $\exists g \in X^*$ such that $\|g\| < \frac{\lambda}{\theta}$ and $g(x_i) = 0$ for $1 \le i \le n$ and $\varphi(g) = \lambda$. Then $f_{n+1} = \frac{\theta}{\lambda} g \in B_{X^*}$ and $f_{n+1}(x_i) = 0$ for $1 \le i \le n$ and $\varphi(f_{n+1}) = \theta$. Since $\varphi(f_i) = \theta$ for $1 \le i \le n+1$ and $\|\varphi\| < 1$, so by Lemma 2, $\exists x_{n+1} \in B_X$ such that $f_i(x_{n+1}) = \theta$ for $1 \le i \le n+1$. Now continue inductively.

(ii) $\implies$ (iii) and (iv) $\implies$ (v) are clear.

We next show (ii) $\implies$ (iv) and (iii) $\implies$ (v). Fix $\theta \in (0,1)$. Assume $\exists (x_i)$ in B_X, (f_i) in B_{X^*} such that

$$f_i(x_j) = \begin{cases} \theta & \text{if } i \le j \\ 0 & \text{if } i > j. \end{cases}$$

Given $n \in \mathbb{N}$ and finite convex combinations $\sum_{i=1}^n t_i x_i$ and $\sum_{i=n+1}^\infty t_i x_i$, we have

$$\left\| \sum_{i=n+1}^\infty t_i x_i - \sum_{i=1}^n t_i x_i \right\| \ge \left| f_{n+1}\Big(\sum_{i=n+1}^\infty t_i x_i - \sum_{i=1}^n t_i x_i \Big) \right| = \sum_{i=n+1}^\infty \theta t_i = \theta.$$

Thus

$$d(\mathrm{conv}\{x_1, \dots, x_n\}, \mathrm{conv}\{x_{n+1}, x_{n+2}, \dots\}) \ge \theta.$$

Finally, we show (v) $\implies$ (i). Assume $\exists \theta \in (0,1)$ and (x_i) in B_X such that (v) holds. Assume for a contradiction that X is reflexive.

For $n \in \mathbb{N}$, let $C_n = \mathrm{conv}\{x_{n+1}, x_{n+2}, \dots\}$. $\overline{C_n}$ ($\|\cdot\|$-closure) is a $\|\cdot\|$-closed, convex subset of B_X. Hence $\overline{C_n}$ is a w-closed subset of B_X. Also $\overline{C_1} \supset \overline{C_2} \supset \overline{C_3} \supset \dots$ and $\overline{C_n} \ne \emptyset$ for all n. Since B_X is w-compact, we have $\bigcap_{n=1}^\infty \overline{C_n} \ne \emptyset$, say it contains x. Since $x \in \overline{C_1}$, there exists $y \in C_1$ such that $\|x - y\| < \frac{\theta}{3}$. Choose n such that $y \in \mathrm{conv}\{x_1, x_2, \dots x_n\}$. Since $x \in \overline{C_n}$, there exists $z \in C_n$ such that $\|x - z\| < \frac{\theta}{3}$. Then

$$\theta \le d(\mathrm{conv}\{x_1, \dots, x_n\}, \mathrm{conv}\{x_{n+1}, x_{n+2}, \dots\}) \le \|y - z\| \le \frac{2\theta}{3},$$

a contradiction. □

Ultrafilters

Fix a set $I \ne \emptyset$. A *filter* on I is a family $\mathcal{F} \subset \mathcal{P}(I)$ such that

(i) $I \in \mathcal{F}$, $\emptyset \notin \mathcal{F}$;

(ii) $A \subset B \subset I, A \in \mathcal{F} \implies B \in \mathcal{F}$;

(iii) $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$.

Remark. One can think of $\mathcal{F}$ as "big sets", or "full-measure".

Example. (i) For $i \in I$, $\mathcal{U}_i = \{A \subset I : i \in A\}$ is a filter – the *principal filter* at i.

(ii) If $|I| = \infty$, then $\{A \subset I : |I \setminus A| < \infty\}$ is a filter – the *cofinite filter* on I.

Definition. If X is a topological space, $f\colon I \to X$ is a function, $\mathcal{F}$ is a filter on I, $x \in X$, then we write $x = \lim_{\mathcal{F}} f$ if for all neighbourhoods U of x in X, $\{i \in I : f(i) \in U\} \in \mathcal{F}$.

Example.

(i) If $I = \mathbb{N}$, $\mathcal{F} =$ cofinite filter on $\mathbb{N}$, then this is just the usual notion of convergence of a sequence.

(ii) If X is Hausdorff and $x = \lim_{\mathcal{F}} f, y = \lim_{\mathcal{F}} f$, then $x = y$.

(iii) If $\mathcal{F} = \mathcal{U}_i$ for some $i \in I$, then $f(i) = \lim_{\mathcal{F}} f$ holds for all $f\colon I \to X$.

