On Stochastic Decompositions of Metric Spaces

Scribe of a talk given at the fall school ”Metric Embeddings : Constructions and Obstructions”. 3-7 November 2014, Paris, France

Ofer Neiman*

Abstract

In this talk we will go over stochastic metric decompositions. These are random partitioning of a metric space into pieces of bounded diameter, such that for each point, a certain ball centered at it has a good chance of being contained in a single cluster. These decompositions play a role in metric embedding, metric Ramsey theory, higher-order Cheeger inequalities for graphs, metric and Lipschitz extension problems and approximation algorithms. We will then see an example of such decompositions for doubling metric spaces.

The second part of the talk will be devoted to embedding finite metrics into normed spaces using these decompositions. We will begin with a basic result due to Rao, and time permitting, the Measure Descent approach.

1 Preliminaries

Let (X, d) be a metric space. For $x \in X$ and $r \geq 0$, denote by $B_X(x, r) = \{z \in X \ : \ d(x, z) \leq r\}$ the closed ball of radius r centered at x (we omit the subscript when clear from context). By $B^\circ(x, r) = \{z \in X \ : \ d(x, z) < r\}$ we mean the open ball. The *diameter* of X is denoted as $\mathrm{diam}(X) = \sup_{x,y\in X} d(x, y)$, and its aspect ratio $\Phi(X) = \frac{\sup_{x,y\in X} d(x,y)}{\inf_{x,y\in X} d(x,y)}$.

Distortion. If (X, d_X) and (Y, d_Y) are metric spaces, a mapping $f : X \to Y$ has distortion at most K if there exists $C > 0$ such that for any $x, y \in X$,

$$\frac{C}{K} \cdot d_X(x, y) \leq d_Y(f(x), f(y)) \leq C \cdot d_X(x, y) \ .$$

The infimum K is called the *distortion* of f. We denote by $c_Y(X)$ the smallest distortion of a mapping from X to Y. In the special case where $Y = \ell_p$ we denote the smallest distortion by $c_p(X)$.

*Ben-Gurion University of the Negev, Israel. Email: `neimano@cs.bgu.ac.il`. Partially funded by the Lynne and William Frankel Center for Computer Sciences.

Stochastic Decompositions. A partition P of X is a pairwise disjoint collection of clusters that covers X. We say that the partition is Δ-bounded if for any cluster $C \in P$, $\mathrm{diam}(C) \leq \Delta$. For $x \in X$, let $P(x)$ denote the unique cluster containing x in P. Denote by $\mathcal{P}$ the collection of all partitions of X. For a distribution $\Pr$ over $\mathcal{P}$, recall that $\mathrm{supp}(\Pr) = \{P \in \mathcal{P} : \Pr[P] > 0\}$.

Definition 1 (Padded-Decomposition)**.** *A stochastic decomposition of a metric space* (X, d) *is a distribution* $\Pr$ *over* $\mathcal{P}$*. Given* $\Delta > 0$*, the decomposition is called* Δ*-bounded if for all* $P \in \mathrm{supp}(\Pr)$*,* P *is* Δ*-bounded. For a function* $\epsilon : X \to [0, 1]$*, a* Δ*-bounded decomposition is called* ϵ-padded *if the following condition holds:*

- *For all* $x \in X$*,* $\Pr[B(x, \epsilon(x) \cdot \Delta) \subseteq P(x)] \geq 1/2$.

Definition 2 (Modulus of Decomposability)**.** *A metric space* (X, d) *is called* α*-decomposable if for every* $\Delta > 0$ *there exists a* Δ*-bounded stochastic decomposition of* X *with padding parameter* $\epsilon(x) = 1/\alpha$*, for all* $x \in X$*. The* modulus of decomposability *of* X *is defined as*

$$\alpha_X = \inf\{\alpha : X \text{ is } \alpha\text{-decomposable}\} .$$

Let $\mathcal{X}$ be a family of metric spaces. If every member of the family has modulus of decomposability at most β, then we say that the family $\mathcal{X}$ is β-decomposable.

In general, every finite metric space has $\alpha_X \leq O(\log |X|)$ [Bar96], which is quantitatively the best possible, as exhibited by the family of expander graphs. However, there are many families of metric space which are decomposable (that is, $O(1)$-decomposable), as described in the next section.

2 Examples of Decomposable Metric Spaces

There are several families of metric spaces which are known to be decomposable, for instance, metrics arising from shortest path on planar graphs, or bounded tree-width graphs, and in general all graphs excluding some fixed minor. Another example is the family of metrics with bounded Negata-Assouad dimension, which contains doubling metric spaces, subsets of compact Riemannian surfaces, Gromov hyperbolic spaces of bounded local geometry, Euclidean buildings, symmetric spaces, and homogeneous Hadamard manifolds. Here for the sake of simplicity, we show the decomposability of the family of doubling metric spaces. The best quantitative result, which is shown here, is due to [GKL03].

2.1 Doubling Metrics

Let λ be a positive integer. A metric space (X, d) has doubling constant λ if for all $x \in X$ and $r > 0$, the ball $B(x, 2r)$ can be covered by λ balls of radius r. The *doubling dimension* of (X, d) is defined as $\log_2 \lambda$.

Comment: The doubling constant may be defined in terms of diameters of sets, rather than radii of balls, but this which affects the dimension by a factor of 2 at most.

Definition 3. *(Nets) For* $r > 0$*, an* r-net *of a metric* (X, d) *is a set* $N \subseteq X$ *satisfying the following properties:*

1. **Packing:** *For every* $u, v \in N$, $d(u,v) > r$.
2. **Covering:** *For every* $x \in X$ *there exists* $u \in N$ *such that* $d(x,u) \le r$.

Proposition 1. *Let* (X,d) *be a metric with doubling constant* λ, *and* N *be an* r*-net of* X. *If* $S \subseteq X$ *is a set of diameter* t, *then*

$$|N \cap S| \le \lambda^{\lceil \log 4t/r \rceil} .$$

Proof. Note that S is contained in a ball of radius $2t$, and that this ball can be covered by λ^k balls of radius $2t/2^k$. Letting $k = \lceil \log 4t/r \rceil$ we get that these small balls have radius at most $r/2$ and thus cannot contain more than a single point of N. □

Theorem 1. *Let* (X,d) *be a metric space with doubling constant* λ, *then* $\alpha_X \le O(\log \lambda)$.

Proof. Fix any $\Delta > 0$, and take N to be a $\Delta/4$-net of X. We now describe the random partition P. Let σ be a random permutation of N, and choose $r \in [\Delta/4, \Delta/2]$ uniformly at random. For each $u \in N$ define a cluster

$$C_u = \{x \in X \ : \ d(x,u) \le r \text{ and } \sigma(u) < \sigma(v) \text{ for all } v \in N \text{ with } d(x,v) \le r\} .$$

In words, every net point in order of σ collects to its cluster all the unassigned points within distance r from it. Then $P = \{C_u\}_{u \in N} \setminus \{\emptyset\}$. Note that this is indeed a Δ-bounded partition, due to the covering property of nets.

Fix some $x \in X$ and let $t = \Delta/(100 \ln \lambda)$, we need to show that the event $\{B(x,t) \nsubseteq P(x)\}$ happens with probability at most $1/2$. Observe that if $u \in N$ has $d(x,u) \ge \Delta$, then $C_u \cap B(x,t) = \emptyset$ for any choice of r (because $r \le \Delta/2$ and $t < \Delta/2$). Let $S = B(x,\Delta) \cap N$, and note that by Proposition 1, $m := |S| \le \lambda^5$. Arrange the points $s_1, s_2, \ldots, s_m \in S$ in order of increasing distance from x. For $j \in [m]$, let I_j be the interval $[d(x,s_j) - t, d(x,s_j) + t]$. We say that the point s_j *cuts* $B(x,t)$ if s_j is the minimal element (of the permutation σ) for which $r \ge d(x,s_j) - t$, and also $r \in I_j$. Observe that if $B(x,t) \nsubseteq P(x)$ then there must be some s_j which cuts $B(x,t)$.

$$\begin{aligned}
\Pr[B(x,t) \nsubseteq P(x)] &\le \sum_{j=1}^{m} \Pr[s_j \text{ cuts } B(x,t)] \\
&\le \sum_{j=1}^{m} \Pr[r \in I_j \wedge \forall_{i<j} \sigma(s_j) < \sigma(s_i)] \\
&= \sum_{j=1}^{m} \Pr[r \in I_j] \cdot \Pr[\forall_{i<j} \sigma(s_j) < \sigma(s_i) \mid r \in I_j] \\
&\le \sum_{j=1}^{m} \frac{2t}{\Delta/4} \cdot \frac{1}{j} \\
&\le \frac{8t}{\Delta} \cdot (1 + \ln m) .
\end{aligned}$$

The third inequality follows from the independent choices of r and σ. Plugging in the estimates for $t = \Delta/(100 \ln \lambda)$ and $m \le \lambda^5$, gives a bound of $1/2$ on the probability, as required. □

3 Embedding Decomposable Metrics into Normed Spaces

In this section we describe an embedding of finite decomposable metrics into ℓ_p space (for any $p \in [1, \infty)$). This is a simplified version of a result of [Rao99], in which the aspect ratio Φ is replaced by n.

Theorem 2. *Let (X, d) be a finite metric space with modulus of decomposability α and aspect ratio Φ, then $c_p(X) = O(\alpha \cdot \log^{1/p} \Phi)$.*

Proof. Let c be a universal constant to be determined later. Assume w.l.o.g that the minimal distance between two distinct points in X is 1 (by appropriate scaling), and thus $\operatorname{diam}(X) = \Phi$. For each $i \in I = \{0, 1, \dots, \lceil \log \Phi \rceil\}$ and $j \in J = [c \log n]$ (where $n = |X|$), let P_{ij} be a 2^i-bounded $1/\alpha$-padded partition sampled from the distribution guaranteed to exist by the decomposability of (X, d). For each $i \in I$, $j \in J$, and each $C \in P_{ij}$ let $\tau(C) \in \{0, 1\}$ be a Bernoulli random variable chosen independently and uniformly. Define a random embedding $f_{ij} : X \to \mathbb{R}$ by

$$f_{ij}(x) = \tau(P_{ij}(x)) \cdot d(x, X \setminus P_{ij}(x)) ,$$

and let $f : X \to \mathbb{R}^{|I| \cdot |J|}$ by $f = \frac{1}{|J|^{1/p}} \bigoplus_{i \in I, j \in J} f_{ij}$.

Expansion. First we bound the expansion of the map f. Fix any $x, y \in X$, and any $i \in I$, $j \in J$. Next we show that $|f_{ij}(x) - f_{ij}(y)| \le d(x, y)$. If it is the case that $P_{ij}(x) = P_{ij}(y)$ then by the triangle inequality

$$f_{ij}(x) - f_{ij}(y) = \tau(P_{ij}(x)) \cdot (d(x, X \setminus P_{ij}(x)) - d(y, X \setminus P_{ij}(x))) \le d(x, y) .$$

Otherwise, if $y \notin P_{ij}(x)$, then

$$f_{ij}(x) - f_{ij}(y) \le f_{ij}(x) \le d(x, X \setminus P_{ij}(x)) \le d(x, y) ,$$

and the bound on the absolute value follows by symmetry. Finally, we obtain that

$$\|f(x) - f(y)\|_p^p = \frac{1}{|J|} \sum_{i=1}^{|I|} \sum_{j=1}^{|J|} |f_{ij}(x) - f_{ij}(y)|^p \le O(d(x, y)^p \cdot \log \Phi) .$$

Contraction. Now we bound the expected contraction of the embedding. Fix any $x, y \in X$, and let $i \in I$ be the unique value such that $2^i < d(x, y) \le 2^{i+1}$. Since P_{ij} is 2^i-bounded, it must be that $P_{ij}(x) \ne P_{ij}(y)$, and as τ is chosen independently, there is probability of $1/4$ for the event $\mathcal{C}_j = \{\tau(P_{ij}(x)) = 1 \wedge \tau(P_{ij}(y)) = 0\}$. Also, by the definition of padded decomposition, we have that the event $\mathcal{D}_j = \{B(x, 2^i/\alpha) \subseteq P_{ij}(x)\}$ happens independently with probability at least $1/2$. Define $\mathcal{E}_j = \mathcal{C}_j \cap \mathcal{D}_j$. Thus with probability at least $1/8$ we have that event $\mathcal{E}_j$ holds and so

$$|f_{ij}(x) - f_{ij}(y)| = f_{ij}(x) = d(x, X \setminus P_{ij}(x)) \ge 2^i/\alpha \ge d(x, y)/(2\alpha) .$$

Note that events $\{\mathcal{E}_j\}_{j \in J}$ are mutually independent. Let Z_j be an indicator random variable for event $\mathcal{E}_j$, and set $Z = \sum_{j \in J} Z_j$. We have that $\mathbb{E}[Z] \ge |J|/8$, and by standard Chernoff bound

$$\Pr[Z \le |J|/16] \le e^{-|J|/128} \le 1/n^2 ,$$

when c is sufficiently large. If indeed $Z \geq |J|/16$ it follows that

$$\|f(x)-f(y)\|_p^p \geq \frac{1}{|J|}\sum_{j\in J}|f_{ij}(x)-f_{ij}(y)|^p \geq \left(\frac{d(x,y)}{32\alpha}\right)^p .$$

By applying a union bound over the $\binom{n}{2}$ pairs, we obtain that with probability at least $1/2$ we have an embedding with distortion $O(\alpha \cdot \log^{1/p}\Phi)$.

□

4 Measured Descent

In this section we enhance the embedding so that the dependence on the aspect ratio is replaced by a dependence on n, and also improve the dependence on the decomposability parameter α. This result was obtained by [KLMN04].

Theorem 3. *For any $1 \leq p \leq \infty$, any finite metric space (X, d) with n points has $c_p(X) = O(\alpha_X^{1-1/p} \cdot \log^{1/p} n)$.*

The distortion guarantee is tight for every possible value of α_X, as shown by [JLM09].

We will need the following lemma, whose proof is similar to that of Theorem 1, which is based on the random partitions of [FRT04, CKR01].

Lemma 2. *For any $\Delta > 0$, any finite metric space (X, d) admits a Δ-bounded ϵ-padded stochastic decomposition, where for each $x \in X$:*

$$\epsilon(x) = \frac{1}{16 + 16\ln\left(\frac{|B(x,\Delta)|}{|B(x,\Delta/8)|}\right)} .$$

Proof. Fix any $\Delta > 0$, and set $\epsilon : X \to [0, 1]$ as defined in the lemma. We now describe the random partition P. Let σ be a random permutation of X, and choose $r \in [\Delta/4, \Delta/2]$ uniformly at random. For each $u \in X$ define a cluster

$$C_u = \{x \in X \ : \ d(x,u) \leq r \text{ and } \sigma(u) < \sigma(v) \text{ for all } v \in X \text{ with } d(x,v) \leq r\} .$$

In words, every point in order of σ collects to its cluster all the unassigned points within distance r from it. Then $P = \{C_u\}_{u\in X} \setminus \{\emptyset\}$.

Fix some $x \in X$ and let $t = \epsilon(x) \cdot \Delta$, we need to show that the event $\{B(x, t) \not\subseteq P(x)\}$ happens with probability at most $1/2$. Let $a = |B(x, \Delta/8)|$ and $m = |B(x, \Delta)|$. Arrange the points $s_1, s_2, \ldots, s_m \in B(x, \Delta)$ in order of increasing distance from x. For $j \in [m]$, let I_j be the interval $[d(x, s_j) - t, d(x, s_j) + t]$. We say that the point s_j *cuts* $B(x, t)$ if s_j is the minimal element (of the permutation σ) for which $r \geq d(x, s_j) - t$, and also $r \in I_j$. Note that if $d(s_j, x) \leq \Delta/8$ then s_j cannot cut $B(x, t)$, because $d(x, s_j) + t \leq \Delta/8 + t < \Delta/4 \leq r$, so r cannot fall in the interval I_j. Also observe that if $u \notin B(x, \Delta)$ then $C_u \cap B(x, t) = \emptyset$ (for any choice of r).

$$
\begin{aligned}
\Pr[B(x,t) \not\subseteq P(x)] &\le \sum_{j=1}^{m} \Pr[s_j \text{ cuts } B(x,t)] \\
&= \sum_{j=a+1}^{m} \Pr[s_j \text{ cuts } B(x,t)] \\
&\le \sum_{j=a+1}^{m} \Pr[r \in I_j] \cdot \Pr[\forall_{i<j} \sigma(s_j) < \sigma(s_i) \mid r \in I_j] \\
&\le \sum_{j=a+1}^{m} \frac{2t}{\Delta/4} \cdot \frac{1}{j} \\
&\le \frac{8t}{\Delta} \cdot (1 + \ln(m/a)) \,.
\end{aligned}
$$

Plugging in the estimate for $t = \frac{\Delta}{16\left(1+\ln\left(\frac{|B(x,\Delta)|}{|B(x,\Delta/8)|}\right)\right)}$, gives a bound of $1/2$ on the probability, as required. □

We will use the following definition of local growth-rate. Intuitively, the embedding will have more coordinates in scales for which there is a significant local growth change, and few (even none) when there is little change in the local cardinality of balls. Fix any $r > 0$, and define the *local growth-rate* of x at scale r as

$$
\mathrm{GR}(x,r) = \frac{|B(x,2r)|}{|B(x,r/512)|} \,.
$$

Proof of Theorem 3. For any integer $k \in \mathbb{Z}$ let P_k be a 2^k-bounded random partition sampled from Lemma 2. Denote by ϵ_k the padding function of P_k. For each k and each $C \in P_k$ let $\tau(C)$ be a $\{0,1\}$ Bernoulli uniform random variable chosen independently. For each $x \in X$ and integer $t > 0$ let $k(x,t) = \max\{k \in \mathbb{Z} \,:\, |B(x,2^k)| < 2^t\}$. Let $I = \{-5,-4,-3,-2,-1,0,1,2,3,4,5\}$. For each $t \in T := \{0,1,\ldots,\lceil \log n \rceil\}$ and $i \in I$ define a set

$$
W_t^i = \{x \in X \,:\, \tau(P_{k(x,t)+i}(x)) = 0\} \,.
$$

The embedding $f : X \to \mathbb{R}^{|I|\cdot|T|}$ is defined as $f(x) = (d(x,W_t^i) \,:\, i \in I, t \in T)$. By the triangle inequality, every coordinate of f is non-expansive, so for any $x,y \in X$

$$
\|f(x)-f(y)\|_p^p \le |I| \cdot |T| \cdot d(x,y)^p = O(\log n) \cdot d(x,y)^p \,.
$$

It remains to show a bound on the contraction. Fix any $x,y \in X$, and let $R = d(x,y)$. It is not hard to verify that

$$
\max\left\{\frac{|B(x,2R)|}{|B(x,R/4)|}, \frac{|B(y,2R)|}{|B(y,R/4)|}\right\} \ge 2 \,. \tag{1}
$$

To see this, note that $B(x,R/4) \cap B(y,R/4) = \emptyset$, while both balls are contained in $B(x,2R)$ and also in $B(y,2R)$. Assume w.l.o.g that $\frac{|B(x,2R)|}{|B(x,R/4)|} \ge 2$. Let t_{lo}, t_{hi} be two integers such that $2^{t_{lo}-1} \le |B(x,R/512)| < 2^{t_{lo}}$ and $2^{t_{hi}} \le |B(x,2R)| < 2^{t_{hi}+1}$. Observe that $t_{hi} -$

$t_{lo} > \log \mathrm{GR}(x,R) - 2$, and due to (1) and our assumption on x, $\log \mathrm{GR}(x,R) \geq 1$, so that $t_{hi} - t_{lo} \geq 0$. Fix any integer $t \in [t_{lo}, t_{hi}]$, and let $k = k(x,t)$. Using the maximality of k in the definition of $k(t,x)$ we obtain that $|B(x, 2^{k+1})| \geq 2^t$, so that $2^k \geq R/1024$ (otherwise $|B(x,2^{k+1})| \leq |B(x,R/512)| < 2^t$). We also have that $2^k < 2R$ (otherwise $|B(x,2^k)| \geq |B(x,2R)| \geq 2^t$). Let $u \in I$ be such that

$$R/32 \leq 2^{k+u} < R/16 \; .$$

(It can be checked that such $u \in I$ exists.)

For any $z \in B(x, R/2048)$, we claim that

$$k - 1 \leq k(t,z) \leq k+2 \; . \tag{2}$$

To see this, note that $d(x,z) \leq R/2048 \leq 2^{k-1}$, and since $|B(x,2^{k+1})| \geq 2^t$ we conclude that $2^t \leq |B(z, 2^{k+1} + d(x,z))| \leq |B(z,2^{k+2})|$, or in other words that $k(t,z) \leq k+2$. Similarly, $2^t > |B(x,2^k)| \geq |B(z, 2^k - d(x,z))| \geq |B(z,2^{k-1})|$, so that $k(t,z) \geq k-1$.