Definition. Let $I \neq \emptyset$ be a set. An *ultrafilter* on I is a maximal filter on I with respect to inclusion: it's a filter $\mathcal{U}$ such that if $\mathcal{F}$ is a filter and $\mathcal{U} \subset \mathcal{F}$ then $\mathcal{U} = \mathcal{F}$.

Example. Any principal filter $\mathcal{U}_i = \{A \subset I : i \in A\}$ is an ultrafilter. If I is finite, then these are the only ones.

In general, any filter is contained in an ultrafilter (use Zorn's lemma).

Definition. A *free ultrafilter* is an ultrafilter that is not a principal ultrafilter.

Example. Any ultrafilter containing the cofinite filter is a free ultrafilter ($|I| = \infty$).

Lemma 6.4. Let $\mathcal{U}$ be an ultrafilter. If $A \cup B \in \mathcal{U}$ then $A \in \mathcal{U}$ or $B \in \mathcal{U}$.

Proof. Assume otherwise, that $\exists C, D \in \mathcal{U}$ such that $A \cap C = B \cap D = \emptyset$. Then $(A \cup B) \cap (C \cap D) = \emptyset$, a contradiction, as $A \cup B, C \cap D \in \mathcal{U}$.

WLOG $A \cap C \neq \emptyset$ for all $C \in \mathcal{U}$. Then

$$\{D \subset I : \exists C \in \mathcal{U}, D \supset A \cap C\}$$

is a filter on I and it contains $\mathcal{U}$, so equals $\mathcal{U}$. So $A \in \mathcal{U}$. □

Remark.

(i) Every free ultrafilter contains the cofinite filter. [For any finite set $A \subset I$, consider $A \cup A^c$ in the lemma above.]

(ii) For an ultrafilter $\mathcal{U}$, define $\mu\colon \mathcal{P}(I) \to \{0, 1\}$ by $\mu(A) = 1_{A \in U}$. Then μ is a finitely additive measure.

Lemma 6.5. Let $\mathcal{U}$ be an ultrafilter and K be a compact topological space. Then for every function $f\colon I \to K$ there exists $x \in K$ such that $x = \lim_{\mathcal{U}} f$ (might not be unique, but if K is Hausdorff then it is). In particular, for every bounded function $f\colon I \to \mathbb{R}$ there exists a unique $x \in \mathbb{R}$ such that $x = \lim_{\mathcal{U}} f$.

Proof. If not, then $\forall x \in K$, $\exists$ open neighbourhood V_x of x such that $A_x = \{i \in I : f(i) \in V_x\} \notin \mathcal{U}$. Since K is compact, there exists a finite $F \subset K$ such that $\bigcup_{x \in F} V_x = K$. Then $\bigcup_{x \in F} A_x = I \in \mathcal{U}$ and by Lemma 4, $\exists x \in F$ such that $A_x \in \mathcal{U}$, a contradiction. □

Remark. Given bounded functions $f, g\colon I \to \mathbb{R}$ we have

$$\lim_{\mathcal{U}}(f + g) = \lim_{\mathcal{U}} f + \lim_{\mathcal{U}} g,$$
$$\lim_{\mathcal{U}}(fg) = \Big(\lim_{\mathcal{U}} f\Big)\Big(\lim_{\mathcal{U}} g\Big),$$

and if $f(i) \leq g(i)$ for all $i \in I$, then

$$\lim_{\mathcal{U}} f \leq \lim_{\mathcal{U}} g.$$

Ultraproduct and Ultrapowers

Definition. Fix a non-empty set I. We are given Banach spaces X_i, $i \in I$. We fix an ultrafilter $\mathcal{U}$ on I. We let

$$\left(\bigoplus_{i\in I} X_i\right)_\infty = \left\{(x_i)_{i\in I} : x_i \in X_i \,\forall i \in I,\ \sup_{i\in I}\|x_i\| < \infty\right\}.$$

This is a Banach space with norm $\|(x_i)\|_\infty = \sup_{i\in I}\|x_i\|$. Define

$$\|(x_i)\|_{\mathcal{U}} = \lim_{\mathcal{U}}\|x_i\|.$$

This defines a seminorm on $\left(\bigoplus_{i\in I} X_i\right)_\infty$. It follows that

$$\mathcal{N}_{\mathcal{U}} = \{(x_i) : \|(x_i)\|_{\mathcal{U}} = 0\}$$

is a subspace of $\left(\bigoplus_{i\in I} X_i\right)_\infty$, and the quotient $\left(\bigoplus_{i\in I} X_i\right)_\infty/\mathcal{N}_{\mathcal{U}}$ becomes a normed space with norm $\|((x_i)_{i\in I})_{\mathcal{U}}\| = \|(x_i)_{i\in I}\|_{\mathcal{U}}$ where for $x \in \left(\bigoplus_{i\in I} X_i\right)_\infty$, $x_{\mathcal{U}} = x + \mathcal{N}_{\mathcal{U}}$. It is easy to check that this is a complete norm. This Banach space is denoted by $\left(\prod_{i\in I} X_i\right)_{\mathcal{U}}$ — called an *ultraproduct* of $(X_i)_{i\in I}$.

If $X_i = X$ for all $i \in I$ for some Banach space X, then the ultraproduct $\left(\prod_{i\in I} X_i\right)_{\mathcal{U}}$ is denoted by $X^{\mathcal{U}}$ — called an *ultrapower* of X.

Proposition 6.6. Any ultrapower $X^{\mathcal{U}}$ of a Banach space X is finitely representable in X.

Proof. Let E be a finite-dimensional subspace of $X^{\mathcal{U}}$. Choose a basis $e_1, e_2, \dots, e_n$ of E. For each $1 \le k \le n$, fix $(x_{k,i})_{i\in I}$, a bounded sequence in X, such that $e_k = ((x_{k,i})_i)_{\mathcal{U}}$. So $\forall (\lambda_k)_{k=1}^n$ in $\mathbb{R}^n$, $\sum \lambda_k e_k = ((\sum \lambda_k x_{k,i})_i)_{\mathcal{U}}$.

Fix $\epsilon > 0$. We seek an injective linear map $T\colon E \to X$ such that $\|T\|\cdot\|T^{-1}\| < 1+\epsilon$ (here $T^{-1}\colon T(E) \to E$). Choose $\delta \in (0, \frac{1}{3})$ such that $\frac{1+\delta}{1-3\delta} < 1+\epsilon$. Let $S \subset \mathbb{R}^n$ be a finite set such that $\tilde{S} = \{\sum_{k=1}^n \lambda_k e_k : (\lambda_k)_{k=1}^n \in S\}$ is a δ-net of S_E.

Since $\|\sum_{k=1}^n \lambda_k e_k\|_{\mathcal{U}} = \lim_{\mathcal{U}} \|\sum_{k=1}^n \lambda_k x_{k,i}\| = 1$ for all $(\lambda_k) \in S$, we have

$$\left\{ i \in I : 1-\delta < \left\|\sum_{k=1}^n \lambda_k x_{k,i}\right\| < 1+\delta \right\} \in \mathcal{U}.$$

Since S is finite, these sets have intersection in $\mathcal{U}$. In particular, $\exists i_0 \in I$ such that

$$1-\delta < \left\|\sum_{k=1}^n \lambda_k x_{k,i_0}\right\| < 1+\delta \qquad \forall (\lambda_k) \in S.$$

Now define $T\colon E \to X$, $T(\sum_{k=1}^n \mu_k e_k) = \sum_{k=1}^n \mu_k x_{k,i_0}$, $(\mu_k) \in \mathbb{R}^n$. Given $x \in S_E$, $\exists z \in \tilde{S}$ such that $\|x - z\| \le \delta$. So

$$\|Tx\| \le \|Tz\| + \|T(x-z)\| \le (1+\delta) + \|T\|\delta.$$

Taking sup over $x \in S_E$, $\|T\| \le 1+\delta+\delta\|T\|$, so $\|T\| \le \frac{1+\delta}{1-\delta}$. It follows that $\|Tx\| \ge \|Tz\| - \|T(x-z)\| \ge 1-\delta-\frac{1+\delta}{1-\delta}\delta = \frac{1-3\delta}{1-\delta}$. Hence $\left\|T^{-1}\right\| \le \frac{1-\delta}{1-3\delta}$ and $\|T\|\,\|T^{-1}\| \le \frac{1+\delta}{1-3\delta} < 1+\epsilon$. $\square$

Theorem 6.7. Let X be a Banach space. Then X is superreflexive $\iff$ whenever Y is crudely finitely representable in X, then Y is reflexive.

Proof. ($\Longleftarrow$): clear from definition. ($\Longrightarrow$): assume Y is non-reflexive and crudely finitely representable in X. Fix $\theta \in (0,1)$. By Theorem 3, $\exists (y_i)_{i=1}^\infty$ in B_Y such that $\forall n$,

$$d(\mathrm{conv}(y_1, \dots, y_n), \mathrm{conv}(y_{n+1}, y_{n+2}, \dots)) \geq \theta.$$

There exists $\lambda > 1$ such that $\forall$ subspace $E \subset Y$, $\dim E < \infty$, $\exists$ linear $T \colon E \to X$ such that

$$\lambda^{-1} \|y\| \leq \|Ty\| \leq \|y\| \qquad \forall y \in E.$$

For $N \in \mathbb{N}$, $\exists$ linear map $T_N \colon \mathrm{span}(y_1, \dots, y_N) \to X$ such that

$$\lambda^{-1} \|y\| \leq \|T_N y\| \leq \|y\| \qquad \forall y \in \mathrm{span}(y_1, \dots, y_N).$$