Let $I' = \{-1,0,1,2\}$, and note that for $i \in I'$, by the assertion of Lemma 2,

$$\begin{aligned}
\epsilon_{k+u+i}(x) &= \frac{1}{16\left(1 + \ln\left(\frac{|B(x,2^{k+u+i})|}{|B(x,2^{k+u+i-3})|}\right)\right)} \\
&\geq \frac{1}{16\left(1 + \ln\left(\frac{|B(x,2^{k+u+2})|}{|B(x,2^{k+u-4})|}\right)\right)} \\
&\geq \frac{1}{16\left(1 + \ln\left(\frac{|B(x,2R)|}{|B(x,R/512)|}\right)\right)} \\
&= \frac{1}{16\left(1 + \ln \mathrm{GR}(x,R)\right)}
\end{aligned} \tag{3}$$

Let $\delta = \frac{1}{2048(1+\ln \mathrm{GR}(x,R))}$. Consider the set $W_t^u = \{z : \tau(P_{k(t,z)+u}(z)) = 0\}$. Let $\mathcal{E}_{big}$ be the event that $d(y, W_t^u) \geq \delta R/2$, and $\mathcal{E}_{small}$ be the event that $d(y, W_t^u) < \delta R/2$. Observe that these events are independent of the value of $\tau(P_{k+u+i}(x))$ for any $i \in I'$, because P_{k+u+i} is 2^{k+u+i}-bounded and $2^{k+u+i} < R/4$, thus for any $z \in B(y, \delta R/2)$, we have that $d(x,z) > 3R/4 \geq \mathrm{diam}(P_{k+u+i}(x))$ (note that $\mathcal{E}_{big}$ and $\mathcal{E}_{small}$ are indeed independent of values τ gives to points outside $B(y, \delta R/2)$).

If it is the case that $\mathcal{E}_{big}$ holds, then there is probability $1/2$ that $\tau(P_{k+u}(x)) = 0$ (independently as we noted above), in such a case $d(x, W_t^u) = 0$, and we obtain that

$$|d(x, W_t^u) - d(y, W_t^u)| \geq \delta R/2 \; .$$

The other case is that $\mathcal{E}_{small}$ holds. Let $\mathcal{E}$ be the event that for each $i \in I'$, $B(x, \epsilon_{k+u+i}(x) 2^{k+u+i}) \subseteq P_{k+u+i}(x)$ and $\tau(P_{k+u+i}(x)) = 1$. These events are clearly mutually independent, and since $2^{k+u+i} < R/4$ and P_{k+u+i} is 2^{k+u+i} bounded, they are also independent of $\mathcal{E}_{small}$. The probability of $\mathcal{E}$ is at least 2^{-8}. Consider any $z \in B(x, \delta R)$. If $\mathcal{E}$ indeed holds, then for each $i \in I'$: since $2^{k+u+i} \geq 2^{k+u-1} \geq R/64$ and due to (3) we conclude that $\epsilon_{k+u+i}(x) 2^{k+u+i} \geq \delta R$ so that $z \in P_{k+u+i}(x)$. From (2) we recall that $k(t,z) \in k + I'$, and as for any $i \in I'$, $\tau(P_{k+u+i}(x)) = 1$ (assuming event $\mathcal{E}$), it follows that $z \notin W_t^u$. We conclude that $d(x, W_t^u) \geq \delta R$, and as $\mathcal{E}_{small}$ holds:

$$|d(x, W_t^u) - d(y, W_t^u)| \geq \delta R/2 \; .$$

We conclude that for each of the (at least) $\min\{1, \log \mathrm{GR}(x, R) - 1\} \geq \log \mathrm{GR}(x, R)/2$ coordinates $t \in [t_{lo}, t_{hi}]$, with constant probability the contribution from a coordinate corresponding to t (and the appropriate value of u) is at least $\delta R/2$, and thus

$$\mathbb{E}[\|f(x) - f(y)\|_p^p] \geq \Omega\left(\frac{R}{\log \mathrm{GR}(x, R)}\right)^p \cdot \log \mathrm{GR}(x, R) \ .$$

Next, we devise another embedding $g : X \to \mathbb{R}$ using the same procedure, while sampling from the $1/\alpha_X$-padded distribution (guaranteed to exists as (X, d) is α_X-decomposable). The same proof holds (defining $\delta = 1/(128\alpha)$), and we obtain that

$$\mathbb{E}[\|f(x) - f(y)\|_p^p] \geq \Omega\left(\frac{R}{\alpha_X}\right)^p \cdot \log \mathrm{GR}(x, R) \ .$$

Finally, observe that choosing at random between f and g, we obtain in expectation the summation of these estimates (divided by 2), which is at least $\Omega(d(x, y)^p/\alpha_X^{p-1})$. This concludes the proof. $\square$

Remark: In order to achieve an actual bound, not only on the expectation, one can use standard sampling and Chernoff bound as in the proof of Theorem 2, and obtain an embedding into R^D with $D = O(\log^2 n)$.

References

[Bar96] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*, pages 184–, Washington, DC, USA, 1996. IEEE Computer Society.

[CKR01] Gruia Calinescu, Howard Karloff, and Yuval Rabani. Approximation algorithms for the 0-extension problem. In *Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, SODA '01, pages 8–16, Philadelphia, PA, USA, 2001. Society for Industrial and Applied Mathematics.

[FRT04] Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. Approximating metrics by tree metrics. *SIGACT News*, 35(2):60–70, 2004.

[GKL03] Anupam Gupta, Robert Krauthgamer, and James R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '03, pages 534–, Washington, DC, USA, 2003. IEEE Computer Society.

[JLM09] Alex Jaffe, James R. Lee, and Mohammad Moharrami. On the optimality of gluing over scales. In *Proceedings of the 12th International Workshop and 13th International Workshop on Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, APPROX '09 / RANDOM '09, pages 190–201, Berlin, Heidelberg, 2009. Springer-Verlag.

[KLMN04] Robert Krauthgamer, James R. Lee, Manor Mendel, and Assaf Naor. Measured descent: A new embedding method for finite metrics. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 434–443, Washington, DC, USA, 2004. IEEE Computer Society.

[Rao99] Satish Rao. Small distortion and volume preserving embeddings for planar and euclidean metrics. In *Proceedings of the fifteenth annual symposium on Computational geometry*, SCG '99, pages 300–306, New York, NY, USA, 1999. ACM.

Compression and Random Walks

following Austin, Naor & Peres

written[1] by Antoine Gournay

The aim of this note is to present the results from Austin, Naor & Peres [4] and Naor & Peres [18]. Namely, we want to show a quantitative upper bound on the compression function of an infinite (finitely generated) group into a Banach space. There are two cases: a bound for any embedding in the case of amenable groups and a bound for equivariant embeddings. These bound are expressed in terms of a property of random walks (the speed, *i.e.* the growth of the expected distance to the identity) and a property of the Banach space (Markov type p in the case of amenable groups and modulus of smoothness of power type p in the equivariant case).

Throughout the text, we will restrict to finitely generated groups.

1 Basic definitions and main results

1.1 Random walks

Let us start with random walk on graphs. The simple random walk on a graph G is a sequence[2] $\{W_n\}_{n=0}^{\infty}$ of random variables taking value in G defined as follows: W_0 is the Dirac mass at the identity element[3] and

$$\mathbb{P}(W_{n+1} = y \mid W_n = x) = \frac{k}{\deg(x)} \qquad \text{if there are } k \text{ edges from } x \text{ to } y,$$

where the degree of x is the number of edges[4] incident at x. Hence, one can compute inductively the law of the W_i. The important point about formulating this inductively is that these random variables are dependant.

We will almost exclusively look at random walks on Γ, a finitely generated group. To do so one constructs its Cayley graph for a generating set S which is finite and symmetric ($s \in S \implies s^{-1} \in S$). The inductive procedure is then written as $\mathbb{P}(W_{n+1} = \eta | W_n = \gamma) = 1/|S|$ if there is a $s \in S$ such that $\eta = \gamma s$ and is $= 0$ otherwise. This data which allows to deduce the $(n+1)^{\text{th}}$ variable from the n^{th} is called the kernel of the Markov process. Here, $K(\eta, \gamma) = \mathbb{1}_S(\eta^{-1}\gamma)/|S|$ where $\mathbb{1}_S$ is the characteristic function of S.

One could also describe the law of W_n as follows:

- (push-forward) Pick uniformly at random n elements in S and multiply them. More formally, look at the map $M_n : S^\infty \to \Gamma$ given by $M_n(s_1, s_2, \ldots, s_n) = s_1 s_2 \cdots s_n \in \Gamma$. Take the uniform measure u on S and the corresponding product measure $\chi = u^{\mathbb{N}}$. The law of W_n is the push-forward by M_n of the latter, *i.e.* for $A \subset \Gamma$, $W_n(A) = \chi(M_n^{-1}(A))$.

[1] Please email comments / corrections / improvements / references / insults / etc... to first-name.name@gmail.com

[2] In fact, a Markov process.

[3] Sometimes, the "initial data" W_0 is some probability distribution rather than a Dirac mass.

[4] Some people count loops as 2 edges and some as 1. It does not matter, as long as this factor is integrated in the degree.

- (Convolution) Given $f, g : \Gamma \to \mathbb{R}$, one defines their convolution as $f*g(\gamma) = \sum_{\eta\in\Gamma} f(\eta)g(\eta^{-1}\gamma)$ (one needs to assume the sum is convergent for any γ). Let $P = \frac{1}{|S|}\mathbb{1}_S$, then the law of W_n is given by $P * P * \cdots * P$ (where P appears n times).

Example 1.1. The simplest example is the random walk on the group $\mathbb{Z}$ with $S = \{\pm 1\}$. Since the walker either moves to the left or the right, one can easily see that the law will be, up to cosmetic differences, a Binomial distribution with n trials and $p = 1/2$. More precisely, if B_n is a random variable with n-trials Binomial law, then

$$\mathbb{P}(W_{2n+1} = k) = \begin{cases} \mathbb{P}(B_{2n+1} = k + \frac{2n+1-k}{2}) & \text{if } k \text{ is odd,} \\ 0 & \text{otherwise;} \end{cases}$$
$$\mathbb{P}(W_{2n} = k) = \begin{cases} \mathbb{P}(B_{2n} = k + \frac{2n-k}{2}) & \text{if } k \text{ is even,} \\ 0 & \text{otherwise.} \end{cases}$$

Note that if one is only interested in "rough" properties of this walk, many of these properties can be easily computed *via* the approximation of the binomial laz by the normal law: in this case we get an approximation by $\mathcal{N}(0, n/4)$ (the normal law with mean 0 and variance $n/4$). ♣

Short history

Polyá's theorem is usually seen as the one of the first result (for infinite groups!). It states that a random walker walking for an infinite time in $\mathbb{Z}^d$ will, with probability 1, visit the origin infinitely many times if and only if $d \leq 2$. It will "go to infinity" with probability 1 if $d \geq 3$. This theorem relies on a good knowledge for the asymptotics of the function $f(n) := \mathbb{P}(W_{2n} = e_\Gamma)$ as $n \to \infty$ (namely $f(n) \asymp Kn^{-d/2}$).

These probability of return came up again in the work of Kesten [15]. He showed that a group is not amenable [see definition below] if and only if $\lim f(n)^{1/2n} < 1$ (where, again, $f(n) = \mathbb{P}(W_{2n} = e_\Gamma)$; the limit exists [exercise]).

On the other hand, bounded harmonic functions on Cayley graphs and random walks are closely related: Avez, Choquet & Deny, Derrienic, Kaimanovich & Vershik, ... This gives a strong link with an ideal completion of the Cayley graph (the Poisson boundary) and some properties of the random walk (speed and entropy). Namely, if the random walker does not flee fast enough to infinity then there are no bounded harmonic functions except the constant function.

Speed

The quantity related to random walks which is of interest here is the speed (also called drift). This is a measure of the expected distance after n-steps to the starting point. A first thing to check (exercise!) is

$$\mathbb{E}\Big(d_S(W_{n+m}, e_\Gamma)\Big) \leq \mathbb{E}\Big(d(W_m, e_\Gamma)\Big) + \mathbb{E}\Big(d_S(W_n, e_\Gamma)\Big),$$

where d_S is the distance in the Cayley graph.

Definition 1.2. Given a group Γ and a generating set S, the [lower] **speed exponent** is

$$\begin{aligned} \beta_{\Gamma,S} &= \sup\{c \in [0,1] \mid \exists K > 0 \text{ such that } \mathbb{E}\Big(d(W_n, e_\Gamma) \geq Kn^c\} \\ &= \liminf_{n\to\infty} \frac{\log \mathbb{E}\Big(d(W_n, e_\Gamma)\Big)}{\log n}. \end{aligned}$$ ★

The speed exponent is tricky to compute. Although it is relatively easy to check that $\beta = 1/2$ for Abelian groups, it was not known until 2005 (an argument of Virág, see Lee & Peres [16]) that $\beta \geq 1/2$ for any infinite group. [Abelian groups are somehow the "smallest" infinite groups.]

It is unknown whether β depends on S (for a fixed Γ). It even unknown if it depends on quasi-isometries (between Cayley graphs!). Nevertheless, Γ and S will often be implicit from the context.

Example 1.3. Let us go back quickly to our example with $\mathbb{Z}$. After n steps, the distribution is well approximated by a centred normal distribution with variance $n/4$. For example, writing N for a random variable with law the standard centred Gaussian,

$$\mathbb{P}(|W_n| \leq A) \simeq \mathbb{P}(|\sqrt{n}N/2| \leq A) = \int_{-2A/\sqrt{n}}^{2A/\sqrt{n}} e^{-x^2/2} dx \simeq 4A/\sqrt{n}.$$

where the $\simeq$ should be read as lower and upper bounds up to constants (near 0, $e^{-x^2/2} \simeq 1$). Similarly, if $F : \mathbb{N} \to \mathbb{R}_{>0}$ is any increasing function with $\lim_{n\to\infty} F(n) = +\infty$,

$$\mathbb{P}\Big(|W_n| \leq \sqrt{n}F(n)\Big) = \int_{-2F(n)}^{2F(n)} e^{-x^2/2} dx \overset{n\to\infty}{\longrightarrow} 1$$

This implies that $\beta_{\mathbb{Z},\{\pm 1\}} \leq 1/2$. On the other hand,

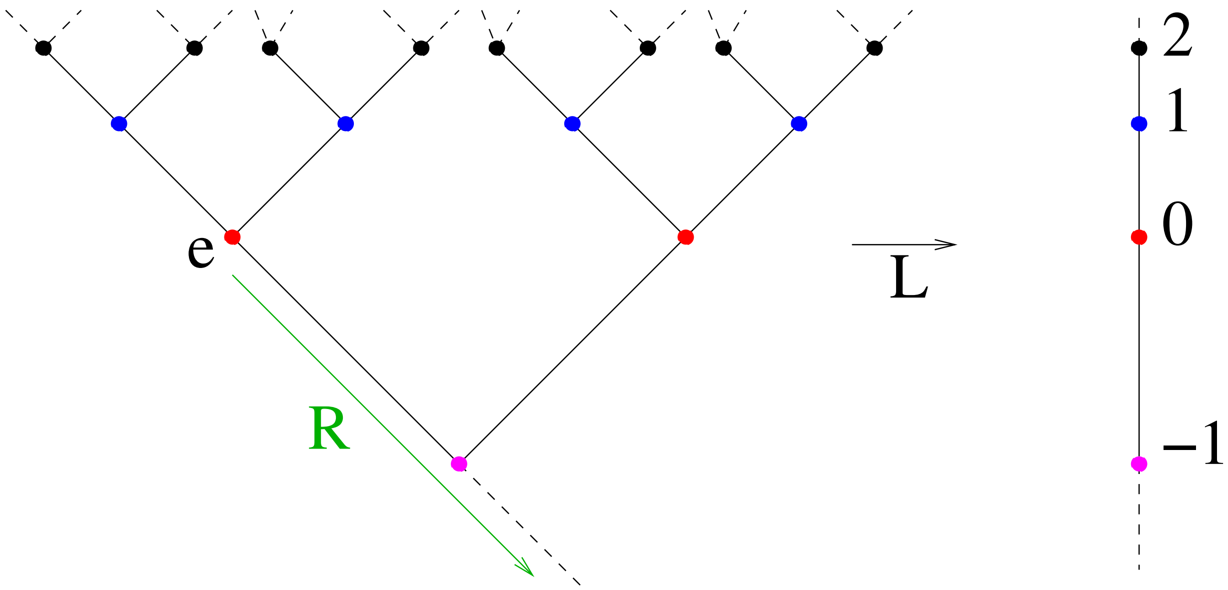
$$\mathbb{P}\Big(|W_n| \geq \sqrt{n}/F(n)\Big) = \int_{-\infty}^{-2/F(n)} + \int_{2/F(n)}^{\infty} e^{-x^2/2} dx \overset{n\to\infty}{\longrightarrow} 1.$$

This also shows $\beta_{\mathbb{Z},\{\pm 1\}} \geq 1/2$. ♣

Knowing the speed for one group implies bound on speed for other groups, see §4.

Let us do an example with high speed.

Example 1.4. Let us show that if the Cayley graph of Γ is a tree T of valency $v \geq 3$ (*e.g.* Γ is a free group on at least two generators, *e.g.* $\Gamma = \mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$) then the speed exponent is 1 (in fact, the expectation grows linearly). To do this pick an infinite [geodesic] path R from e to somewhere at infinity. Define the level of a vertex v in the tree by $L(v) = 2d(v,R) - d(v,e)$.



Note that $L(v) \leq d(v,e)$. L also defines a map $L : T \to \mathbb{Z}$. Let $Z_t = L(W_t)$. Note that Z_t is going to be a random walk on $\mathbb{Z}$ but with a preference for a direction:

$$\mathbb{P}(Z_{t+1} = k+1 \mid Z_t = k) = \frac{v-1}{v} \qquad \text{and} \qquad \mathbb{P}(Z_{t+1} = k-1 \mid Z_t = k) = \frac{1}{v}.$$

As in the previous example, the law of Z_t will be (essentially) a binomial law (with probability of success $\frac{v-1}{v}$). The normal approximation will be $\mathcal{N}(n\frac{v-2}{v}, n\frac{v(v-1)}{v^2})$. Using this[5], it is very easy to show that,

$$\mathbb{P}((1-\epsilon)n\tfrac{v-2}{v} \leq Z_t \leq (1+\epsilon)n\tfrac{v-2}{v}) \to 1.$$

Since $Z_n \leq d(W_n, e)$, one concludes that

$$\mathbb{E}\Big(d(W_n,e)\Big) \geq n(1-\epsilon)\tfrac{v-2}{v}.$$

This implies $\beta \geq 1$ (and so $\beta = 1$, since the linear upper bound is trivially true). ♣

1.2 Amenable groups

For a set $F \subset \Gamma$ the boundary ∂F is the set of edges between F and F^{c}.

Definition 1.5. Assume Γ is finitely generated. A sequence $\{F_n\}$ of subsets of Γ is **Følner** if and only if the F_n are finite and $\lim\limits_{n\to\infty} \dfrac{|\partial F_n|}{|F_n|} = 0$. A group is said to be **amenable** if it has a Følner sequence. ★

Alternatively, a group is not amenable if there exists a $K > 0$ such that, for any finite set F,

$$\frac{|\partial F|}{|F|} > C.$$

This is also known as a "strong" isoperimetric profile (or having a positive isoperimetric constant).

Example 1.6. Here are example of amenable groups:

- finite groups;
- Abelian, nilpotent, polycyclic and solvable groups;
- if B_n is the ball of radius n (around e) in the Cayley graph and $\liminf \frac{1}{n}\log|B_n| = 0$ then the group is amenable (exercise!). [Such groups are called of "subexponential" growth.]

Here are examples of non-amenable groups:

- free groups on at least 2 generators;
- hyperbolic groups (non-elementary ones, *i.e.* except virtually-$\mathbb{Z}$ groups[6]);
- some infinite torsion groups ("Burnside groups"). ♣

Given a few amenable groups there are many ways to build new ones:

[5] If you are into probability, it's probably more natural for you to use the Hoeffding inequality [or Chernoff, or Azuma-Hoeffding, or ...] here.

[6] If P is a property of groups (*e.g.* being $\mathbb{Z}$, being Abelian, being nilpotent, ...), a group group is said to be virtually-P is it contains a subgroup of finite index which has the property P.

Theorem 1.7 (*"The closure properties"*)

Let Γ, N and $\{\Gamma_i\}_{i\geq 0}$ be amenable groups.

(a) *If H is a subgroup[7] of Γ then H is amenable* *"Subgroup"*

(b) *If H is an extension N by Γ (*i.e.* $1 \to N \to H \to \Gamma \to 1$ is an exact sequence) then H is amenable* *"Extension"*

(c) *If $N \lhd \Gamma$ then $H = \Gamma/N$ is amenable* *"Quotient"*

(d) *If H is a direct limit of the Γ_i then H is amenable* *"Direct limit"*

Note that any group containing a non-amenable group is non-amenable (by (a)). The previous properties (with the eventual exception of (d)) were first shown in von Neumann [23] (together with the first definition of amenability).

1.3 The Main results

To fix notations let's recall the definition of compression exponent:

Definition 1.8. Let B be a Banach space. A coarse embedding $f : \Gamma \to B$ is a map such that there exist an unbounded increasing function $\rho_f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and a constant $C > 0$, satisfying $\forall x, y \in \Gamma$

$$\rho_f(d(x,y)|) \leq \|f(x) - f(y)\| \leq Cd(x,y).$$

★

A very important note to make before going on is that the right-hand side is of a very particular form. The reason is the following: in a graph metric, the first non-trivial value of $d(x,y)$ is 1. Hence, if the modulus of distortion[8] ω_f is finite, then the triangle inequality (used along a path from x to y) implies $\|f(x) - f(y)\| \leq d(x,y) \cdot \omega_f(1)$. This is an instance of the Colson-Klee lemma.

Definition 1.9. The embedding is said to be **equivariant** if there is a representation $\pi : \Gamma \to \mathrm{Isom} B$ of Γ in the isometries[9] of B and $f(\gamma x) = \pi(\gamma) f(x)$.

The function $\rho_f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is called the **modulus of compression** (associated to f). The bf compression exponent is

$$\alpha(f) = \sup\{c \in [0,1] \mid \exists K > 0 \text{ such that } \rho_f(n) \geq K n^c\} = \liminf \frac{\log \rho_f(t)}{\log t}.$$

The **compression exponent** of Γ, $\alpha_B(\Gamma)$, is the supremum over all embeddings into B of $\alpha(f)$. The **equivariant compression exponent** of Γ, $\alpha^\sharp_B(\Gamma)$ is the supremum over all equivariant $f : \Gamma \to B$ of $\alpha(f)$. ★

Markov type will be defined soon (in §2.1). For now, just note that L_p has Markov type $\min(2,p)$. [The following theorem can be stated for any metric space, not just Banach space.]

Theorem 1.10

Let Γ be an amenable group and let $\widehat{\beta} = \sup_S \beta_{\Gamma,S}$. If B has Markov type p, then

[8] ω_f is the modulus of distortion of the map f above if, for any $x, y \in \Gamma$, $\|f(x) - f(y)\| \leq \omega_f(d(x,y))$

[9] Here "isometries" stands for surjective isometries. Indeed, in order for this to a representation (*i.e.* a group homomorphism) the target needs to form a group.

$p\alpha_B\widehat{\beta} \leq 1$.

The following theorem restricts to equivariant compression but holds for any (infinite finitely generated) group.

Theorem 1.11

Let $\widehat{\beta} = \sup_S \beta_{\Gamma,S}$. If B has a modulus of smoothness with exponent of power-type p, then $p\alpha^\sharp_B\widehat{\beta} \leq 1$.

Theorem 1.11 is a strengthening of Guentner & Kaminker [14, Theorem 5.3]. Indeed, if $\alpha^\sharp_{L_2} > \frac{1}{2}$ then $\beta < 1$. This implies the "Liouville property" which in turns implies amenability.

2 Proof of Theorem 1.10

2.1 Main ingredient: Markov type p

A Markov chain on Y is a sequence of random variables $\{Z_n\}_{n=0}^\infty$ (with possible values in the state space Y) such that

$$\mathbb{P}(Z_{n+1} = y \mid Z_n = y_n, Z_{n-1} = y_{n-1}, \dots, Z_0 = y_0) = \mathbb{P}(Z_{n+1} = y \mid Z_n = y_n).$$

One usually see Z_n as a random variable which evolves in time. The condition means that the process has no memory (and its evolution is time independent): only the current state determines the (possible) future evolution. Recall that the kernel is defined by $K(x,y) = \mathbb{P}(Z_{n+1} = x \mid Z_n = y)$ (does not depend on n!).

Definition 2.1. A Markov chain on a finite state space (Y is finite) is **stationary** if $\pi(y) := \mathbb{P}(Z_n = y)$ does not depend on n. It is **reversible** if $\pi(x)K(x,y) = \pi(y)K(y,x)$. ★

Example 2.2. Take the kernel of the simple random walk on a finite graph. Define the initial distribution (*i.e.* the law of W_0) to be

$$\mathbb{P}(W_0 = x) = \pi(x) = \frac{\deg x}{2|E|}$$

where $|E|$ is the number of edges[10]. The claim is that this is a stationary and reversible Markov chain. First check it is reversible: if there are k_{xy} edges between x and y,

$$\pi(x)K(x,y) = \frac{k}{2|E|} = \pi(y)K(y,x).$$

Then check it is stationary:

$$\begin{aligned} \mathbb{P}(Z_1 = y) &= \sum_x \mathbb{P}(Z_1 = y \mid Z_0 = x)\mathbb{P}(Z_0 = x) &&= \sum_x \frac{k_{xy}}{\deg x}\pi(x) \\ &= \sum_x \frac{k_{xy}}{\deg x}\frac{\deg x}{2|E|} &&= \frac{1}{2|E|}\sum_x k_{xy} \quad = \frac{\deg y}{2|E|} \end{aligned}$$

This shows $\mathbb{P}(Z_1 = y = \pi(y)$ so that Z_1 has the same law as Z_0. By a trivial induction, Z_n all have the same law. ♣

[10] Because of loops, there might some problem of convention related to how you count the edges. If loops contribute 2 to the degree, then a loop count as one edge. If loops contribute to 1, then a loop count as 1/2 an edge.

The notion of Markov type has been used to attack (successfully) many embedding problems of finite metric spaces. The main idea in Austin, Naor & Peres [4] is to use it on Følner sequences (which are the finite sets giving a "good" approximation of the infinite group). The notion was introduced by K. Ball in [5].[11] Markov type 2 implies [Rademacher or usual] type 2.

Definition 2.3. A metric space (X, d_X) has **Markov type** p ($p \in [1,2]$), if for every stationary Markov chain $\{Z_n\}_{n=0}^{\infty}$ on Y (finite!) and every mapping $f : Y \to X$, one has

$$\mathbb{E}\bigg(d_X\Big(f(Z_n), f(Z_0)\Big)^p\bigg) \leq K^p n \mathbb{E}\bigg(d_X\Big(f(Z_1), f(Z_0)\Big)^p\bigg) \qquad \bigstar$$

A rough interpretation (for $p = 2$) is that the distance grows in expectation as $\sqrt{t}$ times the size of the first step, so that some "central limit" behaviour holds. In K. Ball's own words: "This property was introduced as a nonlinear analogue of the classical type property for normed spaces that arose in the theory of vector-valued central limit theorems and the extension/factorisation theory of Kwapien and Maurey."

K. Ball showed L_p has Markov type p for $p \leq 2$ and it was shown by Naor, Peres, Schramm and Sheffield [20] that Banach space with modulus of smoothness of power-type 2 (*e.g.* L_p for $p > 2$) have Markov type 2. Negatively curved Riemannian manifolds and δ-hyperbolic spaces also have Markov type 2 (see again [20]).

Example 2.4. Let us show that $\mathbb{R}$ has Markov type 2. Let K be the kernel of the Markov chain. It acts on functions $f : Y \to \mathbb{R}$ by convolution:

$$Kf(x) = \sum_{y \in Y} K(x,y) f(y).$$

It turns out that the reversibility of the Markov chain implies that K is self-adjoint in $L_2(Y, \pi)$:

$$\begin{array}{ll} \langle Kf \mid g\rangle = \sum_x Kf(x)g(x)\pi(x) & = \sum_x \Big(\sum_y K(x,y)f(y)\Big)g(x)\pi(x) \\ = \sum_{x,y} \pi(x)K(x,y)f(y)g(x) & \overset{\text{rev.}}{=} \sum_{x,y} \pi(y)K(y,x)f(y)g(x) \\ = \sum_y \Big(f(y)\pi(y)\sum_x K(y,x)g(x)\Big) & = \sum_y f(y)Kg(y) \\ = \langle f \mid Kg\rangle & \end{array}$$

The reverse ("self-adjoint" implies "reversibility") is also true: just let f and g vary over all possible Dirac masses. An important upshot for us is that (as an operator) it has real eigenvalues. Furthermore, it has norm $||K||_{\ell^2 \to \ell^2} \leq 1$ (since K is convolution by a kernel which sums to 1, use Young's inequality). If you don't know Young's inequality, just note that

$$|Kf(x)|^2 = |\sum_y K(x,y)f(y)|^2 \overset{C-S}{\leq} \Big(\sum_y |f(y)|^2 K(x,y)\Big)\Big(\sum_y K(x,y)\Big) = \sum_y |f(y)|^2 K(x,y).$$

where C-S is the Cauchy-Schwartz inequality. This implies

$$||Kf||_{\ell^2}^2 \sum_x |Kf(x)|^2 \leq \sum_x \sum_y |f(y)|^2 K(x,y) = \sum_y \Big(|f(y)|^2 \sum_x K(x,y)\Big) = ||f||_{\ell^2}^2.$$

[11] In this paper he shows that given X a metric space of Markov type 2 and Y a metric space of Markov cotype 2 then any Lipschitz maps $f : Z \to Y$ (where $Z \subset X$) extends to a Lipschitz map $\tilde{f} : X \to Y$.

Let K^t be the kernel for the t-step Markov chain (the same Markov chain, but taking t steps at a time). One still has $\sum_y K^t(x,y) = 1$ (because one must get somewhere in t steps). As a operator, it is the same thing as applying the operator K t times. In particular, K^t is also self-adjoint (so also reversible).

If you rather want to see "directly" that K^t is reversible (for the same π), note that

$$K^t(x,y) = \sum_{x \sim x_1 \sim x_2 \ldots \sim x_{t-1} \sim y} K(x,x_1)K(x_1,x_2)\ldots K(x_{t-1},y).$$

where the sum is over all path of length t between x and y. If one multiplies this by $\pi(x)$, one transform slowly the sum:

$$\begin{aligned}
\pi(x)K^t(x,y) &= \textstyle\sum_{x \sim x_1 \sim x_2 \ldots \sim x_{t-1} \sim y} \pi(x)K(x,x_1)K(x_1,x_2)\ldots K(x_{t-1},y) \\
&= \textstyle\sum_{x \sim x_1 \sim x_2 \ldots \sim x_{t-1} \sim y} K(x_1,x)\pi(x_1)K(x_1,x_2)\ldots K(x_{t-1},y) \\
&\ \ \vdots \\
&= \textstyle\sum_{x \sim x_1 \sim x_2 \ldots \sim x_{t-1} \sim y} K(x_1,x)K(x_2,x_1)\ldots K(y,x_{t-1})\pi(y) \\
&= \textstyle\sum_{x \sim x_1 \sim x_2 \ldots \sim x_{t-1} \sim y} K(y,x_{t-1})\ldots K(x_2,x_1)K(x_1,x)\pi(y) \\
&= \pi(y)K^t(y,x)
\end{aligned}$$

Now let us do some massaging:

$$\begin{aligned}
&\mathbb{E}\Big(d(f(Z_t),f(Z_0))^2\Big) \\
&\quad= \mathbb{E}\Big[\Big(f(Z_t)-f(Z_0)\Big)^2\Big] \\
&\quad= \sum_x \mathbb{E}\Big[\Big(f(Z_t)-f(Z_0)\Big)^2 \mid Z_0 = x\Big]\mathbb{P}(Z_0 = x) \\
&\quad= \sum_x \Big[\sum_y \mathbb{P}(Z_t = y \mid Z_0 = x)\Big(f(y)-f(x)\Big)^2\Big]\pi(x) \\
&\quad= \sum_{x,y} \pi(x)K^t(x,y)\Big(f(y)-f(x)\Big)^2 \\
&\quad= \sum_{x,y} \pi(x)K^t(x,y)f(y)^2 + \sum_{x,y} \pi(x)K^t(x,y)f(x)^2 - 2\sum_{x,y} \pi(x)K^t(x,y)f(y)f(x) \\
&\quad\overset{\text{rev.}}{=} \sum_{x,y} \pi(y)K^t(y,x)f(y)^2 + \sum_{x,y} \pi(x)K^t(x,y)f(x)^2 - 2\sum_{x,y} \pi(x)K^t(x,y)f(y)f(x) \\
&\quad= 2\sum_{x,y} \pi(x)K^t(x,y)f(x)^2 - 2\sum_{x,y} \pi(x)K^t(x,y)f(y)f(x) \\
&\quad= 2\sum_x \pi(x)\Big(\sum_y K^t(x,y)\Big)f(x)^2 - 2\sum_{x,y} \pi(x)K^t(x,y)f(y)f(x) \\
&\quad= 2\sum_x \pi(x)f(x)^2 - 2\sum_{x,y} \pi(x)K^t(x,y)f(y)f(x) \\
&\quad= 2\langle(\mathrm{Id}-K^t)f,f\rangle
\end{aligned}$$

Hence one needs to prove that $\langle(\mathrm{Id}-K^t)f,f\rangle \leq t\langle(\mathrm{Id}-K)f,f\rangle$. If f is an eigenfunction (of eigenvalue λ) this reads $(1-\lambda^t) \leq t(1-\lambda)$ or

$$\sum_{i=0}^{t-1} \lambda^i \leq t.$$

This is true since $|\lambda| \leq 1$. Decomposing a generic f as a sum of eigenfunctions concludes the proof. ♣

A similar argument can be used to show L_2 has Markov type 2, and was first given by K. Ball. He then used this (together with existence of isometric embeddings of L_2 in L_p) to show that L_p has Markov type p for $p \leq 2$. For more see Ball's original paper [5], Lyons with Peres [17, §13.5, Theorem 13.16] or Naor, Peres, Schramm & Sheffield [20].

2.2 The proof

The "story" goes as follows. Assume the target space X has Markov type p and there is an embedding $\Gamma \to X$ which has a compression function with $\rho_f(t) \asymp t^\alpha$. How fast should a random walker go in the image? Well, since this guy has speed exponent β, one "expects" at time t some significant mass at distance $\rho_f(t^\beta)$. On the other hand, the Markov property says expected distance can only grow so fast; one can hope that a non-trivial bound pops up. There are some things to fix, because the Markov type is only for finite spaces.

PROOF OF THEOREM 1.10: Take a coarse embedding $f : \Gamma \to B$ with $\rho_f(t) \geq Kt^\alpha$ for some $K > 0$ and $\alpha \in]\alpha_B(\Gamma) - \epsilon, \alpha_B(\Gamma)[$. Next, take a $\beta \geq \beta_{\Gamma,S} - \epsilon$ (for some $\epsilon > 0$). For some t_0 and any $t > t_0$ we will obtain a bound on the compression function. Let F_n be a Følner sequence. Let $A_n = \cup_{x \in F_n} B(x,t)$ where $B(x,t)$ is the ball of radius t centred at x. It is an exercise to check that the Følner condition implies

$$\lim_{n\to\infty} \frac{|A_n|}{|F_n|} = 1.$$

Consider Z_t to be he random walk restricted to A_n with initial measure the uniform distribution on A_n. By random walk restricted to A_n, we mean that the kernel is exactly as before, except that if the element were to leave A_n from some vertex x, it remains at x (instead of leaving). If you wish, just replace every edge going out of A_n by a loop. More precisely,

$$\mathbb{P}(Z_{i+1} = y \mid Z_i = x) = \begin{cases} 1/|S| & \text{if } y = xs \text{ for some } s \in S \text{ and both } x, y \in A_n; \\ |xS \cap A_n| & \text{if } y = x \in A_n; \\ 0 & \text{otherwise.} \end{cases}$$

Instead of going through the computation to check that this is a reversible and stationary Markov chain, just note that this is the same situation as in Example 2.2. Indeed, since all vertices have the same degree, the initial distribution is the uniform distribution.

We will use the following inequalities: (K is always some constant > 0)

(MTp) Markov type p: $\mathbb{E}\Big(d_X\big(f(Z_t), f(Z_0)\big)^p\Big) \leq Kt\mathbb{E}\Big(d_X\big(f(Z_1), f(Z_0)\big)^p\Big)$ (MTp)

(1-L) The embedding is 1-Lipschitz: $\|f(x) - f(e)\| \leq Kd(x,e)$ (1-L)

(ρ_f) The compression function lower bound: $\|f(x) - f(e)\| \geq Kd(x,e)^{\alpha_0}$ (ρ_f)

(Fol) The Følner condition: $\lim_{n\to\infty} \frac{|A_n|}{|F_n|} = 1$ (Fol)

First use the Markov property to see one cannot get too far:

$$\mathbb{E}\Big(d_B\big(f(Z_t), f(Z_0)\big)^p\Big) \overset{(\text{MT}p)}{\leq} K^p t\mathbb{E}\Big(d_B\big(f(Z_1), f(Z_0)\big)^p\Big) \overset{(\text{1-L})}{\leq} K^p t\mathbb{E}\Big(d\big(Z_1, Z_0\big)^p\Big) \leq K^p t$$

where the last inequality follows since the random walk always makes one or no step: $\big(Z_1, Z_0\big) = 0$ or 1. Next, working in the other direction:

$$K^p t \geq \mathbb{E}\Big(d_B\big(f(Z_t), f(Z_0)\big)^p\Big) \overset{(\rho_f)}{\geq} \mathbb{E}\Big(\rho_f\big(d(Z_t, Z_0)\big)^p\Big) \geq \frac{1}{|A_n|}\sum_{x\in F_n} \mathbb{E}\Big(\rho_f\big(d(Z_t, Z_0)\big)^p \mid Z_0 = x\Big)$$

where the last inequality is obtained by noting that the terms removed in the sum (corresponding to $x \in A_n \setminus F_n$) are ≥ 0. Pick $\beta_0 \in]\beta - \epsilon, \beta[$. Note[12] that there is a $\delta > 0$ such that,

[12] The short way out is to note that there is nothing to prove if $\alpha p \leq 1$ so that one may use Jensen's inequality and (ρ_f) to interchange expectations and $^{p\alpha_0}$. But one can still keep the compression function along for a while.

for any $t > t_0$, $\mathbb{P}\Big(d(W_t, e_\Gamma) > t^{\beta_0}\Big) > \delta$ (otherwise the speed exponent would be less than β_0). If $Z_0 = x \in A_n$, the law of Z_t is exactly the same as the law of W_t (because A_n contains the t-ball around x). This[13] implies:

$$\mathbb{E}\left(\rho_f\Big(d(Z_t, Z_0)\Big)^p \mid Z_0 = x\right) \geq \delta \rho_f(t^{\beta_0})^p.$$

Hence,

$$K^p t \geq \frac{|F_n|}{|A_n|}\delta\rho_f(t^{\beta_0})^p \overset{(Fol)}{\to} \delta\rho_f(t^{\beta_0})^p$$

For any $\alpha_0 \in]\alpha - \epsilon, \alpha[$, there exists[14] $C > 0$ such that $\rho_f(t^{\beta_0})^p \overset{(\rho_f)}{\geq} Ct^{\alpha_0\beta_0 p}$. One gets that:

$$Kt \geq K' t^{\alpha_0\beta_0 p}$$

Remembering the ϵ lost on the way:

$$(\alpha_B(\Gamma) - 2\epsilon)(\widehat{\beta} - 2\epsilon)p \leq 1.$$

Taking $\epsilon \to 0$, yields the conclusion[15]. ■

2.3 Comments and questions

The theorem 1.10 is sharp for many amenable groups. In fact, to compute the exponent α this is a very useful upper bound. The lower bounds can be obtained by explicit coarse embeddings. See Austin, Naor & Peres [4] for a proof that, for coarse embeddings in Hilbert spaces (they are all isomorphic), $\alpha_{L_2}(\mathbb{Z} \wr \mathbb{Z}) = 2/3$. Further computations in wreath products are done by Naor & Peres in [18] and [19].

The bound from Theorem 1.10 is also very useful as a bound on speed rather than a bound on compression. Indeed, to show that $\beta \leq \frac{1}{2}$ (and hence $= \frac{1}{2}$, thanks to the generic lower bound of Virág, see [16]) it suffices to show that $\alpha_{L_2} \geq 1$. It is often easier to produce a good coarse embedding than to evaluate β by brute force. For more along these lines see [13].

There are amenable groups with $\alpha_{L_p} = 0$, see Austin [3] for a first construction (a solvable group) and Bartholdi & Erschler [7] for more (groups of intermediate growth).

Compression, when restricted to amenable groups has many other nice features. First, there is "Gromov's trick" which says that if $f : \Gamma \to \mathcal{H}$ is a coarse embedding in a Hilbert space $\mathcal{H}$, then there is an equivariant coarse embedding $g : \Gamma \to \mathcal{H}'$ such that the function ρ_f for g is the same as the one for f.

In Naor & Peres [19, Theorem 9.1] this is done in the non-Hilbertian setting. Namely fix a $p \in [1, \infty[$. If X is a Banach space and $f : \Gamma \to X$ is a coarse embedding then there is a Banach space Y which is finitely representable[16] in ℓ^p and with $\alpha^\sharp_Y(\Gamma) \geq \alpha^*_X(\Gamma)$. If, furthermore $X = L_p$ then Y may also be taken to be L_p. Hence $\alpha^\sharp_{L_p} = \alpha_{L_p}$ for amenable groups.

Here is a very important corollary of these results:

[13] together with the fact that in a Cayley graph all the vertices are the same

[14] The constant C depends on many things, but what is important is that the only place where a dependency in n or t occurs is in the ration $|F_n|/|A_n|$. Fortunately, taking $n \to \infty$ then makes the dependency on t disappear.

[15] Some constants will explode as $\epsilon \to 0$, but that's not what matters for compression

[16] U is finitely representable in V if for every $\epsilon > 0$ and for every finite dimensional vector subspace F of U, there is a linear operator $T : F \to V$ such that $\|x\|_U \leq \|Tx\|_V \leq (1+\epsilon)\|x\|_U$ for any $x \in F$.

Corollary 2.5

Let Γ be an amenable group. Then $\alpha^\sharp_{L_p}(\Gamma)$ is an invariant of quasi-isometry.

Indeed, it is obvious that $\alpha_{L_p}(\Gamma)$ is an invariant of quasi-isometry. Since $\alpha^\sharp_{L_p} = \alpha_{L_p}$ for amenable groups the conclusion follows. This is FALSE for non-amenable groups: there are groups with $\alpha^\sharp_{L_p} = \frac{1}{p}$ which are quasi-isometric to groups with $\alpha^\sharp_{L_p} = 0$ (see the appendix in Carette [11]).

What is true (for any group) is that $\alpha^\sharp_X(\Gamma)$ does not depend on the choice of generating set.

Question 2.6 (Question 10.2 in Naor & Peres [19]). Find a hypothesis that ensures there is a [equivariant] coarse embedding which realises the compression exponent?

Indeed, there are extremely few cases where this is known to be the case. For example for Abelian groups this is true (they have a bi-Lipschitz embedding in Euclidean space). In de Cornulier, Valette & Tessera [12] it is shown that a large class of groups among those having $\alpha^\sharp_{L_2} = 1$, these are the only ones.

Question 2.7 (Conjecture 1 in de Cornulier, Valette & Tessera [12]). Let Γ be a *compactly* generated group and assume Γ has a bi-Lipschitz embedding in a Hilbert space. Does Γ act co-compactly on some Euclidean space?