Let $x_{N,i} = T_N(y_i)$ for $1 \leq i \leq N$. Note that for $1 \leq m < n \leq N$ and for convex combinations $\sum_{i=1}^m t_i x_{N,i}$, $\sum_{i=m+1}^n t_i x_{N,i}$, we have

$$\left\| \sum_{i=1}^m t_i x_{N,i} - \sum_{i=m+1}^n t_i x_{N,i} \right\| \geq \frac{1}{\lambda} \left\| \sum_{i=1}^m t_i y_i - \sum_{i=m+1}^n t_i y_i \right\| \geq \frac{\theta}{\lambda}.$$

Note also that $\|x_{N,i}\| \leq 1$ for all $1 \leq i \leq N$. WLOG replace θ/λ by θ. Now fix a free ultrafilter $\mathcal{U}$ on $\mathbb{N}$. Define

$$\tilde{x}_{N,i} = \begin{cases} x_{N,i} & \text{if } i \leq N \\ 0 & \text{if } i > N, \end{cases} \qquad \tilde{x}_i = ((\tilde{x}_{N,i})_{N=1}^\infty)_{\mathcal{U}}.$$

Given $1 \leq m < n$ and convex combinations $z = \sum_{i=1}^m t_i \tilde{x}_i$ and $w = \sum_{i=m+1}^n t_i \tilde{x}_i$ in $X^{\mathcal{U}}$, we have $\forall N \in \mathbb{N}, N \geq n$,

$$\left\| \sum_{i=1}^m t_i \tilde{x}_{N,i} - \sum_{i=m+1} t_i \tilde{x}_{N,i} \right\| \geq \theta.$$

It follows that $\|z - w\| \geq \theta$. Then

$$d(\mathrm{conv}\{\tilde{x}_1, \dots, \tilde{x}_m\}, \mathrm{conv}\{\tilde{x}_{m+1}, \dots\}) \geq \theta.$$

By Theorem 3, $X^{\mathcal{U}}$ is non-reflexive. By Proposition 6, $X^{\mathcal{U}}$ is finitely representable in X, and hence X is not superreflexive. $\square$

Definition. A Banach space X is *strictly convex* if $\forall x, y \in S_X$, $x \neq y$, $\left\| \frac{x+y}{2} \right\| < 1$. Say X is *uniformly convex* if $\forall \epsilon \in (0,2]$, $\exists \delta > 0$, $\forall x, y \in S_X$, $\|x - y\| \geq \epsilon \implies 1 - \left\| \frac{x+y}{2} \right\| \geq \delta$. The *modulus of uniform convexity* of X is the function $\delta_X \colon [0,2] \to \mathbb{R}^+$ defined by

$$\delta_X(\epsilon) = \inf\{1 - \left\| \frac{x+y}{2} \right\| : x, y \in S_X, \|x - y\| \geq \epsilon\}.$$

Example. (i) ℓ_2 is uniformly convex: given $x, y \in S_{\ell_2}$ with $\|x - y\| \geq \epsilon$, we have, by the parallelogram rule,

$$4 = 2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2 \geq \|x+y\|^2 + \epsilon^2.$$

So $1 - \left\| \frac{x+y}{2} \right\| \geq 1 - \sqrt{1 - \frac{\epsilon^2}{4}} \approx \frac{\epsilon^2}{8}$.

(ii) Choose $1 < p_n < 2$, $p_n \to 1$. Let $X = \left(\bigoplus_{n=1}^{\infty} \ell_{p_n}^2\right)_{\ell_2}$. Then X is strictly convex, but not uniformly convex. However, $X \sim (\bigoplus_{n=1}^{\infty} \ell_2^n)_{\ell_2} \cong \ell_2$. So uniform convexity is not an isomorphic property.

(iii) c_0, ℓ_1, ℓ_∞ are not strictly convex.

Theorem 6.8 (Milman-Pettis)**.** If X is uniformly convex, then X is reflexive.

Remark. Recall Goldstine's Theorem: $\overline{B_X}^{w^*} = B_{X^{**}}$. In fact, if $\dim X = \infty$, then $\overline{S_X}^{w^*} = B_{X^{**}}$.