In particular, are the only amenable groups with a [equivariant] bi-Lipschitz embedding in Hilbert spaces virtually-Abelian groups? [In [12], it is shown that $\alpha_{L_2} = \alpha^\sharp_{L_2}$ for compactly generated amenable groups.]

Here is another "particular case" of 2.6. There are quite a few groups which are known to have a bi-Lipschitz embedding in L_1 (*e.g.* Abelian groups, Free groups, $\mathbb{Z}_2 \wr \mathbb{Z}$, ...). It is known that $\alpha_{L_1}(\mathbb{Z}_2 \wr \mathbb{Z}^2) = 1$.

Question 2.8 (Question 10.1 in Naor & Peres [19]). Has $\mathbb{Z}_2 \wr \mathbb{Z}^2$ a bi-Lipschitz embedding in L_1?

The compression of wreath products is largely unknown when the base has exponential growth:

Question 2.9 (Just before Question 10.7 in Naor & Peres [19]). Compute $\alpha_{L_p}\Big(\mathbb{Z}_2 \wr (\mathbb{Z}_2 \wr \mathbb{Z}^2)\Big)$.

Arzhantseva, Druţu & Sapir [2] constructed for any $a \in [0,1]$, groups with $\alpha_{L_2}(\Gamma) = a$. However, these groups are non-amenable. For amenable groups, the values computed for α_{L_2} fall in a very small set.

Question 2.10 (Question 7.6 in Naor & Peres [18]). Is there a finitely generated amenable group Γ with $\alpha_{L_2}(\Gamma) \in]\frac{2}{3}, 1[$?

In fact, the only values which are known to be taken so far are $\{2^{k-1}/(2^k-1)\}_{k\geq 0}$, $\frac{1}{2}$ and 0. For β, more is known: there are groups with speed exponent $\frac{1}{2}$, $\{1-\frac{1}{2^k}\}_{k\geq 1}$ and 1. Furthermore, Amir & Virág [1] showed that for any $b \in]\frac{3}{4}, 1[$ there is a group Γ (of intermediate growth) such that $\beta_{\Gamma,S} = b$ (for some S). The range $]\frac{1}{2}, \frac{3}{4}[$ (which corresponds in compression to the range $]\frac{2}{3}, 1[$) remains unknown.

Here is somehow a more fuzzy question:

Question 2.11 (Question 10.4 in Naor & Peres [19]). Can one say something about the set of values defined by $\alpha^\sharp_{L_p}(\Gamma)$ as Γ runs over all finitely presented groups? (except that it is countable)

3 Proof of theorem 1.11

3.1 Cocycles and the main "lemma"

An equivariant coarse embedding is, in fact, very constrained. Indeed, one may (by translating everything) always put $f(e) = 0 \in B$ for simplicity. Next, recall that a surjective isometry of a [real] Banach space is always affine[17] (Mazur-Ulam theorem). Write

(MU) $$\pi(y)v = \lambda(y)v + b(y)$$

where λ is a map from Γ into the linear isometries of B and b is a map from Γ to B. Note that $f(y) = \pi(y)f(e) = \pi(y)0 = b(y)$. Using (MU) on each side of the equality $\pi(xy)v = \pi(x)\pi(y)v$ (which holds for all $v \in B$) implies

$$\lambda(xy)v + b(xy) = \lambda(x)\big(\lambda(y)v + b(y)\big) + b(x) = \lambda(x)\lambda(y)v + \lambda(x)b(y) + b(x).$$

This means λ is a homomorphism and b satisfies the cocycle relation:

$$b(xy) = \lambda(x)b(y) + b(x).$$

It is left as an exercise to check that a cocycle is always 1-Lipschitz.

The main "lemma" (a very nice theorem) enables to retrieve the bound coming from Markov type by assuming instead the equivariance of the cocycle (and the smoothness). Let us start with a simple case in the Hilbertian setting.

Lemma 3.1 (*Hilbert case*)

Let $\mathscr{H}$ be a Hilbert space and let b be a cocycle for [the linear representation] $\lambda : \Gamma \to \mathrm{Isom}(\mathscr{H})$, then, for any $k > 0$,

$$\mathbb{E}\Big(||b(W_{2^k})||^2\Big) \leq 2^k \mathbb{E}\Big(||b(W_1)||^2\Big).$$

PROOF : Denote by σ_i each of the uniformly distributed letter, *i.e.* $W_t = \prod_{i=1}^t \sigma_i$. Next consider $W_t^{-1} = \sigma_t^{-1} \cdots \sigma_1^{-1}$ and $W_t^{-1}W_{2t} = \sigma_{t+1} \cdots \sigma_{2t}$. These two variables are i.i.d., hence so are $Y_1 = b(W_t^{-1})$ and $Y_2 = b(W_t^{-1}W_{2t})$. Use the cocycle relation on Y_2 to get

(*) $$Y_2 = b(W_t^{-1}W_{2t}) = b(W_t^{-1}) + \lambda(W_t^{-1})b(W_{2t}) \quad \text{hence} \quad Y_2 - Y_1 = \lambda(W_t^{-1})b(W_{2t})$$

Remembering that λ is isometric and that W_t, Y_1 and Y_2 are i.i.d,

$$\begin{aligned}
\mathbb{E}||b(W_{2t})||^2 &\overset{\text{isom.}}{=} \mathbb{E}||\lambda(W_t^{-1})b(W_{2t})||^2 &&\overset{(*)}{=} \mathbb{E}||Y_2 - Y_1||^2 \\
&= \mathbb{E}\Big(||Y_1||^2 + ||Y_2||^2 - 2\langle Y_1 \mid Y_2\rangle\Big) &&\overset{\text{i.i.d.}}{=} 2\mathbb{E}\Big(||Y_1||^2\Big) - 2\langle \mathbb{E}(Y_1) \mid \mathbb{E}(Y_2)\rangle \\
&\overset{\text{i.i.d.}}{=} 2\mathbb{E}\Big(||Y_1||^2\Big) - 2||\mathbb{E}(Y_1)||^2 &&\leq 2\mathbb{E}\Big(||Y_1||^2\Big)
\end{aligned}$$

To conclude, recall Y_1 has the same distribution as $b(W_t)$ and use induction. ■

[17] surjectivity is automatically assume here, since $\pi(\gamma)$ must have an inverse: $\pi(\gamma^{-1})$.

3.2 Smoothness, martingales ...

Proving the full statement will require another important theorem due to Pisier [21]. This result requires to recall two notions.

Definition 3.2. A Banach space X is said to have a modulus of **smoothness of power type** p if there exists a $K > 0$ such that

$$\rho_X(\tau) := \sup\left\{ \frac{||x+\tau y|| - ||x-\tau y||}{2} - 1 \,\middle|\, x, y \in X \text{ and } ||x|| = ||y|| = 1 \right\} \leq K\tau^p. \qquad \bigstar$$

Being uniformly smooth is equivalent to $\lim_{t\to 0} \frac{\rho_X(t)}{t} = 0$. One can check that necessarily $p \leq 2$ (exercise!).

Proposition 3.3 (*see Proposition 8 in [6]*)

X has modulus of smoothness of power type p if and only if there is a constant $S > 0$ such that

$$||x+y||^p + ||x-y||^p \leq 2||x||^p + S_p^p ||y||^p.$$

for any $x, y \in X$.

For the record, L_p has a modulus of smoothness with power type $\min(2,p)$ for $p \in [1, \infty[$. See Benyamini & Lindenstrauss [9, Appendix A] for more on this topic.

As for martingales, we will present the definition in a simplified context. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and $\mathscr{G} \subset \mathscr{F}$ be a sub-σ-algebra of $\mathscr{F}$ and let $X : (\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$ be a (real-valued!) random variable. One could ask what is the $\mathscr{G}$-measurable content of X?

Assume (for simplicity) that $X \in L_2(\Omega, \mathscr{F}, \mathbb{P})$ (*i.e.* $\mathbb{E}(X^2) < +\infty$) and denote $\langle X \mid Y \rangle = \mathbb{E}(XY)$. Define $\mathbb{E}(X \mid \mathscr{G})$ to be (the random variable) given by the projection of X on $L_2(\Omega, \mathscr{G}, \mathbb{P})$. In other words (this is done by picking $Y = \mathbb{1}_A$), this is equivalent to

$$\forall A \in \mathscr{G}, \int_A X \mathrm{d}\mathbb{P} = \int_A \mathbb{E}(X \mid \mathscr{G}) \mathrm{d}\mathbb{P}.$$

The important consequence of this (using $A = \Omega \in \mathscr{G}$) is

$$\mathbb{E}\Big(\mathbb{E}(X \mid \mathscr{G})\Big) = \mathbb{E}(X).$$

Example 3.4. If X is already $\mathscr{G}$-measurable then $\mathbb{E}(X \mid \mathscr{G}) = X$.

Let $\mathscr{X}$ be the σ-algebra generated by $X^{-1}(U)$ (for $U \subset \mathbb{R}$). If $\mathscr{X}$ and $\mathscr{G}$ are independent[18], then $\mathbb{E}(X \mid \mathscr{G}) = \mathbb{E}(X)$. Indeed, since $X - \mathbb{E}(X \mid \mathscr{G})$ is orthogonal to $L_2(\Omega, \mathscr{G}, \mathbb{P})$, for any $B \in \mathscr{G}$,

$$0 = \langle X - \mathbb{E}(X \mid \mathscr{G}) \mid \mathbb{1}_B \rangle = \mathbb{E}\Big(\big(X - \mathbb{E}(X \mid \mathscr{G})\big)\mathbb{1}_B\Big) \overset{\text{indep.}}{=} \mathbb{E}\big(X - \mathbb{E}(X \mid \mathscr{G})\big)\mathbb{E}(\mathbb{1}_B). \qquad \clubsuit$$

Definition 3.5. A sequence of random variables $(X_n)_{n\geq 0}$ is a **martingale** with respect to the filtration[19] $\mathscr{F}_n$ if

$$\mathbb{E}(X_{n+1} \mid \mathscr{F}_n) = X_n. \qquad \bigstar$$

[18] This means that $\forall A \in \mathscr{X}$ and $\forall B \in \mathscr{G}$, one has $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

[19] increasing sequence of σ-algebras.

The general idea is that $\mathscr{F}_n$ describes the total information that will be available at time n. The interpretation (in its original appearance, where two players play a game with N rounds and X_n is the gain or loss after n rounds) is that the expected future gain cannot be predicted from past information.

Example 3.6. Let Γ be a group, S a finite generating set and $\mu = \mathbb{1}_S/|S|$. The underlying σ-algebra for the random walk W_t is given by the "cylindrical" sets: subsets of $S^{\mathbb{N}}$ of the form $x \times S^{\mathbb{N}}$ where $x \in S^k$ (for some $k \in \mathbb{N}$). Let $\mathscr{F}_n = \{x \times S^{\mathbb{N}} \mid x \in S^n\}$.

A function $f : \Gamma \to \mathbb{R}$ is said to be harmonic if it satisfies the mean value property: $f(\gamma) = \frac{1}{|S|}\sum_{s \in S} f(\gamma s)$ or $f = f * \mu$. The claim is that $f(W_t)$ is a martingale (for $\mathscr{F}_n$):

$$\mathbb{E}\Big(f(W_{n+1}) \mid \mathscr{F}_n\Big) = \mathbb{E}\Big(f(W_n\sigma_{n+1}) \mid \mathscr{F}_n\Big) = \frac{1}{|S|}\sum_{s \in S} f(W_n s) \overset{\text{harm.}}{=} f(W_n).$$

where the second equality comes from the fact that the σ-algebra of σ_{n+1} is independent of $\mathscr{F}_n$ and that W_n is $\mathscr{F}$-measurable (see Example 3.4). ♣

3.3 ... and the full "lemma"

Let us now state the theorem of Pisier [21] which will be crucial for the lemma. (The constant below is taken from Theorem 4.2 in Naor, Peres, Schramm & Sheffield [20].)

Theorem 3.7

Let $1 \leq p \leq 2$ and X be a Banach space with modulus of smoothness of power type 2. Let $\{M_t\}_{t=0}^n$ be a martingale with value in X, then

$$\mathbb{E}(||M_n - M_0||^p) \leq K_X \sum_{t=0}^{n-1} \mathbb{E}(||M_{k+1} - M_k||^p).$$

where $K_X = S_p(X)^p/(2^{p-1}-1)$ and $S_p(X)$ is the constant from Proposition 3.3.

Here is the full statement of the main "lemma":

Theorem 3.8

Let X be a Banach space with modulus of smoothness with power type p, let b be a cocycle for [the linear representation] $\lambda : \Gamma \to \mathrm{Isom}(X)$. Then, for any $t \geq 1$,

$$\mathbb{E}(||b(W_t)||^p) \leq C_p(X) t \mathbb{E}(||b(W_1)||^p),$$

where $C_p(X) = 2^{2p} S_p(X)^p/(2^{p-1}-1)$.

Proof : Recall that σ_i are the i^{th} randomly chosen letter, *i.e.* $W_t = \prod_{i=1}^t \sigma_i$. Also the σ_i are i.i.d. uniformly in S.

First try would be to use the cocycle identity repeatedly on W_t:

$$\begin{aligned} b(W_t) &= b(W_{t-1}) + \lambda(W_{t-1})b(\sigma_t) \\ &= b(W_{t-2}) + \lambda(W_{t-2})b(\sigma_{t-1}) + \lambda(W_{t-1})b(\sigma_t) \\ &\vdots \\ &= \textstyle\sum_{j=1}^t \lambda(W_{j-1})b(\sigma_j), \end{aligned}$$

with the convention that $W_0 = e$. Unfortunately, this is not a Martingale, so there is not much hope to get anywhere.

Second try is to change this a bit to be a martingale: write $v = \mathbb{E}\Big(b(W_1)\Big)$ and let

$$\begin{aligned} M_t &= \sum_{j=1}^{t} \lambda(W_{j-1})\Big(b(\sigma_j) - v\Big) \\ &= \sum_{j=1}^{t} \lambda(\sigma_1\sigma_2\cdots\sigma_{j-1})\Big(b(\sigma_j) - v\Big) \end{aligned}$$

Let's check that M_t is a martingale:

$$\begin{aligned} \mathbb{E}(M_t \mid \sigma_1, \dots, \sigma_{t-1}) &= \mathbb{E}\Big(M_{t-1} + \lambda(\sigma_1\cdots\sigma_{t-1})\Big(b(\sigma_t) - v\Big) \Big| \sigma_1, \dots, \sigma_{t-1}\Big) \\ &\overset{\sigma_t \perp \sigma_j}{=} M_t + \lambda(\sigma_1\cdots\sigma_{t-1})\Big(\mathbb{E}\Big(b(\sigma_t)\Big) - v\Big) \\ &\overset{W_1 \sim \sigma_t}{=} M_t. \end{aligned}$$

Thus M_t is a martingale. However, we are adding t terms with a v, so there is not much chance of getting anywhere either.

Third try is to note that there is another way of writing the cocycle relation:
$(\cdot)$ $b(e) = b(e^2) = b(e) + \lambda(e)b(e) = 2b(e)$ (because $\lambda(e)$ is the identity) so that $b(e) = 0$.
$(\cdot\cdot)$ $0 = b(e) = b(x^{-1}x) = b(x^{-1}) + \lambda(x^{-1})b(x) = b(x^{-1}) + \lambda(x)^{-1}b(x)$, so that $b(x) = -\lambda(x)b(x^{-1})$.

Using $(\cdot\cdot)$ we have that

$$b(xy) = b(x) - \lambda(xy)b(y^{-1}).$$

Iterating this on $b(W_t)$ as above gives:

$$b(W_t) = -\sum_{j=1}^{t} \lambda(W_j)b(\sigma_j^{-1}).$$

What is great in the above expression is the "$-$" sign. Indeed, when adding the terms in v in M_t one can add them symmetrically in this expression. In the meantime, just tag the sum of our two identities by

$$\text{(C)} \qquad 2b(W_t) = \sum_{j=1}^{t} \lambda(W_{j-1})b(\sigma_j) - \sum_{j=1}^{t} \lambda(W_j)b(\sigma_j^{-1}).$$

Next introduce a quantity similar to M_t:

$$\begin{aligned} N_t &:= \sum_{j=1}^{t} \lambda(W_t^{-1}W_j)\Big(b(\sigma_j^{-1}) - v\Big) \\ &= \sum_{j=1}^{t} \lambda(\sigma_t^{-1}\sigma_{t-1}^{-1}\cdots\sigma_{j+1}^{-1})\Big(b(\sigma_j) - v\Big) \end{aligned}$$

Since S is symmetric, σ_i and σ_j^{-1} have the same distribution (and are independent unless $i = j$). Thus, M_t and N_t actually have the same distribution! Furthermore, (C) reads:

$$\text{(C')} \qquad 2b(W_t) = M_t + \lambda(W_t)N_t - v + \lambda(W_t)v.$$

Now we are ready to finish. Since

$$\text{(TI}p\text{)} \qquad ||\sum_{i=1}^{k} a_i||^p \leq k^{p-1} \sum_{i=1}^{k} ||a_i||^p$$

and $\lambda(\cdot)$ is an isometry:

$$\begin{array}{rcl} 2^p\mathbb{E}(||b(W_t)||^p) & \overset{\text{(C')}}{\leq} & \mathbb{E}(||M_t + \lambda(W_t)N_t - v + \lambda(W_t)v||^p) \\ & \overset{\text{(TI}p\text{)}}{\leq} & 4^{p-1}\mathbb{E}(||M_t||^p) + 4^{p-1}\mathbb{E}(||\lambda(W_t)N_t||^p) + 4^{p-1}||v||^p + 4^{p-1}||\lambda(W_t)v||^p \\ & \overset{\text{isom.}}{\leq} & 4^{p-1}\mathbb{E}(||M_t||^p) + 4^{p-1}\mathbb{E}(||N_t||^p) + 2\cdot 4^{p-1}||v||^p \\ & \overset{M_t \sim N_t}{=} & 2\cdot 4^{p-1}\mathbb{E}(||M_t||^p) + 2\cdot 4^{p-1}||\mathbb{E}\big(b(W_1)\big)||^p \\ & \leq & 2\cdot 4^{p-1}\mathbb{E}(||M_t||^p) + 2\cdot 4^{p-1}\mathbb{E}\Big(||b(W_1)||^p\Big) \end{array}$$

Note that the very last term on the right is already a bound like the one we are after. Hence a bound on $\mathbb{E}(||M_t||^p)$ is the only thing which stands in the way. This is where Theorem 3.7 comes in ($M_0 = 0$): indeed, since M_t is a martingale, then

$$\begin{array}{rclcl} \mathbb{E}(||M_t||^p) & \overset{\text{Th.3.7}}{\leq} & K\sum_{k=0}^{t-1}\mathbb{E}(||M_{k+1} - M_k||^p) & = & K\sum_{k=0}^{t-1}\mathbb{E}(||b(\sigma_k) - v||^p) \\ & \overset{\text{TI}p}{\leq} & Kt2^{p-1}\mathbb{E}(||b(\sigma_k)||^p + ||v||^p) & \overset{\sigma_k \sim W_1}{=} & Kt2^p\mathbb{E}(||b(W_1)||^p). \end{array}$$

This completes the proof. ■

3.4 The proof

With Theorem 3.8 in hand, the proof of Theorem 1.11 basically goes as the one of Theorem 1.10. Recall the classical result

Theorem 3.9 (*Jensen's inequality*)

> *If Z is a (real-valued) random variable and $g : \mathbb{R} \to \mathbb{R}$ is convex then $g\big(\mathbb{E}(Z)\big) \leq \mathbb{E}\big(g(Z)\big)$.*

Functions $x \mapsto x^\lambda$ are convex for $\lambda \geq 1$.

Proof of Theorem 1.11: Note that the desired inequality is trivial if $\alpha^\sharp \leq 1/p$ (since $\beta \leq 1$) or $\widehat{\beta} = 0$. So pick $\alpha_0 \in [\frac{1}{p}, \alpha^\sharp[$, $\beta_0 \in]0, \widehat{\beta}[$ and some generating set with $\beta_0 < \beta_{\Gamma,S}$. Pick b a cocycle with $\rho_f(t) \geq Kt^{\alpha_0}$.

As before, we will use the following ingredients: (K is always some constant > 0)

(tmL)	Theorem 3.8: $\mathbb{E}(\|\|b(W_t)\|\|^p) \leq Kt\mathbb{E}(\|\|b(W_1)\|\|^p)$	(tmL)
(1-L)	The embedding is 1-Lipschitz: $\|\|b(x) - b(e)\|\| \leq Kd(x,e)$	(1-L)
(ρ_f)	The compression function lower bound: $\|\|b(x) - b(e)\|\| \geq Kd(x,e)^{\alpha_0}$	(ρ_f)
(spd)	The speed: $\mathbb{E}\Big(d(W_t, e)\Big) \geq Kt^{\beta_0}$	(spd)
($\cdot$)	Cocycle are normalised: $b(e) = 0$.	($\cdot$)
(Jen)	Jensen's inequality: $\mathbb{E}(Z^{p\alpha_0}) \geq \Big(\mathbb{E}(Z)\Big)^{p\alpha_0}$.	($\cdot$)

Note that Jensen's inequality require $p\alpha_0 \geq 1$. On one side:

$$\mathbb{E}(||b(W_t)||^p) \overset{\text{tmL}}{\leq} K_1t\mathbb{E}(||b(W_1)||^p) \overset{(\cdot)}{=} K_1t\mathbb{E}(||b(W_1) - b(e)||^p) \overset{\text{(1-L)}}{=} K_2t\mathbb{E}\Big(d(W_1, e)\Big)||^p) = K_2t.$$

On the other:

$$\mathbb{E}(\|b(W_t)\|^p) \overset{(\cdot)}{=} \mathbb{E}(\|b(W_t) - b(e)\|^p) \overset{(\rho_f)}{\geq} K_3\mathbb{E}(d(W_t, e)^{p\alpha_0}) \overset{(\mathrm{Jen})}{\geq} K_3\mathbb{E}\Big(d(W_t, e)\Big)^{p\alpha_0} \geq K_4 t^{p\alpha_0\beta_0}.$$

Combining these two equations, taking $\alpha_0 \to \alpha$ and $\beta_0 \to \widehat{\beta}$ gives the conclusion ■

3.5 Comments and Questions

Theorem 1.11 is sharp for the free groups on 2 generators, where $\widehat{\beta} = 1$ and $\alpha^\sharp_{L_p} \min(2, p) = 1$ (see Naor & Peres [18, §2] for details).

It is absolutely unclear what happens to the compression exponents under relatively generic group operations, *e.g.* semi-direct products. The best example is the semi-direct product $\Gamma = \mathbb{Z}^2 \rtimes \mathrm{SL}_2(\mathbb{Z})$ (with the obvious action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{Z}^2$). Kazhdan showed it has no proper action on a Hilbert space (more precisely it has property (T)). As a consequence it has equivariant compression exponent 0, although the equivariant compression of $\mathbb{Z}^2$ and $\mathrm{SL}_2(\mathbb{Z})$ are well-known (respectively 1 and $\frac{1}{2}$).

Naor & Peres also show in [18, Lemma 2.3] that $\alpha^\sharp_{L_p}(\Gamma) \geq \alpha^\sharp_{L_2}(\Gamma)$ for any $p \geq 1$. In the non-equivariant case, more is known: for $1 \leq p \leq q \leq 2$, L_q embeds isometrically in L_p, so $\alpha_{L_p}(\Gamma) \geq \alpha_{L_q}(\Gamma)$. Also L_2 embeds isometrically in L_p for any p, so $\alpha_{L_p}(\Gamma) \geq \alpha_{L_2}(\Gamma)$ for all $p \geq 1$. See [19, Paragraph before Question 10.4] for more inequalities (they involve isometric embeddings between "snowflaked" L_ps).

Yu [24] has shown that hyperbolic groups have $\alpha^\sharp_{L_p} \geq \frac{1}{p}$ for some p large enough. See also Bourdon [10].

Question 3.10 (Question 7.7 in Naor & Peres [18]). Assume Γ is hyperbolic. Is it true that $\alpha^\sharp_{L_p}(\Gamma) \geq \frac{1}{2}$ for some p? or at least $\alpha^\sharp_{L_p}(\Gamma) \to \frac{1}{2}$ as $p \to \infty$?

There are also some potential subtleties which are not investigated. In the cases where a computation is done, the upper bound is done in L_p but the actual coarse embedding is in ℓ_p. This enables to show $\alpha^?_{L_p} = \alpha_{\ell_p}$. Yet, it needs not be the case in general. Baudier [8, Corollary 14] showed that $\alpha_{L_p} = \alpha_{\ell_p}$.

Question 3.11 (Question 10.7 in Naor & Peres [19]). Is there a group with $\alpha^\sharp_{L_p} \neq \alpha^\sharp_{\ell_p}$?

Lastly, let us mention a theorem from the article where compression exponents were introduced: if $\alpha_{L_2}(\Gamma) > \frac{1}{2}$ then Γ is exact (see Guentner & Kaminker [14, Theorem 3.2]).

4 More about random walks

The speed exponent satisfies a monotonicity for surjective homomorphism. However, in order to state it correctly, one needs to allow other measures for the choice of the random letter σ_i than the uniform measure. Hence, instead of having P being uniform, it may only be some finitely supported[20] measure which satisfies $P(s) = P(s^{-1})$.

[20] One may also say interesting things about non-finitely supported measure if some "moment" condition holds.

Lemma 4.1

Assume $\phi : \Gamma \twoheadrightarrow H$ is a surjective homomorphism. Let S be generating for G be such that $\phi(S)$ generates H. Then $\beta_{\Gamma,S} \geq \beta_{H,P}$ where $P(h) = |\phi^{-1}(h)|/|S|$.

In particular, since it is relatively easy to show $\beta_{\mathbb{Z},P} = 1/2$ for any symmetric P, one gets that $\beta_{\Gamma,S} \geq 1/2$ for any Γ with non-trivial map to $\mathbb{Z}$. As mentioned above, there is a way of proving that $\beta_{\Gamma,S} \geq 1/2$ for any group (due to Virág).

Proof : Let d_H be the distance of the Cayley graph with respect to S_H = support of P. Define the function $d' : \Gamma \to \mathbb{N}$ by $d'(\gamma) = d_H\big(\phi(\gamma), e\big)$. Note that $d'(\gamma) \leq d_\Gamma(\gamma, e)$: indeed $d_H(h_1, h_2) = d_\Gamma\big(\phi^{-1}(h_1), \phi^{-1}(h_2)\big)$, so that $d'(\gamma) = d_\Gamma(\gamma N, N)$ where $N = \ker\phi$. Let W_n^Γ be the random walker on Γ and W_n^H be the random walker on H (which moves according to P as in the statement). Note that $\mathbb{P}\big(d(W_n^H, e) = i\big) = \mathbb{P}\big(d'(W_n^\Gamma) = i\big)$. This implies

$$\mathbb{E}\big(d_H(W_n, e)\big) = \mathbb{E}\big(d'(W_n^\Gamma)\big) \leq \mathbb{E}\big(d_\Gamma(W_n^\Gamma, e)\big) \qquad \blacksquare$$

Let us mention a last (easy) example.

Example 4.2. Let us try to compute the return probability in $\mathbb{Z}^d$. Pick some symmetric generating set S. If p_n is the law of W_n, recall that $p_n = P * \cdots * P$ (with n appearances of P). Because the Fourier transform turn convolution in multiplication, let, for $\Theta \in [-\pi, \pi]^d$,

$$\phi(\Theta) = \sum_{s \in S} e^{s \cdot \Theta} P(x).$$

Fourier analysis tells us

$$p_n(x) = (2\pi)^{-d} \int_{[-1,1]^d} e^{-ix \cdot \Theta} \phi(\Theta)^n \mathrm{d}\Theta.$$

Now, pick $S^+ \subset S$ so that $S = S^+ \cup -S^+$.

$$\phi(\Theta) = \sum_{s \in S} e^{is \cdot \Theta} P(s) = \sum_{s \in S^+} \cos(s \cdot \Theta) P(s)$$

This shows ϕ is real and $\phi(\Theta) = 1$ if and only if $\Theta = 0$. This can be made even more visible by writing

$$\phi(\Theta) = 1 - \sum_{s \in S^+} \big(1 - \cos(s \cdot \Theta)\big) P(s).$$

Pick $b_1, \ldots, b_d \in S$ a basis of $\mathbb{R}^d$, then,

$$\phi(\Theta) \leq 1 - \sum_{s \in \{b_1, \ldots, b_d\}} \big(1 - \cos(s \cdot \Theta)\big) P(s) \leq 1 - C|\Theta|^2$$

for some $C > 0$ depending on $\{b_1, \ldots, b_d\}$ and $|S|$. This last expression is very useful, for example one gets, using $1 - C|\Theta|^2 \leq e^{-C|\Theta|^2}$:

$$p_n(0) = (2\pi)^{-d} \int_{[-1,1]^d} \phi(\Theta)^n \mathrm{d}\Theta \leq (2\pi)^{-d} \int_{[-1,1]^d} e^{-nC|\Theta|^2} \mathrm{d}\Theta \leq K/n^{d/2}$$

for some $K > 0$. ♣

Further computations lead to $\beta_{\mathbb{Z}^d,S} = 1/2$.

5 Examples of equivariant compression

The simplest example is for $\mathbb{Z}^d$

Example 5.1. Let $\Gamma = \mathbb{Z}^d$ and $f : \Gamma \to \mathbb{R}^d$ be the identity map. Put the ℓ^p-norm on $\mathbb{R}^d$ and note that $||d_\Gamma(x,y)|| = ||x-y||_{\ell^1}$

$$d^{\frac{1}{p}-1}||x-y||_{\ell^1} \leq ||x-y||_{\ell^p} \leq ||x-y||_{\ell^1}.$$

Hence ρ_f can be taken to be a linear function. It turns out this is also a "equivariant" embedding. Indeed, this is the cocycle obtained by looking at Γ acting by translation on $\mathbb{R}^d$ (as a subgroup). The linear representation corresponding to this action is the trivial representation (cocycles for the trivial representation are just homomorphisms). ♣

Free groups were also among the very first case investigated in Guentner & Kaminker [14].

Example 5.2. Assume Γ is a free group with generators $a_1, \ldots, a_d$. Take ϕ = characteristic functions of words beginning with a_1, *i.e.*

$$\phi(g) = \begin{cases} 1 & \text{if } g = aw \text{ where } w \text{ is a word and } aw \text{ is reduced} \\ 0 & \text{else.} \end{cases}$$

Note that this function can be used to define a cocycle with values in $\ell^p(\Gamma)$ (for the right-regular representation $\rho_{\ell^p\Gamma}$): Indeed,

$$b(\gamma)(g) := \phi(g) - \rho_\gamma\phi(g) = \phi(g) - \phi(g\gamma) = \begin{cases} -1 & \text{if } g = w^{-1} \text{ and } \gamma = waw'(\text{reduced}) \\ 1 & \text{if } g = aw^{-1} \text{ where } \gamma = wa^{-1}w'(\text{reduced}) \\ 0 & \text{else.} \end{cases}$$

so $||b(\gamma)||_p^p$ =number of appearance of the letter $a_1^{\pm 1}$ in γ.

Now take another cocycle for each letter a_i and look at the cocycle b' given by the direct sum (*i.e.* Γ acts on $\oplus_{i=1}^d \ell^p\Gamma$ diagonally). One has $||b(\gamma)||_p^p \geq K|g|$ (for some $K > 0$) so that $\alpha_p^\sharp(F_d) \geq 1/p$. Theorem 1.11 shows this is an equality for $p \in [1,2]$. For $p > 2$, the correct value is $\frac{1}{2}$, see Naor & Peres [18, Lemma 2.3]. ♣

The "trick" in the previous example is sometimes referred to as "virtual coboundary". Indeed, if ϕ would be in $\ell^p\Gamma$, b would be a [usual] coboundary. Here, $\phi \notin \ell^p\Gamma$ but nevertheless b is well-defined.

Example 5.3. Let $\mathcal{B} := \ell^p\Gamma$ and let Γ act diagonally on each factor by the right-regular representation. Define a cocycle via a "virtual coboundary", *i.e.* first put $f = \sum_{n\geq 0} a_n\phi_n$ where $\phi_n \in \ell^p\Gamma$ and $a_n \in \mathbb{R}$ are to be chosen. We would like to define a cocycle by

$$b(\gamma) = f - \rho_\gamma f.$$

We need to insure that, for each γ, $b(\gamma)$ is indeed in $\ell^p\Gamma$. Using the cocycle relation, it suffices to check this for $s \in S$: we need

$$||\rho_s f - f||_p^p = \sum_n \alpha_n^p ||\rho_s\phi_n - \phi_n||_p^p < +\infty.$$

Just let $||\nabla\phi_n||_p := \sum_{s\in S}||\rho_s\phi_n - \phi_n||_p$ and pick $\alpha_n^p = ||\nabla\phi_n||_p^{-p}n^{-1-\epsilon}$, where $\epsilon > 0$. Next (see exercises), we only need a lower bound on $||b(g)||$.

$$||b(g)||_p^p = \sum_{n\geq 1} n^{-1-\epsilon}\frac{||\phi_n - \rho_g\phi_n||_p^p}{||\nabla\phi_n||_p^p} \geq \sum_{n|\operatorname{diam}\sup_\phi n<|g|} n^{-1-\epsilon}\frac{||\phi_n||_p^p}{||\nabla\phi_n||_p^p}.$$

Take $\phi_n = \mathbb{1}_{F_n}$ where F_n is a sequence of finite sets such that $F_n \subsetneq F_{n+1}$ and $\partial F_n \cap \partial F_{n+1} = \emptyset$ (here ∂F is the set of edges between F and F^c). Then $||\nabla f||_p \leq K|\partial F_n|$ for some $K > 0$.

$$\sum \alpha_n^p|\partial F_n| < +\infty \quad\Rightarrow\quad \alpha_n^p := n^{-1-\epsilon}|\partial F_n|^{-1}.$$

where $\epsilon > 0$ is arbitrary. This gives, for $n = \sup\{l \mid \operatorname{diam} F_n < |g|\}$,

$$||b(g)||_p^p \geq \sum_{k=1}^n\left(\sum_{i=k}^n\alpha_i\right)^p(|F_k| - |F_{k-1}|)$$

with $F_0 = \emptyset$. Now, assuming further the α_n are decreasing, for $p \in \mathbb{R}_{\geq 1}$ the inner sum can be written as

$$||b(g)||_p^p \geq \sum_{k=1}^n(n-k+1)^p\alpha_n^p(|F_k| - |F_{k-1}|)$$

Since $\sum_{k=1}^c a_k(b_k - b_{k-1}) = b_ca_c + \sum_{k=1}^{c-1}b_k(a_k - a_{k+1})$ (given $b_0 = 0$), one has

$$||b(g)||_p^p \geq \alpha_n^p|F_n| + \alpha_n^p\sum_{k=1}^{n-1}|F_k|((n-k+1)^p - (n-k)^p)$$

So let

$$R_n = \frac{1}{|\partial F_n|}\sum_{k=1}^n|F_k|((n-k+1)^p - (n-k)^p)\Big)$$

then $||b(g)||_p^p \geq R_nn^{-1-\epsilon}$ for $n = \sup\{l \mid \operatorname{diam} F_n < |g|\}$.

Apply this to the case where Γ has polynomial growth, the F_n can *essentially*[21] be chosen to be sequence of balls. One finds,

$$R_n \simeq \frac{1}{n^{d-1}}\sum_{k=1}^n k^d(n-k)^{p-1} \simeq n^{p+1}$$

where the last equality was using Euler-Maclaurin with $\int_0^T x^a(T-x)^b \simeq \mathrm{cst}T^{a+b+1}$. Thus, with $n = |g|$,

$$||b(g)||_p^p \geq Kn^{p-\epsilon} \qquad \text{for some } K > 0.$$

Taking $\epsilon \to 0$ shows that $\alpha^\sharp_{\ell^p}(\Gamma) = 1$ for groups of polynomial growth. ♣

Note the above bound on compression (together with the fact that nilpotent groups surject on $\mathbb{Z}$), gives a proof that $\beta_{\Gamma,S} = 1/2$ for any generating set.

More careful computations can be used to show $\alpha^\sharp_{\ell^p}(\mathbb{Z}_2 \wr \mathbb{Z}) \geq 1/p$ (which is not sharp!.. unless $p = 1$) and gives some result for groups of intermediate growth. See Tessera [22] for a more general class of groups with $\alpha^\sharp_{\ell^p}(\Gamma) = 1$ and further techniques.

[21] one needs to pick some careful subsequence

6 Exercises

Everywhere: Γ is a finitely generated group. For hints look at the end.

Exercise 1 Let $f(n) = \mathbb{P}(W_{2n} = e_\Gamma)$. Show that $f(n+m) \geq f(n) \cdot f(m)$. Deduce that the limit of $f(n)^{1/2n}$ exists.

Exercise 2 Let B_n be the ball of radius n in a Cayley graph of Γ. Show that $\liminf_{n\to\infty} \frac{|B_{n+1}|}{|B_n|} = 1$ implies the group is amenable.

Exercise 3 Show that the Følner condition does not depend on the generating set S, *i.e.* $\{F_n\}$ are Følner sets for S they are also Følner sets for T.

Exercise 4 Show that $\mathbb{E}\Big(d(W_{n+m}, e_\Gamma)\Big) \leq \mathbb{E}\Big(d(W_m, e_\Gamma)\Big) + \mathbb{E}\Big(d(W_n, e_\Gamma)\Big)$.

Exercise 5 Given a cocycle b, many of the conditions for a coarse embedding are automatically satisfied.

- Show that it suffice to check that $\rho_f\Big(d(e,\gamma)\Big) \leq ||b(\gamma)|| \leq Cd(e,\gamma) + C$ for all $\gamma \in \Gamma$.
- Show that a cocycle is always Lipschitz (*i.e.* the upper bound always hold): $||b(\gamma)|| \leq K d(e,\gamma)$ for some $K > 0$ (K depends on b).

Conclude that the only inequality to establish to see that a cocycle is a coarse embedding is $||b(\gamma)|| \geq \rho_f(|\gamma|)$.

Exercise 6 Show that there cannot be a Banach space with modulus of smoothness of power type p for $p > 2$. Likewise for Markov type.

Exercise 7

Hint[s] for 1: One has more chances of returning at identity after time $n+m$ as returning at time n and $n+m$. Let $g(n) = -\log f(n)$ then $g(n+m) \leq g(n) + g(m)$ and by Fekete's subadditive Lemma $\lim \frac{g(n)}{n}$ exists.

Hint[s] for 2: $\frac{|B_{n+1} \triangle B_n|}{|B_n|} = \frac{|B_{n+1}|}{|B_n|} - 1$ and $|\partial B_n| \leq |S| \cdot |B_{n+1} \triangle B_n|$

Hint[s] for 3: Write the letters of T as words in S (and vice-versa) to see that the T-boundary of a set is at most some thickening of the S boundary. The size of a regular tree being bounded, this yields a bound on the T-boundary in terms of the S-boundary.

Hint[s] for 4: All points in a Cayley graph are the same. In particular, from each possible position at time n, one has the same future after m steps. Check for the "worst case scenario" (triangle inequality).

Hint[s] for 5: Use that λ is isometric and the cocycle relation. For the second point, write γ as a word, and use the cocycle relation many times.

Hint[s] for 6:

Hint[s] for 7:

References

[1] G. Amir and B. Virág, Speed exponents for random walks on groups, arXiv:1203.6226

[2] G. N. Arzhantseva, C. Druţu, and M.V. Sapir, Compression functions of uniform embeddings of groups into Hilbert and Banach spaces, *J. Reine Angew. Math.*, **633**:213–235, 2009.

[3] T. Austin, Amenable groups with very poor compression into Lebesgue spaces, *Duke Math. J.* **159**(2):187–222, 2011.

[4] T. Austin, A. Naor and Y. Peres, The wreath product of $\mathbb{Z}$ with $\mathbb{Z}$ has Hilbert compression exponent 2/3, *Proc. Amer. Math. Soc.* **137**(1):85–90, 2009.

[5] K. Ball, Markov chains, Riesz transforms and Lipschitz maps. Geom. Funct. Anal. **2**(2), 137–172, 1992.

[6] K, Ball, E. A. Carlen and E. H. Lieb, Sharp uniform convexity and smoothness inequalities for trace norms, *Invent. Math.* **115**(3):463–482.

[7] L. Bartholdi and A. Erschler, Imbeddings into groups of intermediate growth, arXiv:1403.5584

[8] F. Baudier, On the metric geometry of stable metric spaces, arXiv:1409.7738

[9] Y. Benyamini and J. Lindenstrauss, *Geometric nonlinear functional analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*, American Mathematical Society, Providence, RI, 2000.

[10] M. Bourdon, Cohomologie et actions isométriques propres sur les espaces L^p, to appear in *Geometry, Topology and Dynamics, Proceedings of the 2010 Bangalore conference*, available at http://math.univ-lille1.fr/ bourdon/papiers/coho.pdf

[11] M. Carette (appendix by S. Arnt, T. Pillon and A. Valette), The Haagerup property is not invariant under quasi-isometry, arXiv:1403.5446

[12] Y. de Cornulier, R. Tessera and A. Valette, Isometric group actions on Hilbert spaces: growth of cocycles, *Geom. Funct. Anal.* **17**(3):770–792, 2007.

[13] A. Gournay, The Liouville property via Hilbertian compression, arXiv:1403.1195

[14] E. Guentner and J. Kaminker, Exactness and uniform embeddability of discrete groups, *J. London Math. Soc. (2)* **70**(3):703–718, 2004.

[15] H. Kesten, Full Banach mean values on countable groups, *Math. Scand.*, **7**:146–156, 1959.

[16] J. Lee and Y. Peres, Harmonic maps on amenable groups and a diffusive lower bound for random walks, *Ann. Probab.* **41**(5):3392–3419, 2013.

[17] R. Lyons with Y. Peres *Probability on Trees and Networks*, Cambridge University Press, (2014). In preparation. Current version available at `http://mypage.iu.edu/~rdlyons/`

[18] A. Naor and Y. Peres, Embeddings of discrete groups and the speed of random walks, *Int. Math. Res. Not.* IMRN 2008, Art. ID rnn 076, 34 pp.

[19] A. Naor and Y. Peres, L_p-compression, traveling salesmen, and stable walks, *Duke Math. J.* **157**(1):53–108, 2011.

[20] A. Naor, Y. Peres, O. Schramm and S. Sheffield, Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces, *Duke Math. J.* **134**(1):165–197, 2006.

[21] G. Pisier, Martingales with values in uniformly convex spaces, *Israel J. Math.* **20**(3-4), 326–350, 1975.

[22] R. Tessera, Asymptotic isoperimetry on groups and uniform embeddings into Banach spaces, *Comment. Math. Helv.*, **86**(3):499–535, 2011.

[23] J. von Neumann, Zur allgemein Theorie des Masses, *Fund. Math.* **13**:73–116, 1929.

[24] G. Yu, Hyperbolic groups and affine isometric actions on ℓ^p-spaces, *Geom. Funct. Anal.* **15**(5):1144-1151, 2005.

LOCAL VERSUS GLOBAL EMBEDDABILITY OF LOCALLY FINITE METRIC SPACES

after M. I. Ostrovskii [8]
written by
Sheng Zhang

Abstract. We will present the techniques used by M. I. Ostrovskii to prove that the Lipschitz (resp. coarse) embeddability into an infinite dimensional Banach space of a locally finite metric space is determined by its finite subsets.