Proof. Let $\varphi \in B_{X^{**}}$ and U be a w^*-neighbourhood of φ. WLOG $\exists n \in \mathbb{N}$, $f_1, \dots, f_n \in X^*$, $\epsilon > 0$ such that $U = \{\psi \in X^{**} : |(\psi - \varphi)(f_i)| < \epsilon, \forall i\}$. Choose $x \in B_X \in U$ by Goldstine. Fix $z \in \bigcap_{i=1}^n \ker f_i$, $z \neq 0$ ($\dim X = \infty$). Then $x + \lambda z \in U$ $\forall \lambda \in \mathbb{R}$, and $\exists \lambda \in \mathbb{R}$ such that $\|x + \lambda z\| = 1$. □

Proof of Theorem 8. WLOG $\dim X = \infty$. Fix $\varphi \in S_{X^{**}}$. We show that $\varphi \in X$. Then we'll be done. Fix $\epsilon \in (0, 2)$ and let $\delta = \delta_X(\epsilon) > 0$. Then $\forall x, y \in S_X$ if $\|x + y\| \geq 2 - \delta$, then $1 - \left\|\frac{x+y}{2}\right\| \leq \frac{\delta}{2} < \delta$, and hence $\|x - y\| < \epsilon$. Choose $f_\epsilon \in B_{X^*}$ such that $\varphi(f_\epsilon) > 1 - \frac{\delta}{2}$. Let $V_\epsilon = \{\psi \in X^{**} : \psi(f_\epsilon) \geq 1 - \frac{\delta}{2}\}$. This is a w^*-closed neighbourhood of φ. Hence $W_\epsilon = V_\epsilon \cap S_X$ is non-empty and $\|\cdot\|$-closed subset of X. Also, given $x, y \in W_\epsilon$, $\|x + y\| \geq f_\epsilon(x + y) \geq 2 - \delta$, and hence $\|x - y\| < \epsilon$. Thus, $\operatorname{diam}(W_\epsilon) \leq \epsilon$. Now for $n \in \mathbb{N}$, let

$$A_n = \bigcap_{k=1}^{n} W_{1/k} = \{\psi \in X^{**} : \psi(f_{1/k}) \geq 1 - \frac{\delta_X(1/k)}{2} \text{ for } k = 1, \dots, n\} \cap S_X.$$

So A_n is a non-empty, $\|\cdot\|$-closed subset of X of diameter at most $\operatorname{diam}(W_{1/n}) \leq \frac{1}{n}$. Also, $A_n \supset A_{n+1}$ for all n, and X is complete, so by Cantor's intersection Theorem, $\cap_{n=1}^{\infty} A_n = \{x\}$ for some $x \in S_X$.

We show that $\varphi = \hat{x}$. If not, then $\exists g \in X^*$, $\eta = \varphi(g) - g(x) > 0$. Let

$$\begin{aligned} B_n &= A_n \cap \{\psi : |\varphi(g) - \psi(g)| \leq \frac{\eta}{2}\} \\ &= \underbrace{\{\psi : \psi(f_{1/k}) \geq 1 - \frac{\delta_X(1/k)}{2} \text{ for } k = 1, \dots, n, \ |\varphi(g) - \psi(g) \leq \frac{\eta}{2}\}}_{w^*\text{-closed neighbourhood of } \varphi} \cap S_X, \end{aligned}$$

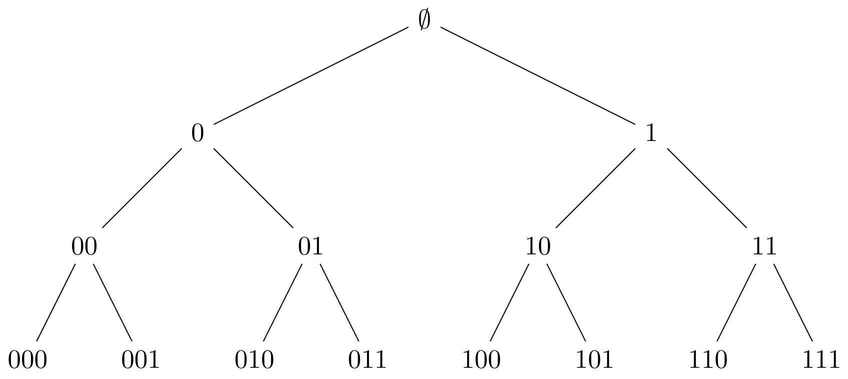
so B_n is nonempty, $\|\cdot\|$-closed and $\operatorname{diam}(B_n) \leq \operatorname{diam}(A_n) \to 0$. So $\cap_{n=1}^{\infty} B_n = \{x\}$, so $|\varphi(g) - g(x)| \leq \frac{\eta}{2}$, a contradiction. □

Fact (Enflo)**.** $(X, \|\cdot\|)$ is superreflexive $\iff$ $\exists$ equivalent norm $\|\cdot\|'$ on X such that $(X, \|\cdot\|')$ is uniformly convex. Recall norm equivalence means $\exists a, b > 0$ such that

$$a\|x\| \leq \|x\|' \leq b\|x\|.$$

Example. $\ell_2 \oplus_2 \ell_1^2 \sim \ell_2 \oplus_2 \ell_2^2 \cong \ell_2$, which is superreflexive but $\ell_2 \oplus_2 \ell_1^2$ is not strictly convex.