1. Introduction

The purpose of this short note is to prove the following theorem by M. I. Ostrovskii. Recall that a metric space X is said to be locally finite if every ball in X contains only finitely many points.

Theorem 1.1 (Ostrovskii [8]). *Let A be a locally finite metric space whose finite subsets admit equi-Lipschitz (resp. equi-coarse) embeddings into a Banach space X. Then A admits a Lipschitz (resp. coarse) embedding into X.*

The main ingredients of the proof contain the following:

- Ultraproduct techniques in Banach space theory;
- Approaches to the selection of good-behaving subsequence;
- The gluing technique of embeddings.

The gluing technique was first introduced by F. Baudier to prove the following characterization of superreflexivity.

Theorem 1.2 (Baudier [1]). *A Banach space X is not superreflexive if and only if the infinite binary tree B_∞ equipped with the shortest path metric admits a Lipschitz embedding into X.*

Here the infinite binary tree is defined by $B_\infty = \bigcup_{i=0}^{\infty} \Omega_i$, where $\Omega_i = \{0,1\}^i$ for $i \geq 1$ and $\Omega_0 = \{\emptyset\}$, and the finite binary tree with n levels is defined similarly by $B_n = \bigcup_{i=0}^{n} \Omega_i$. By Ostrovskii's theorem one can easily see that Baudier's theorem is indeed equivalent to J. Bourgain's early result:

Theorem 1.3 (Bourgain [3]). *A Banach space X is not superreflexive if and only if the finite binary trees (B_n) equipped with the shorted path metric admit equi-Lipschitz embeddings into X.*

The next theorem, due to F. Baudier and G. Lancien, is another application of Ostrovskii's theorem.

Theorem 1.4 (Baudier-Lancien [2]). *Each locally finite metric space admits a Lipschitz embedding into any Banach space without cotype.*

This theorem is an immediate consequence of Theorem 1.1, Theorem 2.6 and the fact that each metric space with n elements admits an isometric embedding into ℓ_∞^n (see [4]).

The last application of Ostrovskii's theorem we want to mention is the following theorem by Ostrovskii. The proof follows immediately from Theorem 1.1 and the Dvoretzky's theorem (Theorem 2.4).

Theorem 1.5 (Ostrovskii [7]). *Let M be a locally finite subset of a Hilbert space. Then M admits a Lipschitz embedding into any infinite-dimensional Banach space.*

2. Preliminaries

In this section we will first list some definitions and facts that will be used throughout this note, and then briefly introduce ultraproduct of Banach spaces. A thorough discussion of ultraproduct techniques in Banach space theory can be found in [5].

2.1. Definitions and Facts.

Definition 2.1. A map $f : X \to Y$ between two metric spaces X and Y is called a Lipschitz embedding if there exists a constant $C \geq 1$ such that for all $u, v \in X$,

$$\frac{1}{C} d(u,v) \leq d(f(u), f(v)) \leq C d(u,v).$$

If this inequality holds for $C = 1$ then f is called an isometric embedding.

Let (X_n) be a sequence of metric spaces. A sequence of maps $f_n : X_n \to Y$ are called equi-Lipschitz embeddings if there exists a constant $C \geq 1$ such that for all n and all $u, v \in X_n$,

$$\frac{1}{C} d(u,v) \leq d(f_n(u), f_n(v)) \leq C d(u,v).$$

Definition 2.2. A map $f : X \to Y$ between two metric spaces X and Y is called a coarse embedding if there exist two nondecreasing functions $\rho_1, \rho_2 : [0, +\infty) \to [0, +\infty)$ with $\lim_{t\to\infty} \rho_1(t) = \infty$ such that for all $u, v \in X$,

$$\rho_1(d(u,v)) \leq d(f(u), f(v)) \leq \rho_2(d(u,v)).$$

Let (X_n) be a sequence of metric spaces. A sequence of maps $f_n : X_n \to Y$ are called equi-coarse embeddings if there exist two nondecreasing functions $\rho_1, \rho_2 : [0, +\infty) \to [0, \infty)$ with $\lim_{t\to\infty} \rho_1(t) = +\infty$ such that for all n and all $u, v \in X_n$,

$$\rho_1(d(u,v)) \leq d(f_n(u), f_n(v)) \leq \rho_2(d(u,v)).$$

Definition 2.3. A Banach space X is said to be finitely representable in a Banach space Y if for any $\varepsilon > 0$ and any finite-dimensional subspace $E \subset X$ there exists a finite-dimensional subspace $F \subset Y$ such that $d_{BM}(E, F) < 1 + \varepsilon$, where d_{BM} is the Banach-Mazur distance defined by

$$d_{BM}(E, F) = \inf\{\|T\|\|T^{-1}\| : T : E \to F \text{ is an isomorphism}\}.$$

A Banach space X is called superreflexive if every Banach space Y that is finitely representable in X is reflexive

Theorem 2.4 (Dvoretzky)**.** *ℓ_2 is finitely representable in each infinite-dimensional Banach space.*

Definition 2.5. A Banach space X is said to have (Rademacher) cotype q, $2 \le q \le \infty$, if there exists a constant $C_q > 0$ such that for every $n \in \mathbb{N}$ and every $x_1, ..., x_n \in X$,

$$\left(\operatorname*{Average}_{\varepsilon_i=\pm 1}\left\|\sum_{i=1}^{n}\varepsilon_i x_i\right\|^q\right)^{\frac{1}{q}} \ge \frac{1}{C_q}\left(\sum_{i=1}^{n}\|x_i\|^q\right)^{\frac{1}{q}}.$$

It is easy to see that every Banach space has cotype ∞ with constant 1. If a Banach space X has some cotype $q < \infty$ then we say that X has nontrivial cotype, otherwise X is said to be without cotype.

Theorem 2.6 (Maurey-Pisier)**.** *A Banach space X has only trivial cotype if and only if ℓ_∞ is finitely representable in X.*

2.2. **Ultraproduct of Banach spaces.**

Definition 2.7. A filter $\mathcal{F}$ on an infinite set I is a subset of $\mathcal{P}(I)$ (the set of all subsets of I) satisfying the following conditions:

(1) $\emptyset \notin \mathcal{F}$;
(2) $\mathcal{F}$ is closed under finite intersection.
(3) If $A \in \mathcal{F}$, then $B \in \mathcal{F}$ for each $B \supset A$.

An ultrafilter $\mathcal{U}$ on I is a maximal filter with respect to inclusion. An ultrafilter is called free if the intersection of all the sets in it is empty.

Definition 2.8. Let $\mathcal{U}$ be an ultrafilter on I. X is a topological space and $(x_i)_{i\in I} \subset X$. We say that $(x_i)_{i\in I}$ converges to $x \in X$ through $\mathcal{U}$ and write $\lim_{\mathcal{U}} x_i = x$ if $\{i \in I : x_i \in U\} \in \mathcal{U}$ for any open neighborhood U of x.

Lemma 2.9. *Let $\mathcal{U}$ be an ultrafilter on I and K be a compact set. Then any $(x_i)_{i\in I} \subset K$ converges to some $x \in K$ through $\mathcal{U}$. In particular, any bounded real-valued collection $(x_i)_{i\in I}$ converges to some $x \in \mathbb{R}$ through $\mathcal{U}$.*

Let $(X_i)_{i\in I}$ be a family of Banach spaces and $\mathcal{U}$ be a free ultrafilter on I. Consider the ℓ_∞-sum of $(X_i)_{i\in I}$, i.e., the Banach space

$$(\bigoplus_{i\in I} X_i)_\infty = \{(x_i)_{i\in I} : x_i \in X_i \text{ and } \sup_{i\in I}\|x_i\| < \infty\}$$

with the norm $\|(x_i)_{i\in I}\|_\infty = \sup_{i\in I}\|x_i\|$. In view of Lemma 2.9, for each $(x_i)_{i\in I} \in (\bigoplus_{i\in I} X_i)_\infty$, $\lim_{\mathcal{U}}\|x_i\|$ exists and defines a seminorm on $(\bigoplus_{i\in I} X_i)_\infty$. It is easy to check that the subspace of $(\bigoplus_{i\in I} X_i)_\infty$ on which the seminorm is equal to 0, denoted by $N_{\mathcal{U}}$, is closed.

Definition 2.10. The ultraproduct of $(X_i)_{i\in I}$ with respect to the free ultrafilter $\mathcal{U}$, denoted by $(\prod_{i\in I} X_i)_{\mathcal{U}}$, is the quotient space $(\bigoplus_{i\in I} X_i)_\infty / N_{\mathcal{U}}$ with the norm $\|(x_i)_{\mathcal{U}}\| = \lim_{\mathcal{U}}\|x_i\|$, where $(x_i)_{\mathcal{U}}$ is the element in $(\prod_{i\in I} X_i)_{\mathcal{U}}$ corresponding to $(x_i)_{i\in I} \in (\bigoplus_{i\in I} X_i)_\infty$. If all X_i's are the same Banach space X, then the ultraproduct is called an ultrapower of X and denoted by $X_{\mathcal{U}}$.

Proposition 2.11. *Let X be a Banach space and $\mathcal{U}$ be a free ultrafilter on I. If X is finite dimensional then $X_{\mathcal{U}}$ is of the same dimension; if X is infinite dimensional then $X_{\mathcal{U}}$ is finitely representable in X.*

3. Proof of Theorem 1.1

Fix a point O in A and consider finite subsets $A_i = \{a \in A : d(O,a) \leq 2^i\}$. By the assumption there exist equi-Lipschitz (equi-coarse) embeddings $f_i : A_i \to X$. Without loss of generality we can assume that $f_i(O) = 0$ for all $i \in \mathbb{N}$.

Let $\mathcal{U}$ be a free ultrafilter on $\mathbb{N}$. For $a \in A$, define $\widetilde{f}_i(a) = f_i(a)$ if $a \in A_i$ and $\tilde{f}_i(a) = 0$ otherwise. Then it is easy to check that the map $f : A \to X_{\mathcal{U}}$, $a \mapsto (\widetilde{f}_i(a))_{\mathcal{U}}$ is a Lipschitz (coarse) embedding. Note that in both cases the image of A under f, denoted by N, is a locally finite subset of $X_{\mathcal{U}}$ containing the origin, so we may assume that every nonzero element in N has norm at least 1. Then the theorem follows from the following lemma. □

Lemma 3.1. *N (and hence every locally finite subset of $X_{\mathcal{U}}$) admits a Lipschitz embedding into X.*

Proof. The case when X is finite dimensional is trivial by Proposition 2.11, so we assume that X is of infinite dimension. Consider finite sets $N_i = \{u \in N : \|u\| \leq 2^i\}$. Again by Proposition 2.11 there exist maps $s_i : N_i \to X$ such that $s_i(0) = 0$ and for all $u, v \in N_i$,

$$\|u - v\| \leq \|s_i(u) - s_i(v)\| \leq 2\|u - v\|. \tag{3.1}$$

To find a Lipschitz embedding, we first introduce a gluing map $\varphi : N \to X$, which pastes s_i's in the sense that for $2^{i-1} \leq \|a\| < 2^i$,

$$\varphi(a) = \frac{2^i - \|a\|}{2^{i-1}} s_i(a) + \frac{\|a\| - 2^{i-1}}{2^{i-1}} s_{i+1}(a). \tag{3.2}$$

Clearly $\|\varphi(a)\| \leq 2\|a\|$, but φ is not a Lipschitz embedding. For technical reason (see the claim below) we need another Lipschitz map $\tau : \mathbb{R}_+ \to X$ so that the map $\widehat{\varphi} : N \to X$ defined by $\widehat{\varphi}(a) = \varphi(a) + \tau(\|a\|)$ is almost the desired Lipschitz embedding (we say "almost" because the definition of φ given by (3.2) needs a small modification later, but at this moment we use (3.2) for the reason of easy understanding). To this end, consider the finite sets $T_i = \{\varphi(u) : u \in N_{i+1}\}$. Let $F_1 = \mathrm{span} T_1$ and choose $p_1 \in S_X$ so that $\mathrm{dist}(p_1, F_1) = 1$. Let $F_2 = \mathrm{span}(T_2 \cup \{p_1\})$ and choose $p_2 \in S_X$ so that $\mathrm{dist}(p_2, F_2) = 1$. Let $F_3 = \mathrm{span}(T_3 \cup \{p_1, p_2\})$...Since X is infinite dimensional, we can continue this process to get a sequence (F_i) of finite-dimensional subspaces of X and a sequence (p_i) so that $p_i \in F_{i+1}$ and $\mathrm{dist}(p_i, F_i) = 1$ for all i. Then the $\tau : \mathbb{R}_+ \to X$ is defined in the following way:

$$\tau(t) = \begin{cases} tp_1 & \text{if } 0 \leq t < 2, \\ 2p_1 + \sum_{j=2}^{k} (2^j - 2^{j-1}) p_j + (t - 2^k) p_{k+1} & \text{if } 2^k \leq t < 2^{k+1} \text{ for some } k \geq 1. \end{cases}$$

It is easy to check that τ is 1-Lipschitz. Moreover, the following claim holds.

Claim: There exists a constant $C > 0$ such that for any $a, b \in N$,

$$\|\widehat{\varphi}(a) - \widehat{\varphi}(b)\| = \|\varphi(a) - \varphi(b) + \tau(\|a\|) - \tau(\|b\|)\| \geq C(\|a\| - \|b\|).$$

We prove this claim by considering three cases.

Case 1. $2^{i-1} \leq \|b\| \leq \|a\| < 2^i$

In this case we have

$$\begin{aligned}\|\varphi(a) - \varphi(b) + \tau(\|a\|) - \tau(\|b\|)\| &= \|(\|a\| - \|b\|)p_i + \varphi(a) - \varphi(b)\| \\ &\geq \|a\| - \|b\|.\end{aligned}$$

The last inequality holds because $\mathrm{dist}(p_i, F_i) = 1$ and $\varphi(a) - \varphi(b) \in F_i$.

Case 2. $2^{i-1} \leq \|b\| < 2^i \leq \|a\| < 2^{i+1}$

In this case we have

$$\begin{aligned}&\|\varphi(a) - \varphi(b) + \tau(\|a\|) - \tau(\|b\|)\| \\ &\qquad = \|(\|a\| - 2^i)p_{i+1} + (2^i - \|b\|)p_i + \varphi(a) - \varphi(b)\|. \qquad (3.3)\end{aligned}$$

Consider two subcases:

$$\|a\| - 2^i \geq \frac{1}{4}(\|a\| - \|b\|), \qquad (3.4)$$

$$\|a\| - 2^i < \frac{1}{4}(\|a\| - \|b\|). \qquad (3.5)$$

In subcase (3.4) we have

$$(3.3) \geq \|a\| - 2^i \geq \frac{1}{4}(\|a\| - \|b\|) \qquad (3.6)$$

because $\mathrm{dist}(p_{i+1}, F_{i+1}) = 1$ and $p_i, \varphi(a) - \varphi(b) \in F_{i+1}$.

In subcase (3.5) we have

$$\begin{aligned}(3.3) &\geq \|(2^i - \|b\|)p_i + \varphi(a) - \varphi(b)\| - (\|a\| - 2^i) \\ &\geq (2^i - \|b\|) - (\|a\| - 2^i) \geq \frac{1}{2}(\|a\| - \|b\|)\end{aligned}$$

since $\mathrm{dist}(p_i, F_i) = 1$ and $\varphi(a) - \varphi(b) \in F_i$.

Case 3. $2^{k-1} \leq \|b\| < 2^k < 2^i \leq \|a\| < 2^{i+1}$

In this case we have

$$\begin{aligned}&\|\varphi(a) - \varphi(b) + \tau(\|a\|) - \tau(\|b\|)\| \\ &\qquad = \|(\|a\| - 2^i)p_{i+1} + (2^i - 2^{i-1})p_i + r + \varphi(a) - \varphi(b)\|, \qquad (3.7)\end{aligned}$$

where r is an element in F_i. Consider two subcases:

$$\|a\| - 2^i \geq \frac{1}{4}(2^i - 2^{i-1}), \qquad (3.8)$$

$$\|a\| - 2^i < \frac{1}{4}(2^i - 2^{i-1}). \qquad (3.9)$$

In subcase (3.8) we have

$$(3.7) \geq \|a\| - 2^i \geq \frac{1}{4}(2^i - 2^{i-1}) \geq \frac{1}{16}(\|a\| - \|b\|)$$

because $\operatorname{dist}(p_{i+1}, F_{i+1}) = 1$ and $p_i, r, \varphi(a) - \varphi(b) \in F_{i+1}$.

In subcase (3.9) we have

$$\begin{aligned}(3.7) &\geq \|(2^i - 2^{i-1})p_i + r + \varphi(a) - \varphi(b)\| - (\|a\| - 2^i)\\ &\geq (2^i - 2^{i-1}) - \frac{1}{4}(2^i - 2^{i-1}) \geq \frac{3}{16}(\|a\| - \|b\|)\end{aligned}$$

since $\operatorname{dist}(p_i, F_i) = 1$ and $r, \varphi(a) - \varphi(b) \in F_i$.

Remark 3.2. Note that the proof of the claim has nothing to do with the expression (3.2) of φ.

Now we have shown that $\|\widehat{\varphi}(a) - \widehat{\varphi}(b)\| \geq C(\|a\| - \|b\|)$ for some $C > 0$. For convenience we will henceforth assume $C = 1$. The rest of the proof is dedicated to show that $\widehat{\varphi}$ is a Lipschitz embedding from N into X. Again, we proceed by considering the above three cases, but in a reverse way (from the easiest to the hardest).

Case 3. $2^{k-1} \leq \|b\| < 2^k < 2^i \leq \|a\| < 2^{i+1}$

In this case we have

$$\frac{2^i - 2^k}{2^{i+1} + 2^k} \leq \frac{\|a\| - \|b\|}{\|a\| + \|b\|} \leq \frac{\|\widehat{\varphi}(a) - \widehat{\varphi}(b)\|}{\|a - b\|} \leq \frac{3(\|a\| + \|b\|)}{\|a\| - \|b\|} \leq \frac{3(2^{i+1} + 2^k)}{2^i - 2^k}.$$

Note that

$$\frac{2^{i+1} + 2^k}{2^i - 2^k} \leq \frac{2^{i+2}}{2^{i-1}} = 8,$$

so we conclude that $\widehat{\varphi}$ is a Lipschitz embedding.

Case 2. $2^{i-1} \leq \|b\| < 2^i \leq \|a\| < 2^{i+1}$

In this case

$$\begin{aligned}\varphi(a) - \varphi(b) = &-\frac{2^i - \|b\|}{2^{i-1}} s_i(b) + \frac{\|a\| - 2^i}{2^i} s_{i+2}(a)\\ &+ \frac{2^{i+1} - \|a\|}{2^i} s_{i+1}(a) - \frac{\|b\| - 2^{i-1}}{2^{i-1}} s_{i+1}(b).\end{aligned}$$

The first and the second terms both have norms at most $4(\|a\|-\|b\|)$. The norm of the last two terms can be estimated as follows:

$$
\begin{aligned}
&\|\frac{2^{i+1}-\|a\|}{2^i}s_{i+1}(a)-\frac{\|b\|-2^{i-1}}{2^{i-1}}s_{i+1}(b)\| \\
&=\|\frac{2^i-(\|a\|-2^i)}{2^i}s_{i+1}(a)+\frac{(2^i-\|b\|)-2^{i-1}}{2^{i-1}}s_{i+1}(b)\| \\
&=\|s_{i+1}(a)-s_{i+1}(b)-\frac{\|a\|-2^i}{2^i}s_{i+1}(a)+\frac{2^i-\|b\|}{2^{i-1}}s_{i+1}(b)\| \qquad (3.10)\\
&\le 2\|a-b\|+4(\|a\|-2^i)+4(2^i-\|b\|) \\
&\le 6\|a-b\|.
\end{aligned}
$$

These along with the fact that τ is 1-Lipschitz imply that $\widehat{\varphi}$ is Lipschitz.

To estimate from below, we use (3.10) and get

$$
\begin{aligned}
\|\varphi(a)-\varphi(b)\| &\ge \|s_{i+1}(a)-s_{i+1}(b)-\frac{\|a\|-2^i}{2^i}s_{i+1}(a)+\frac{2^i-\|b\|}{2^{i-1}}s_{i+1}(b)\| \\
&\quad -\frac{2^i-\|b\|}{2^{i-1}}\|s_i(b)\|-\frac{\|a\|-2^i}{2^i}\|s_{i+2}(a)\| \\
&\ge \|s_{i+1}(a)-s_{i+1}(b)\|-8(\|a\|-\|b\|) \\
&\ge \|a-b\|-8(\|a\|-\|b\|). \qquad (3.11)
\end{aligned}
$$

At this step, if $\|a\|-\|b\|<\frac{1}{10}\|a-b\|$ then

$$
\begin{aligned}
\|\widehat{\varphi}(a)-\widehat{\varphi}(b)\| &\ge \|\varphi(a)-\varphi(b)\|-\|\tau(\|a\|)-\tau(\|b\|)\| \\
&\ge \|a-b\|-9(\|a\|-\|b\|)\ge\frac{1}{10}\|a-b\|.
\end{aligned}
$$

If $\|a\|-\|b\|\ge\frac{1}{10}\|a-b\|$, then by the claim we have

$$\|\widehat{\varphi}(a)-\widehat{\varphi}(b)\|\ge\|a\|-\|b\|\ge\frac{1}{10}\|a-b\|.$$

Therefore $\widehat{\varphi}$ is a Lipschitz embedding.