Recall that the *binary tree* of depth n, B_n, has vertex set $\cup_{k=0}^{n} \{0, 1\}^k$ and $\epsilon = (\epsilon_1, \dots, \epsilon_k) \in \{0, 1\}^k$, $k < n$, is joined to $(\epsilon_1, \dots, \epsilon_k, i)$, $i = 0, 1$.



Notation. Given $\epsilon = (\epsilon_1, \ldots, \epsilon_k)$, $\delta = (\delta_1, \ldots, \delta_\ell)$, we write $\epsilon \preceq \delta$ if $k \leq \ell$ and $\epsilon_i = \delta_i$ for $1 \leq i \leq k$. We also let $|\epsilon| = k$ denote the *length of* ϵ.

Definition. We say a Banach space X has the *finite tree property* if $\exists \theta > 0$, $\forall n \in \mathbb{N}$, $\exists \{x_\epsilon : \epsilon \in B_n\} \subset B_X$ such that $x_\epsilon = \frac{1}{2}(x_{\epsilon 0} + x_{\epsilon 1})$ for all $\epsilon \in B_{n-1}$, $\|x_\epsilon - x_{\epsilon i}\| \geq \theta$ $\forall \epsilon \in B_{n-1}$, $i = 0, 1$.

Theorem 6.9. For a Banach space, the following are equivalent:

(a) X is not superreflexive;

(b) X has the finite tree property;

(c) $\exists \theta > 0$, $\forall n \in \mathbb{N}$, $\exists \{x_1, \ldots, x_n\} \subset B_X$ such that

$$\left\| \sum_{i=1}^{n} a_i x_i \right\| \geq \theta \left| \sum_{i=\ell}^{m} a_i \right| \quad \forall a_1, \ldots, a_n \in \mathbb{R}, \quad 1 \leq \ell \leq m \leq n.$$

Remark. Let $S = \{(a_i)_{i=1}^\infty \subset \mathbb{R} : \sum_{i=1}^\infty a_i \text{ is convergent}\}$. This becomes a normed space with

$$\|(a_i)\| = \sup \left\{ \left| \sum_{i=\ell}^{m} a_i \right| : 1 \leq \ell \leq m \right\}.$$

This is called the *summing norm*. Note $S \sim c_0$, via the map

$$(a_i)_{i=1}^\infty \longmapsto \left(\sum_{i=n}^{\infty} a_i \right)_{n=1}^{\infty}.$$

Definition. Given a convex set C in a Banach space Z, a point $w \in C$ is *strongly exposed* if $\exists f \in Z^*$ such that

(i) $f(u) < f(w) \quad \forall u \in C$, $u \neq w$;

(ii) $\operatorname{diam}\{u \in C : f(w) - \epsilon < f(u)\} \to 0$ as $\epsilon \to 0$.

Theorem 6.10. Every non-empty, w-compact convex subset of a separable Banach space has a strongly exposed point.

Proof. Omitted. Theorem is also true for non-separable spaces. □

Proof of Theorem 9. (a) $\Longrightarrow$ (b): There exists a non-reflexive Z finitely representable in X. Fix $\theta \in (0,1)$. By Theorem 3, $\exists (z_i)$ in B_Z such that $d(\operatorname{conv}\{z_1,\dots,z_n\}, \operatorname{conv}\{z_{n+1},\dots\}) \geq \theta$ for all $n \in \mathbb{N}$. For $\epsilon = (\epsilon_1,\dots,\epsilon_n) \in B_n$, let $k(\epsilon) = 1 + \sum_{i=1}^n 2^{n-i}\epsilon_i$. This is an enumeration of the leaves. For $\delta \in B_n$, let $I_\delta = \{k(\epsilon) : \delta \preceq \epsilon, |\epsilon| = n\}$, set of nth generation descendants of δ. Let $z_\delta = 2^{|\delta|-n}\sum_{k\in I_\delta} z_k$. Since $|I_\delta| = 2^{n-|\delta|}$, we have $z_\delta \in \operatorname{conv}\{z_k : k \in I_\delta\} \subset B_Z$. For $\delta \in B_{n-1}$, $I_\delta = I_{\delta,0} \cup I_{\delta,1}$ and $I_{\delta,0} \cap I_{\delta,1} = \emptyset$, and moreover, $\forall k \in I_{\delta,0}$, $\forall \ell \in I_{\delta,1}$, $k < \ell$. It follows that $z_\delta = \frac{1}{2}(z_{\delta 0} + z_{\delta 1})$, and for $i = 0,1$, we have $\|z_\delta - z_{\delta,i}\| = \frac{1}{2}\|z_{\delta 0} - z_{\delta 1}\| \geq \frac{1}{2} d(\operatorname{conv}\{z_k : k \in I_{\delta 0}\}, \operatorname{conv}\{z_k : k \in I_{\delta 1}\}) \geq \frac{\theta}{2}$. So Z has the finite tree property, and hence so does X since Z is finitely representable in X.

(b) $\Longrightarrow$ (a): $\exists \theta > 0$, $\forall n$, $\exists \{x_\epsilon^n : \epsilon \in B_n\} \subset B_X$ such that $x_\epsilon^n = \frac{1}{2}(x_{\epsilon 0}^n + x_{\epsilon 1}^n)$ $\forall \epsilon \in B_{n-1}$ and $\|x_\epsilon^n - x_{\epsilon i}^n\| \geq \theta$ $\forall \epsilon \in B_{n-1}$, $i = 0,1$. Let $\mathcal{U}$ be a free ultrafilter and let B_∞ be the ∞ binary tree with vertex set $\bigcup_{k=0}^\infty \{0,1\}^k$ and ϵ joined to ϵi $\forall \epsilon \in B_\infty$, $i = 0,1$. Let

$$\tilde{x}_\epsilon^n = \begin{cases} x_\epsilon^n & \text{if } |\epsilon| \leq n \\ 0 & \text{if } n < |\epsilon|. \end{cases} \quad \text{and} \quad \tilde{x}_\epsilon = ((\tilde{x}_\epsilon^n)_n)_{\mathcal{U}}.$$

It's easy to see that $\tilde{x}_\epsilon = \frac{1}{2}(\tilde{x}_{\epsilon 0} + \tilde{x}_{\epsilon 1})$ and $\|\tilde{x}_\epsilon - \tilde{x}_{\epsilon i}\| \geq \theta$ $\forall \epsilon \in B_\infty$, $i = 0,1$. Let $Z = \overline{\operatorname{span}}\{\tilde{x}_\epsilon : \epsilon \in B_\infty\}$. This is a separable subspace of $X^{\mathcal{U}}$. Assume for contradiction that X is superreflexive. Then by Proposition 6, Z is reflexive. Then B_Z is w-compact. Let $C = \overline{\operatorname{conv}}\{\tilde{x}_\epsilon : \epsilon \in B_\infty\}$. Then C is a $\|\cdot\|$-closed convex subset of B_Z, and hence w-compact. By Theorem 10, C has a strongly exposed point w. So $\exists f \in Z^*$ such that $f(u) < f(w)$ $\forall u \in C, u \neq w$ and $\exists \eta > 0$ $\{u \in C : f(u) > f(w) - \eta\}$ has diameter $< \frac{\theta}{2}$. Since $\{u \in C : f(u) \leq f(w) - \eta\}$ is $\|\cdot\|$-closed and convex and $\subsetneq C$, it cannot contain $\tilde{x}_\epsilon$ $\forall \epsilon$. So $\exists \epsilon \in B_\infty$ such that $f(\tilde{x}_\epsilon) > f(w) - \eta$. Then $\frac{1}{2}(f(\tilde{x}_{\epsilon 0}) + f(\tilde{x}_{\epsilon 1})) = f(\tilde{x}_\epsilon)$, so $\exists i \in \{0,1\}$ such that $f(\tilde{x}_{\epsilon i}) > f(w) - \eta$. Thus $\|\tilde{x}_\epsilon - \tilde{x}_{\epsilon i}\| < \frac{\theta}{2}$, a contradiction.

(a) $\Longrightarrow$ (c): Let Z be non-reflexive and finitely representable in X. By Theorem 2, $\exists \theta \in (0,1)$ and (z_i) in B_Z, (h_i) in B_{Z^*} such that

$$h_i(z_j) = \begin{cases} \theta & i \leq j \\ 0 & i > j. \end{cases}$$

Given scalars $(a_i)_{i=1}^n$, $|\sum_{i=\ell}^n a_i| = \left|\frac{1}{\theta} h_\ell\left(\sum_{i=1}^n a_i z_i\right)\right| \leq \frac{1}{\theta}\|\sum_{i=1}^n a_i z_i\|$. If $1 \leq \ell \leq m \leq n$, then

$$\left|\sum_{i=\ell}^m a_i\right| \leq \left|\sum_{i=\ell}^n a_i\right| + \left|\sum_{i=m+1}^n a_i\right| \leq \frac{2}{\theta}\left\|\sum_{i=1}^n a_i z_i\right\|.$$

Since Z is finitely representable in X, $\forall \lambda > \frac{2}{\theta}$, $\forall n$, $\exists x_1,\dots,x_n \in B_X$ such that

$$\left|\sum_{i=\ell}^m a_i\right| \leq \lambda\left\|\sum_{i=1}^n a_i x_i\right\| \qquad \forall a_1,\dots,a_n \in \mathbb{R}, 1 \leq \ell \leq m.$$