Case 1. $2^{i-1}\le\|b\|\le\|a\|<2^i$

In this case

$$
\begin{aligned}
\varphi(a)-\varphi(b)&=\frac{2^i-\|a\|}{2^{i-1}}(s_i(a)-s_i(b))+\frac{\|a\|-2^{i-1}}{2^{i-1}}(s_{i+1}(a)-s_{i+1}(b)) \\
&\quad+\frac{\|b\|-\|a\|}{2^{i-1}}s_i(b)+\frac{\|a\|-\|b\|}{2^{i-1}}s_{i+1}(b),
\end{aligned}
$$

so by (3.1) we have

$$
\begin{aligned}
\|\varphi(a)-\varphi(b)\|&\le\frac{2^i-\|a\|}{2^{i-1}}2\|a-b\|+\frac{\|a\|-2^{i-1}}{2^{i-1}}2\|a-b\| \\
&\quad+4(\|a\|-\|b\|)+4(\|a\|-\|b\|)\le 10\|a-b\|,
\end{aligned}
$$

and hence

$$\begin{aligned}\|\widehat{\varphi}(a)-\widehat{\varphi}(b)\| &\le \|\varphi(a)-\varphi(b)\|+\|\tau(\|a\|)-\tau(\|b\|)\| \\ &\le 10\|a-b\|+(\|a\|-\|b\|)\le 11\|a-b\|.\end{aligned}$$

In order to estimate $\|\widehat{\varphi}(a)-\widehat{\varphi}(b)\|$ from below, a subsequence of the maps $(s_i)_{i=1}^{\infty}$ with good behavior is needed, and hence we have to slightly modify the definition of φ by changing the index in (3.2). The technique used here dates back to [6] by Kadets and Pełczyński.

First we may assume that X is separable, otherwise we can simply replace X by $\overline{\operatorname{span}}(\bigcup_{i=1}^{\infty} s_i(N_i))$. Let $(x_n)_{n=1}^{\infty}$ be a sequence of nonzero vectors which is dense in X. By Hahn-Banach theorem we can pick $x_n^* \in S_{X^*}$ so that $x_n^*(x_n)=\|x_n\|$ for each n. Then it is easy to check that the sequence $(x_n^*)_{n=1}^{\infty}$ is norming in X, meaning that $\|x\|=\sup_n |x_n^*(x)|$ for all $x\in X$. Let M be the closed subpace generately by $(x_n^*)_{n=1}^{\infty}$, then the natural embedding from X into M^* is a linear isometric embedding, so we may identify X with its image under this embedding. The selection of subsequences of $(s_i)_{i=1}^{\infty}$ is presented in the following two steps.

Step 1. Since every closed ball in M^* is compact and metrizable in the weak* topology, and also note that the sets N_j's are finite and increasing, we can choose a subsequence (still denoted by $(s_i)_{i=1}^{\infty}$) such that for each j the sequence $(s_i(a))_{i=j}^{\infty}$ is weak*-convergent for all $a\in N_j$. Denote the weak*-limit of this sequence by $m(a)$.

Step 2. We choose a sequence $(k_j)\subset \mathbb{N}$ by induction as follows:

First choose k_1 such that for each pair $a,b\in N_1$ with $m(a)\neq m(b)$,

$$|f(s_n(a)-s_n(b)-(m(a)-m(b)))|\le \frac{1}{100}\|m(a)-m(b)\|$$

for all $n\ge k_1$, where $f=f_{a,b}$ is a fixed element in S_M so that

$$f(m(a)-m(b))\ge \frac{99}{100}\|m(a)-m(b)\|.$$

This can be achieved because N_1 is finite and $(s_i(a)-s_i(b))_{i=1}^{\infty}$ converges to $m(a)-m(b)$ in the weak* topology.

Suppose that k_j has been chosen, we pick $q_j>k_j$ such that for each pair $a,b\in N_j$ satisfying $s_{k_j}(a)-s_{k_j}(b)-(m(a)-m(b))\neq 0$,

$$|g(s_n(a)-s_n(b)-(m(a)-m(b)))|\le \frac{1}{1000}\|a-b\| \tag{3.12}$$

for all $n\ge q_j$, where $g=g_{k_j,a,b}$ is a fixed element in S_M so that

$$g(s_{k_j}(a)-s_{k_j}(b)-(m(a)-m(b)))\ge \frac{99}{100}\|s_{k_j}(a)-s_{k_j}(b)-(m(a)-m(b))\|. \tag{3.13}$$

This can be achieved because N_j is finite and $(s_i(a)-s_i(b))_{i=j}^{\infty}$ converges to $m(a)-m(b)$ in the weak* topology.

Now choose $k_{j+1} \geq q_j$ such that for each pair $a, b \in N_{j+1}$ with $m(a) \neq m(b)$,

$$|f(s_n(a) - s_n(b) - (m(a) - m(b)))| \leq \frac{1}{100}\|m(a) - m(b)\| \tag{3.14}$$

for all $n \geq k_{j+1}$, where $f = f_{a,b}$ is a fixed element in S_M so that

$$f(m(a) - m(b)) \geq \frac{99}{100}\|m(a) - m(b)\|. \tag{3.15}$$

This can be achieved because N_{j+1} is finite and $(s_i(a) - s_i(b))_{i=j+1}^{\infty}$ converges to $m(a) - m(b)$ in the weak* topology.

We redefine φ in the following way: for $a \in N$ with $2^{i-1} \leq \|a\| < 2^i$,

$$\varphi(a) = \frac{2^i - \|a\|}{2^{i-1}} s_{k_i}(a) + \frac{\|a\| - 2^{i-1}}{2^{i-1}} s_{k_{i+1}}(a).$$

To estimate $\|\widehat{\varphi}(a) - \widehat{\varphi}(b)\|$ from below, it suffices to estimate $\|\varphi(a) - \varphi(b)\|$ and get an inequality of the form as (3.11), which along with the claim will allow us to consider two subcases separately and complete the argument, just as shown right after (3.11). We write

$$\begin{aligned}\varphi(a) - \varphi(b) = m(a) - m(b) &+ \frac{2^i - \|a\|}{2^{i-1}}(s_{k_i}(a) - s_{k_i}(b) - (m(a) - m(b))) \\ &+ \frac{\|a\| - 2^{i-1}}{2^{i-1}}(s_{k_{i+1}}(a) - s_{k_{i+1}}(b) - (m(a) - m(b))) \\ &+ \frac{\|b\| - \|a\|}{2^{i-1}} s_{k_i}(b) + \frac{\|a\| - \|b\|}{2^{i-1}} s_{k_{i+1}}(b).\end{aligned} \tag{3.16}$$

First we consider the case when $\|m(a) - m(b)\| \geq \frac{1}{100}\|a - b\|$. By (3.14) and (3.15) we have

$$\begin{aligned}\|\varphi(a) - \varphi(b)\| &\geq f_{a,b}(\varphi(a) - \varphi(b)) \\ &= f_{a,b}(m(a) - m(b)) + f_{a,b}\left(\frac{2^i - \|a\|}{2^{i-1}}(s_{k_i}(a) - s_{k_i}(b) - (m(a) - m(b)))\right) \\ &\quad + f_{a,b}\left(\frac{\|a\| - 2^{i-1}}{2^{i-1}}(s_{k_{i+1}}(a) - s_{k_{i+1}}(b) - (m(a) - m(b)))\right) \\ &\quad + \frac{\|b\| - \|a\|}{2^{i-1}} f_{a,b}(s_{k_i}(b)) + \frac{\|a\| - \|b\|}{2^{i-1}} f_{a,b}(s_{k_{i+1}}(b)) \\ &\geq \frac{99}{100}\|m(a) - m(b)\| - \frac{1}{100}\|m(a) - m(b)\| - 8(\|a\| - \|b\|) \\ &\geq \frac{98}{10000}\|a - b\| - 8(\|a\| - \|b\|).\end{aligned}$$

For the case when $\|m(a) - m(b)\| < \frac{1}{100}\|a - b\|$, we separate into two subcases:

$$\frac{2^i - \|a\|}{2^{i-1}}\|s_{k_i}(a) - s_{k_i}(b) - (m(a) - m(b))\| \geq \frac{1}{10}\|a - b\|, \tag{3.17}$$

$$\frac{2^i - \|a\|}{2^{i-1}}\|s_{k_i}(a) - s_{k_i}(b) - (m(a) - m(b))\| < \frac{1}{10}\|a - b\|. \tag{3.18}$$

In the case (3.17) we use (3.12) and (3.13) and get

$$\begin{aligned}
\|\varphi(a)-\varphi(b)\| &\geq g_{k_i,a,b}(\varphi(a)-\varphi(b)) \\
&= g_{k_i,a,b}(m(a)-m(b)) + g_{k_i,a,b}\left(\frac{2^i-\|a\|}{2^{i-1}}(s_{k_i}(a)-s_{k_i}(b)-(m(a)-m(b)))\right) \\
&\quad + g_{k_i,a,b}\left(\frac{\|a\|-2^{i-1}}{2^{i-1}}(s_{k_{i+1}}(a)-s_{k_{i+1}}(b)-(m(a)-m(b)))\right) \\
&\quad + \frac{\|b\|-\|a\|}{2^{i-1}}g_{k_i,a,b}(s_{k_i}(b)) + \frac{\|a\|-\|b\|}{2^{i-1}}g_{k_i,a,b}(s_{k_{i+1}}(b)) \\
&\geq -\frac{1}{100}\|a-b\| + \frac{99}{1000}\|a-b\| - \frac{1}{1000}\|a-b\| - 8(\|a\|-\|b\|) \\
&= \frac{88}{1000}\|a-b\| - 8(\|a\|-\|b\|).
\end{aligned}$$

On the other hand, in the case (3.18) we have

$$\begin{aligned}
\frac{1}{10}\|a-b\| &> \frac{2^i-\|a\|}{2^{i-1}}(\|s_{k_i}(a)-s_{k_i}(b)\| - \|(m(a)-m(b))\|) \\
&\geq \frac{2^i-\|a\|}{2^{i-1}}\left(\|a-b\| - \frac{1}{100}\|a-b\|\right),
\end{aligned}$$

which implies that $\frac{2^i-\|a\|}{2^{i-1}} < \frac{10}{99}$, and hence $\frac{\|a\|-2^{i-1}}{2^{i-1}} > \frac{89}{99}$. Apply triangle inequality to (3.16) we get

$$\begin{aligned}
\|\varphi(a)-\varphi(b)\| &\geq \frac{\|a\|-2^{i-1}}{2^{i-1}}\|s_{k_{i+1}}(a)-s_{k_{i+1}}(b)-(m(a)-m(b))\| \\
&\quad - \frac{2^i-\|a\|}{2^{i-1}}\|s_{k_i}(a)-s_{k_i}(b)-(m(a)-m(b))\| \\
&\quad - \|m(a)-m(b)\| - 8(\|a\|-\|b\|) \\
&> \frac{89}{99}\left(\|a-b\| - \frac{1}{100}\|a-b\|\right) - \frac{1}{10}\|a-b\| - \frac{1}{100}\|a-b\| - 8(\|a\|-\|b\|) \\
&= \frac{78}{100}\|a-b\| - 8(\|a\|-\|b\|).
\end{aligned}$$

$\square$

References

1. F. Baudier, *Metrical characterization of super-reflexivity and linear type of Banach spaces*, Arch. Math. (Basel) **89** (2007), no. 5, 419–429.
2. F. Baudier and G. Lancien, *Embeddings of locally finite metric spaces into Banach spaces*, Proc. Amer. Math. Soc. **136** (2008), no. 3, 1029–1033.
3. J. Bourgain, *The metrical interpretation of super-reflexivity in Banach spaces*, Israel J. Math. **56** (1986), 221–230.
4. M. Fréchet, *Les dimensions dun ensemble abstrait*, Math. Ann. **68** (1910), no. 2, 145–168.

5. S. Heinrich, *Ultraproducts in Banach space theory*, J. Reine Angew. Math. **313** (1980), 72–104.
6. M. I. Kadets and A. Pełczyński, *Basic sequences, biorthogonal systems and norming sets in Banach and Fréchet spaces*, Studia Math. **25** (1965), 297–323.
7. M. I. Ostrovskii, *Coarse embeddability into Banach spaces*, Topology Proc. **33** (2009), 163–183.
8. M. I. Ostrovskii, *Embeddability of locally finite metric spaces into Banach spaces is finitely determined*, Proc. Amer. Math. Soc. **140** (2012), no. 8, 2721–2730.

Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA

E-mail address: z1986s@math.tamu.edu

ON ASSOUAD'S EMBEDDING TECHNIQUE

MICHAL KRAUS

Abstract. We survey the standard proof of a theorem of Assouad stating that every snowflaked version of a doubling metric space admits a bi-Lipschitz embedding into $\mathbb{R}^n$ for some $n \in \mathbb{N}$.

1. Introduction

Let $(X, d_X), (Y, d_Y)$ be metric spaces and $f: X \to Y$ be a mapping. Then f is called a *bi-Lipschitz embedding* if there are constants $A, B > 0$ such that

$$Ad_X(x, y) \leq d_Y(f(x), f(y)) \leq Bd_X(x, y) \text{ for all } x, y \in X. \tag{1}$$

If f is a bi-Lipschitz embedding, then the *distortion* of f is defined to be the infimum of $\frac{B}{A}$ over all constants $A, B > 0$ for which (1) holds.

A metric space (X, d) is said to have a *doubling constant* $K \geq 1$ if for every $r > 0$ every closed ball in X of radius r can be covered by at most K closed balls of radius $\frac{r}{2}$. By a closed ball of radius r we mean a set of the form $B(x, r) = \{y \in X : d(y, x) \leq r\}$, where $x \in X$ is the center of $B(x, r)$. The space X is called *doubling* if it has a doubling constant K for some $K \geq 1$. Note that doubling metric spaces are separable.

If (X, d) is a metric space and $\alpha \in (0, 1)$, then d^α is clearly also a metric on X and the space (X, d^α) is called the *α-snowflaked version* of (X, d). Note that (X, d) is doubling if and only if (X, d^α) is doubling (possibly with a different doubling constant).

An important open problem in embedding theory is to characterize intrinsically those metric spaces that admit a bi-Lipschitz embedding into $\mathbb{R}^n$ for some $n \in \mathbb{N}$ (we will always consider the Euclidean norm and metric on $\mathbb{R}^n$). It is easy to see that if a metric space admits a bi-Lipschitz embedding into $\mathbb{R}^n$ for some $n \in \mathbb{N}$, then it must be doubling. It is known that the converse does not hold. For example, the 3-dimensional Heisenberg group with its Carnot metric is doubling but does not admit a bi-Lipschitz embedding into $\mathbb{R}^n$ for any $n \in \mathbb{N}$ (see [Se, Theorem 7.1]). However, Assouad [As, Proposition 2.6] proved the following fundamental theorem.

Theorem 1.1 (Assouad, 1983)**.** *Let (X, d) be a doubling metric space and $\alpha \in (0, 1)$. Then (X, d^α) admits a bi-Lipschitz embedding into $\mathbb{R}^n$ for some $n \in \mathbb{N}$.*

Assouad's proof of Theorem 1.1, which we will present here, actually gives the following stronger quantitative statement.

Theorem 1.2 (Quantitative version of Assouad's theorem)**.** *For every $K \geq 1$ and $\alpha \in (0, 1)$, there is an $N = N(K, \alpha) \in \mathbb{N}$ and $D = D(K, \alpha) \geq 1$ such that for every metric space (X, d) with a doubling constant K, the space (X, d^α) admits a bi-Lipschitz embedding into $\mathbb{R}^N$ with distortion at most D.*

Let us mention that in the original paper of Assouad [As], Theorem 1.1 is stated for metric spaces of finite Assouad dimension instead of for doubling metric spaces. Let (X, d) be a metric space. The *Assouad dimension* of X (called the *metric*

Date: October 15, 2014.

dimension in [As]) is the infimum of those $\beta \geq 0$ for which there is $C > 0$ such that for every $0 < a < b$, for every $Y \subset X$ such that $d(x,y) > a$ whenever $x, y \in Y, x \neq y$, and for every $Z \subset X$ such that $\operatorname{diam} Z \leq b$, we have $|Y \cap Z| \leq C \left(\frac{b}{a}\right)^{\beta}$ (if M is a set, we denote by $|M|$ the cardinality of M). However, it is not difficult to prove that X is of finite Assouad dimension if and only if it is doubling. It is actually the value of the doubling constant, and not so much the Assouad dimension, that is relevant to the proof of Assouad's theorem, and so it seems more natural to state the theorem in the present form.

The purpose of this survey is to present in detail the proof of Theorem 1.2. In Section 2 we recall the notion of the tensor product of Hilbert spaces, which will be used in the proof. The proof itself is presented in Section 3. In Section 4, we discuss some questions concerning the dimension of the receiving space $\mathbb{R}^N$ in Theorem 1.2.

2. Tensor products of Hilbert spaces

In this section, we briefly recall the notion of the tensor product of Hilbert spaces, which will be used in the proof of Theorem 1.2. Those who are familiar with tensor products can skip this section.

Let us first describe the algebraic tensor product of linear spaces. Let V, W be linear spaces over $\mathbb{R}$. We denote by $\Lambda(V \times W)$ the set of all formal finite linear combinations of members of the Cartesian product $V \times W$, that is, the set of all expressions of the form $\sum_{i=1}^n a_i(e_i, f_i)$, where $a_i \in \mathbb{R}, e_i \in V, f_i \in W, i = 1, \dots, n$, and $n \in \mathbb{N}$. We identify $\sum_{i=1}^n a_i(e_i, f_i)$ and $\sum_{i=1}^n a_{\pi(i)}(e_{\pi(i)}, f_{\pi(i)})$ for any permutation π of $\{1, \dots, n\}$ and we also identify $\sum_{i=1}^{n+1} a_i(e_i, f_i)$ and $\sum_{i=1}^n a_i(e_i, f_i)$ if $a_{n+1} = 0$. We make $\Lambda(V \times W)$ into a linear space by defining

$$a\left(\sum_{i=1}^n a_i(e_i, f_i)\right) = \sum_{i=1}^n a a_i(e_i, f_i)$$

and

$$\sum_{i=1}^n a_i(e_i, f_i) + \sum_{i=1}^n b_i(e_i, f_i) = \sum_{i=1}^n (a_i + b_i)(e_i, f_i).$$

Furthermore, we denote by $\Lambda_0(V \times W)$ the linear subspace of $\Lambda(V \times W)$ generated by the elements of the form

$$(a_1 e_1 + a_2 e_2, b_1 f_1 + b_2 f_2) - \sum_{i,j=1}^2 a_i b_j (e_i, f_j).$$

The *algebraic tensor product* of V and W, denoted by $V \otimes W$, is the linear quotient space $\Lambda(V \times W)/\Lambda_0(V \times W)$. Elements of $V \otimes W$ are called *tensors*. We denote by $e \otimes f$ the tensor containing (e, f), that is, the equivalence class $(e, f) + \Lambda_0(V \times W)$. So any tensor from $V \otimes W$ can be written as $\sum_{i=1}^n a_i e_i \otimes f_i$, where $a_i \in \mathbb{R}, e_i \in V, f_i \in W$ and $n \in \mathbb{N}$. The purpose of taking the quotient is that now we have

$$(a_1 e_1 + a_2 e_2) \otimes (b_1 f_1 + b_2 f_2) = \sum_{i,j=1}^2 a_i b_j e_i \otimes f_j.$$

It is not hard to show that if $(e_i)_{i \in \Gamma_1}$ is a basis of V and $(f_j)_{j \in \Gamma_2}$ is a basis of W, then $(e_i \otimes f_j)_{(i,j) \in \Gamma_1 \times \Gamma_2}$ is a basis of $V \otimes W$. In particular, if V and W are finite dimensional, then $\dim(V \otimes W) = \dim V \dim W$.

Now, let H_1, H_2 be real Hilbert spaces with inner products $\langle .,. \rangle_1, \langle .,. \rangle_2$ and norms $\|.\|_1, \|.\|_2$ respectively. We define an inner product on the algebraic tensor product $H_1 \otimes H_2$ by setting $\langle e_1 \otimes f_1, e_2 \otimes f_2 \rangle = \langle e_1, e_2 \rangle_1 \langle f_1, f_2 \rangle_2$ for all $e_1, e_2 \in H_1, f_1, f_2 \in H_2$, and by extending bilinearly to all of $H_1 \otimes H_2$. It is of course necessary to check that $\langle .,. \rangle$ is well-defined and that it is indeed an inner product. As usual, the inner product $\langle .,. \rangle$ gives rise to a norm on $H_1 \otimes H_2$ defined by $\|x\| = \sqrt{\langle x, x \rangle}$ for $x \in H_1 \otimes H_2$. The completion of $H_1 \otimes H_2$ under this norm, which is of course a Hilbert space, is called the *tensor product* of H_1 and H_2 and is also denoted by $H_1 \otimes H_2$ (from now on, we will always use this symbol for tensor products of Hilbert spaces, so no confusion should arise). Note that for any $e \in H_1, f \in H_2$ we have $\|e \otimes f\| = \|e\|_1 \|f\|_2$. Note also that if H_1 and H_2 are finite dimensional, then their algebraic tensor product is also finite dimensional and so the completion leaves the space unchanged. Hence in this case $\dim(H_1 \otimes H_2) = \dim H_1 \dim H_2$.

3. Proof of Theorem 1.2

Let us prove Theorem 1.2. We will basically follow the lines of the original proof of Assouad [As] (see also [He, Theorem 12.2] for an exposition in English). We will use the following lemma.