(c) $\Longrightarrow$ (a):$\exists \theta > 0$, $\forall n \in \mathbb{N}$, $\exists \{x_1^n,\dots,x_n^n\} \subset B_X$ such that

$$\left\|\sum_{i=1}^n a_i x_i^n\right\| \geq \theta\left|\sum_{i=\ell}^m a_i\right| \qquad \forall a_1,\dots,a_n \in \mathbb{R}, \quad 1 \leq \ell \leq m \leq n.$$

Given a free ultrafilter $\mathcal{U}$ on $\mathbb{N}$, the usual process yields an infinite sequence $(\tilde{x}_i)_{i=1}^\infty$ in $B_{X^\mathcal{U}}$ such that $\forall n \in \mathbb{N}$, $\forall a_1, \dots, a_n \in \mathbb{R}$, $\forall 1 \le \ell \le m \le n$,

$$\left\| \sum_{i=1}^n a_i \tilde{x}_i \right\| \ge \theta \left| \sum_{i=\ell}^m a_i \right|.$$

It follows that $\forall i \in \mathbb{N}$,

$$h_i(\tilde{x}_j) = \begin{cases} \theta & i \le j \\ 0 & i > j \end{cases}$$

extends to a well-defined linear functional on $X^\mathcal{U}$ with $\|h_i\| \le 1$ [also uses Hahn-Banach]. By Theorem 3, $X^\mathcal{U}$ is not reflexive. By Proposition 6, $X^\mathcal{U}$ is finitely representable in X, so X is not superreflexive. □

Theorem 6.11 (Metric Characterization of Superreflexivity)**.** Let X be a Banach space. The following are equivalent:

(a) X not superreflexive;

(b) The sequence (D_n) of diamond graphs embeds uniformly bilipschitzly into X.

Sketch proof. (non-examinable) (b) $\implies$ (a): Have $f_n \colon D_n \to X$ $\sup_n \operatorname{dist}(f_n) < \infty$. WLOG $\exists \delta > 0$, $\forall n$, $\forall x, y \in D_n$, $\delta 2^{-n} d_n(x, y) \le \|f_n(x) - f_n(y)\| \le 2^{-n} d_n(x, y)$. Let $D_0 = tb$, $D_1 = tb\ell r$, and D_n is a union of 4 copies of D_{n-1}. Fix n, $f = f_n$. Let $x_\emptyset = f(t) - f(b)$. Then $\|x_\emptyset\| \le 2^{-n} d_n(t, b) = 1$. Consider $\|[(f(t) - f(\ell)) - (f(\ell) - f(b))] - [(f(t) - f(r)) - (f(r) - f(b))]\| = \|2(f(r) - f(\ell))\| \ge 2\delta 2^{-n} d_n(\ell, r) = 2\delta$. WLOG $\|(f(t) - f(\ell)) - (f(\ell) - f(b))\| \ge \delta$. Let $x_0 = 2(f(\ell) - f(b))$, $x_1 = 2(f(t) - f(\ell))$. Then $x_\emptyset = \frac{1}{2}(x_0 + x_1)$ and $\|x_\emptyset - x_0\| = \frac{1}{2}\|x_1 - x_0\| \ge \delta$. Continue inductively.

(a) $\implies$ (b): $\exists \theta > 0$, $\forall n$, $\exists x_1, \dots, x_{2^n} \in B_X$ with lower summing norm estimate. First embed $f_n \colon D_n \to \{0, 1\}^{2^n} \subset \ell_1^{2^n}$. For D_0, do $t = 1, b = 0$. For D_1, do $t = 11, \ell = 01, b = 00, r = 10$. If $xy \in E_{n-1}$, $f_{n-1}(x), f_{n-1}(y) \in \{0, 1\}^{2^{n-1}}$ differ in one digit, say j. Consider $yuxv$ in D_n. If $\nu \in \{x, y, u, v\}$, $(f_n(\nu))_{2i-1} = (f_n(\nu))_{2i} = (f_{n-1}(x))_i$. $f_n(\nu)_{2j-1}, f_n(\nu)_{2j}$ will be $00, 11, 01, 10$ for $\nu = x, y, u, v$ $(f_{n-1}(x))_j = 0$.

Let $g_n \colon D_n \to X$ given by

$$g_n(x) = \sum_{j=1}^{2^n} \epsilon_j x_j, \qquad (\epsilon_j) = f_n(x).$$

If x is in top left, y is in bottom right, then$f_n(x) = (f_{n-1}(x), \underbrace{1, \dots, 1}_{2^{n-1}})$, $f_n(y) = (f_{n-1}(y), \underbrace{0, \dots, 0}_{2^{n-1}})$. □

Exam will be 4 questions, answer 3 in 3 hours. Mostly bookwork.