Lemma 3.1. *Let $\alpha, \tau \in (0,1)$, $A, B > 0$ and $m \in \mathbb{N}$. Then there is an $N \in \mathbb{N}$ and $D \geq 1$ such that if (X, d) is a metric space and there are mappings $\varphi_i \colon X \to \mathbb{R}^m$, $i \in \mathbb{Z}$, satisfying*

(1) $\|\varphi_i(s) - \varphi_i(t)\| \geq A$ *if* $\tau^{i+1} < d(s,t) \leq \tau^i$,
(2) $\|\varphi_i(s) - \varphi_i(t)\| \leq B \min\{\tau^{-i} d(s,t), 1\}$ *for all* $s, t \in X$,

then (X, d^α) admits a bi-Lipschitz embedding into $\mathbb{R}^N$ with distortion at most D.

Proof. Let (X, d) be a metric space and suppose that there are mappings $\varphi_i \colon X \to \mathbb{R}^m$, $i \in \mathbb{Z}$, satisfying the conditions (1) and (2). We will also work with the space $\mathbb{R}^{2n}$, where $n \in \mathbb{N}$ will be chosen later. Let $e_1, \ldots, e_{2n}$ be an orthonormal basis of $\mathbb{R}^{2n}$ (for example the canonical basis) and extend the sequence (e_i) $2n$-periodically to all of $\mathbb{Z}$, that is, $e_{i+2n} = e_i$ for every $i \in \mathbb{Z}$. Also, fix an arbitrary $s_0 \in X$.

We define a mapping $f \colon X \to \mathbb{R}^m \otimes \mathbb{R}^{2n}$ by

$$f(s) = \sum_{i \in \mathbb{Z}} \tau^{i\alpha} (\varphi_i(s) - \varphi_i(s_0)) \otimes e_i.$$

(By $\mathbb{R}^m \otimes \mathbb{R}^{2n}$ we mean the tensor product of the Hilbert spaces $\mathbb{R}^m$ and $\mathbb{R}^{2n}$ as described in Section 2. It is linearly isometric to $\mathbb{R}^{2mn}$.) The convergence of the series will follow from the first estimate bellow. Let us show that for large enough n the mapping f is a bi-Lipschitz embedding of (X, d^α) into $\mathbb{R}^m \otimes \mathbb{R}^{2n}$.

Let $s, t \in X, s \neq t$, and let $k \in \mathbb{Z}$ be such that $\tau^{k+1} < d(s,t) \leq \tau^k$. Let us first estimate $\|f(s) - f(t)\|$ from above. We have

$$\begin{aligned}
\|f(s) - f(t)\| &\leq \sum_{i>k} \tau^{i\alpha} \|\varphi_i(s) - \varphi_i(t)\| + \sum_{i \leq k} \tau^{i\alpha} \|\varphi_i(s) - \varphi_i(t)\| \\
&\leq \sum_{i>k} \tau^{i\alpha} B + \sum_{i \leq k} \tau^{i\alpha} B \tau^{-i} d(s,t) \\
&= B\tau^{(k+1)\alpha} \sum_{i=0}^{\infty} \tau^{i\alpha} + B d(s,t) \tau^{k(\alpha-1)} \sum_{i=0}^{\infty} \tau^{i(1-\alpha)} \\
&= B\tau^{(k+1)\alpha} \frac{1}{1-\tau^\alpha} + B d(s,t) \tau^{k(\alpha-1)} \frac{1}{1-\tau^{1-\alpha}} \\
&\leq B d(s,t)^\alpha \frac{1}{1-\tau^\alpha} + B d(s,t) d(s,t)^{\alpha-1} \frac{1}{1-\tau^{1-\alpha}} \\
&= B \left(\frac{1}{1-\tau^\alpha} + \frac{1}{1-\tau^{1-\alpha}} \right) d(s,t)^\alpha.
\end{aligned}$$

Note that no restriction on n was needed in this estimate.

Now, let us estimate $\|f(s)-f(t)\|$ from bellow. We have

$$\|f(s)-f(t)\| \geq \left\| \sum_{k-n<i\leq k+n} \tau^{i\alpha}(\varphi_i(s)-\varphi_i(t))\otimes e_i \right\| - \sum_{i>k+n} \tau^{i\alpha}\|\varphi_i(s)-\varphi_i(t)\| - \sum_{i\leq k-n} \tau^{i\alpha}\|\varphi_i(s)-\varphi_i(t)\|.$$

For the first sum we have

$$\left\| \sum_{k-n<i\leq k+n} \tau^{i\alpha}(\varphi_i(s)-\varphi_i(t))\otimes e_i \right\| \geq \tau^{k\alpha}\|\varphi_k(s)-\varphi_k(t)\| \geq \tau^{k\alpha}A \geq d(s,t)^\alpha A,$$

where the first inequality holds since the summands on the left hand side are mutually orthogonal. The second sum satisfies

$$\sum_{i>k+n} \tau^{i\alpha}\|\varphi_i(s)-\varphi_i(t)\| \leq \sum_{i>k+n} \tau^{i\alpha}B = B\tau^{(k+n+1)\alpha}\sum_{i=0}^{\infty}\tau^{i\alpha} = B\tau^{(k+n+1)\alpha}\frac{1}{1-\tau^\alpha} \leq Bd(s,t)^\alpha \frac{\tau^{n\alpha}}{1-\tau^\alpha},$$

and for the last sum we have

$$\sum_{i\leq k-n} \tau^{i\alpha}\|\varphi_i(s)-\varphi_i(t)\| \leq \sum_{i\leq k-n} \tau^{i\alpha}B\tau^{-i}d(s,t) = Bd(s,t)\tau^{(k-n)(\alpha-1)}\sum_{i=0}^{\infty}\tau^{i(1-\alpha)} = Bd(s,t)\tau^{(k-n)(\alpha-1)}\frac{1}{1-\tau^{1-\alpha}} \leq Bd(s,t)^\alpha\frac{\tau^{n(1-\alpha)}}{1-\tau^{1-\alpha}}.$$

Hence we obtain

$$\|f(s)-f(t)\| \geq \left(A - B\left(\frac{\tau^{n\alpha}}{1-\tau^\alpha} + \frac{\tau^{n(1-\alpha)}}{1-\tau^{1-\alpha}}\right)\right) d(s,t)^\alpha.$$

Now if n is large enough so that the constant on the right hand side is positive (which depends only on α, τ, A and B), then the mapping f is a bi-Lipschitz embedding of (X, d^α) into $\mathbb{R}^m \otimes \mathbb{R}^{2n}$ and both the dimension of the target space and the distortion of f depend only on α, τ, A, B and m. □

Proof of Theorem 1.2. Let $K \geq 1$ and fix an arbitrary $\tau \in (0,1)$. Let (X,d) be a metric space with a doubling constant K and let $i \in \mathbb{Z}$. We will construct a mapping $\varphi = \varphi_i \colon X \to \mathbb{R}^m$ for some $m \in \mathbb{N}$ such that the conditions (1) and (2) in Lemma 3.1 will be satisfied for some $A, B > 0$, and A, B and m will depend only on K and our choice of τ. Lemma 3.1 will then complete the proof of Theorem 1.2.

Let $c = \frac{1}{4}\tau^{i+1}$ and take a c-net Y in X. By a c-net we mean a maximal subset of X such that all pairs of its distinct points have distance at least c. By Zorn's lemma, such a set exists. It is then clear that for every $y \in Y$ we have

$$\left|\left\{ z \in Y : d(z,y) \leq \left(\frac{4}{\tau}+4\right)c \right\}\right| \leq m,$$

where $m \in \mathbb{N}$ depends only on the doubling constant K and the choice of τ (we can take any $m \geq K^{2+\log_2\left(\frac{4}{\tau}+4\right)}$). Let $k \colon Y \to \{1,\dots,m\}$ be an $(m, (\frac{4}{\tau}+4)c)$-colouring of Y, that is, $k(y) \neq k(y')$ if $y, y' \in Y, y \neq y'$, and $d(y,y') \leq (\frac{4}{\tau}+4)c$. Such a mapping clearly exists. Indeed, since Y is clearly countable, we can make it into a sequence (y_j) and define $k(y_j)$ inductively by choosing a value from $\{1,\dots,m\}$ not taken by those y_l for $l < j$ for which $d(y_l, y_j) \leq (\frac{4}{\tau}+4)c$. Since there are at most $m-1$ such y_l, this is always possible.

Let $e_1, \dots, e_m$ be an orthonormal basis of $\mathbb{R}^m$. We define $\varphi\colon X \to \mathbb{R}^m$ by

$$\varphi(s) = \sum_{y \in Y} g_y(s) e_{k(y)},$$

where

$$g_y(s) = \frac{1}{2c} \max\{2c - d(s,y), 0\}.$$

Let us verify that φ is the desired mapping.

First, if $s \in X$, let $B_s = \{y \in Y : g_y(s) \neq 0\} = \{y \in Y : d(y,s) < 2c\}$. Then clearly $|B_s| \leq m$. This in particular shows that the sum in the definition of $\varphi(s)$ is in fact finite, hence convergent. Let $s, t \in X$. It is clear that for every $y \in Y$ we have

$$|g_y(s) - g_y(t)| \leq \frac{1}{2c} d(s,t) = \frac{2}{\tau} \tau^{-i} d(s,t),$$

and therefore

$$\|\varphi(s) - \varphi(t)\| \leq \sum_{y \in B_s \cup B_t} |g_y(s) - g_y(t)| \leq 2m \frac{2}{\tau} \tau^{-i} d(s,t) = \frac{4m}{\tau} \tau^{-i} d(s,t).$$

Furthermore, we have

$$\|\varphi(s) - \varphi(t)\| \leq \|\varphi(s)\| + \|\varphi(t)\| = \left\| \sum_{y \in B_s} g_y(s) e_{k(y)} \right\| + \left\| \sum_{y \in B_t} g_y(t) e_{k(y)} \right\| \\ \leq 2m \leq \frac{4m}{\tau}.$$

Hence

$$\|\varphi(s) - \varphi(t)\| \leq \frac{4m}{\tau} \min\{\tau^{-i} d(s,t), 1\},$$

and therefore the condition (2) in Lemma 3.1 is satisfied with $B = \frac{4m}{\tau}$.

Now, let $s, t \in X$ be such that $4c = \tau^{i+1} < d(s,t) \leq \tau^i = \frac{4}{\tau} c$. Then $B_s \cap B_t = \emptyset$ and the vectors $e_{k(y)}$ for $y \in B_s \cup B_t$ are mutually orthogonal, and therefore

$$\|\varphi(s) - \varphi(t)\|^2 = \sum_{y \in B_s} |g_y(s)|^2 + \sum_{y \in B_t} |g_y(t)|^2.$$

Since Y is a c-net, there is a $y \in Y$ such that $d(y,s) < c$. Then $g_y(s) \geq \frac{1}{2}$, and therefore $\|\varphi(s) - \varphi(t)\| \geq \frac{1}{2}$. Hence the condition (1) in Lemma 3.1 is satisfied with $A = \frac{1}{2}$ and the proof is complete. □

4. The dimension of the receiving Euclidean space

Let $K \geq 2$ and $\alpha \in (0,1)$ be fixed. Let us inspect the above proof of Theorem 1.2 to see how large the dimension $N(K, \alpha)$ it gives. We are interested in an estimate from below. At the beginning of the proof we choose an arbitrary $\tau \in (0,1)$. Then we take $m \in \mathbb{N}$ such that $m \geq K^{2 + \log_2\left(\frac{4}{\tau} + 4\right)}$, and $A = \frac{1}{2}$ and $B = \frac{4m}{\tau}$. The dimension $N(K, \alpha)$ is then equal to $2mn$, where $n \in \mathbb{N}$ is such that

$$\frac{\tau^{n\alpha}}{1 - \tau^\alpha} + \frac{\tau^{n(1-\alpha)}}{1 - \tau^{1-\alpha}} < \frac{A}{B} = \frac{\tau}{8m}.$$

Then we must have

$$\frac{\tau^{n\alpha}}{1 - \tau^\alpha} < \frac{\tau}{8m},$$

and therefore

$$\begin{aligned} n &> \frac{\log_2\left(\frac{\tau(1-\tau^\alpha)}{8m}\right)}{\alpha\log_2\tau} = \frac{\log_2\left(\frac{8m}{\tau(1-\tau^\alpha)}\right)}{\alpha\log_2\frac{1}{\tau}} = \frac{\log_2 8+\log_2 m-\log_2(\tau(1-\tau^\alpha))}{\alpha\log_2\frac{1}{\tau}} \\ &\geq \frac{\log_2 m}{\alpha\log_2\frac{1}{\tau}} \geq \frac{\log_2 K^{2+\log_2\left(\frac{4}{\tau}+4\right)}}{\alpha\log_2\frac{1}{\tau}} = \frac{\left(2+\log_2\left(\frac{4}{\tau}+4\right)\right)\log_2 K}{\alpha\log_2\frac{1}{\tau}} \\ &\geq \frac{\log_2\frac{1}{\tau}\log_2 K}{\alpha\log_2\frac{1}{\tau}} = \frac{\log_2 K}{\alpha}. \end{aligned}$$

Similarly, we must have

$$\frac{\tau^{n(1-\alpha)}}{1-\tau^{1-\alpha}} < \frac{\tau}{8m},$$

and therefore

$$n > \frac{\log_2 K}{1-\alpha}.$$

Hence

$$n > \log_2 K \max\left\{\frac{1}{\alpha}, \frac{1}{1-\alpha}\right\}.$$

It follows that no matter which $\tau \in (0,1)$ we choose at the beginning of the proof we obtain

$$N(K,\alpha) = 2mn \geq 2K^4\log_2 K \max\left\{\frac{1}{\alpha}, \frac{1}{1-\alpha}\right\}$$

(here we used the fact that $m \geq K^{2+\log_2\left(\frac{4}{\tau}+4\right)} \geq K^{2+\log_2 4} = K^4$). In particular, the construction gives $N(K,\alpha)\to\infty$ as $\alpha\to 0$ and also as $\alpha\to 1$. Is this necessary?

To answer this question, we can start by trying to optimize the constants that come into the construction. For example, it is not clear at first sight whether we can take some $m < K^{2+\log_2\left(\frac{4}{\tau}+4\right)}$ that would work as well. However, let us take a different point of view. In this context, the notion of Assouad dimension introduced after Theorem 1.2 proves useful. Let us denote the Assouad dimension of a metric space (X,d) by $\dim_A(X,d)$. It is not difficult to prove the following facts (see also [As]).

- $\dim_A(\mathbb{R}^n) = n$ for every $n\in\mathbb{N}$.
- $\dim_A(X,d^\alpha) = \frac{1}{\alpha}\dim_A(X,d)$ for every $\alpha\in(0,1)$.
- If (X,d) admits a bi-Lipschitz embedding into a metric space (Y,δ), then $\dim_A(X,d)\leq\dim_A(Y,\delta)$.

It follows that if (X,d) is a doubling metric space and $\alpha\in(0,1)$, then in order to have a bi-Lipschitz embedding of (X,d^α) into $\mathbb{R}^n$ we must have $n\geq\frac{1}{\alpha}\dim_A(X,d)$. In particular, in Theorem 1.2 we must have $N(K,\alpha)\to\infty$ as $\alpha\to 0$ for any $K\geq 2$ (by taking e.g. $X=\mathbb{R}$). However, note that if $\alpha\in(b,1)$ for some $b\in(0,1)$, then this method does not show any obstruction for having a bi-Lipschitz embedding of (X,d^α) into $\mathbb{R}^n$ for some $n\in\mathbb{N}$ independent of α. It turns out that this is not accidental. Indeed, Naor and Neiman [NN, Theorem 1.2] proved the following theorem.

Theorem 4.1 (Naor, Neiman, 2012). *For every $K\geq 1$ there is an $N=N(K)\in\mathbb{N}$ and for every $K\geq 1$ and $\alpha\in(\frac{1}{2},1)$ there is a $D=D(K,\alpha)\geq 1$ such that for every metric space (X,d) with a doubling constant K, the space (X,d^α) admits a bi-Lipschitz embedding into $\mathbb{R}^N$ with distortion at most D.*

Note that the theorem holds true if we replace $\frac{1}{2}$ with any fixed constant $b\in(0,1)$. The point is to have α bounded away from 0. Let us mention that the proof in [NN] gives the estimates

$$N(K) \leq C \log K \text{ and } D(K,\alpha) \leq C \left(\frac{\log K}{1-\alpha} \right)^2,$$

where $C > 0$ is some absolute constant. We will not discuss the proof of Theorem 4.1 here. Let us just say that the proof of Naor and Neiman is probabilistic. Later, David and Snipes [DS] found a non-probabilistic proof of Theorem 4.1 based on the original construction of Assouad.

References

[As] P. Assouad, *Plongements lipschitziens dans* $\mathbb{R}^n$, Bull. Soc. Math. France 111 (1983), no. 4, 429–448.

[DS] G. David and M. Snipes, *A non-probabilistic proof of the Assouad embedding theorem with bounds on the dimension*, Anal. Geom. Metr. Spaces 1 (2013), 36–41.

[He] J. Heinonen, *Lectures on Analysis on Metric Spaces*, Universitext. Springer-Verlag, New York, 2001.

[NN] A. Naor and O. Neiman, *Assouad's theorem with dimension independent of the snowflaking*, Rev. Mat. Iberoam. 28 (2012), no. 4, 1123–1142.

[Se] S. Semmes, *On the nonexistence of bilipschitz parameterizations and geometric problems about A_∞-weights*, Rev. Mat. Iberoam. 12 (1996), no. 2, 337–410.

Institute of Mathematics AS CR, Žitná 25, 115 67 Praha 1, Czech Republic
E-mail address: kraus@math.cas.cz

OBSTRUCTION TO UNIFORM OR COARSE EMBEDDABILITY INTO REFLEXIVE BANACH SPACES

A. DALET

Abstract. This paper is based on the paper [11] of N. J. Kalton. The main result is that c_0 cannot be uniformly or coarsely embedded into a reflexive Banach space. In order to prove it, we will present a Ramsey type argument and Kalton's property $\mathcal{Q}$, which used together permit to rule out coarse or uniform embeddings into reflexive Banach spaces.

1. Introduction

Let (M,d), (N,δ) be metric spaces and $f : M \to N$ be any map. For $t > 0$, define

$$\varphi_f(t) = \inf\{\delta(f(x), f(y));\ d(x,y) \geq t\}$$

and

$$\omega_f(t) = \sup\{\delta(f(x), f(y));\ d(x,y) \leq t\}.$$

The map f is said to be:

- a coarse embedding if $\lim\limits_{t\to+\infty} \varphi_f(t) = +\infty$ and $\omega_f(t) < +\infty$, $\forall t > 0$. Then M coarsely embeds into N.
- a uniform embedding if $\lim\limits_{t\to 0} \omega_f(t) = 0$ and $\varphi_f(t) > 0$, $\forall t > 0$. Then M uniformly embeds into N.
- a strong uniform embedding if f is a coarse and a uniform embedding.
- a Lipschitz embedding if there exist $A, B > 0$ such that for every $x, y \in M$,

$$Ad(x,y) \leq \delta(f(x), f(y)) \leq Bd(x,y).$$

In 1974, Aharoni [1] proved that every separable metric space can be Lipschitz embedded into c_0. There exist quantitative versions of this result due to Assouad [4], Pelant [17] and finally the sharp constant of distortion is 2 and is given by Kalton and Lancien in [13]. It is an open question to know whether there exist other Banach spaces into which every separable metric spaces can be Lipschitz embedded.

This question is equivalent to the following: if c_0 Lipschitz-embeds into a Banach space, does it imply that it linearly embeds into this space? In [10] Kalton proved that there exists a Banach space into which c_0 strong uniformly embeds but does not linearly embed. More precisely, for any non trivial gauge ω and any metric space (M,d), the Lipschitz-free space over $(M, \omega \circ d)$, denoted $\mathcal{F}_\omega(M)$, is a Schur space. Now $\omega : \begin{array}{ccc} \mathbb{R}_+ & \to & \mathbb{R}_+ \\ t & \mapsto & \begin{cases} t^\alpha, & t \leq 1 \\ t, & t \geq 1 \end{cases} \end{array}$ is non trivial, thus $\mathcal{F}_\omega(c_0)$ is a Schur space. Moreover it is easy to see that the identity from $(c_0, \|\cdot\|_\infty)$ to $(c_0, \omega \circ \|\cdot\|_\infty)$ is a strong uniform embedding. It is known from [9] that $(c_0, \omega \circ \|\cdot\|_\infty)$ isometrically embeds into its Lipschitz-free space. Finally, we conclude that c_0 strong uniformly embeds into $\mathcal{F}_\omega(c_0)$, which is a Schur space, hence c_0 cannot be linearly embedded into it.

It was proved independently by Christensen [7], Mankiewicz [15] and Aronszajn [3] in the 70's that if a separable Banach space X Lipschitz embeds into a space Y with the Radon-Nikodym property, the embedding admits a point of Gâteaux-differentiability and one can deduce that X linearly embeds into Y. Thus, because every reflexive space has the RNP, it is not possible to find a reflexive Banach space which is universal for Lipschitz embeddings of separable metric space, but one can ask whether there exists a reflexive Banach space into which every separable metric space could be uniformly or coarsely embedded. Following a paper of Kalton [11] (see also [14] or [8]) we will prove that there exists no reflexive Banach space containing uniformly or coarsely the space c_0. More precisely we will define a property, failed by c_0, and prove that a Banach space failing this property cannot be uniformly or coarsely embedded into a reflexive Banach space. This implies a previous result: Mendel and Naor proved in [16] that c_0 cannot be coarsely embedded into a super-reflexive Banach space. However Baudier obtained in [5] that any Banach space without cotype contains strongly uniformly every proper metric space. In particular $\left(\oplus_{n=1}^{+\infty}\ell_\infty^n\right)_2$, which is reflexive, contains strongly uniformly every proper metric space.

Section 2 is about Ramsey theory and is devoted to the proof of a Ramsey type argument due to Kalton [11]. In section 3 we introduce the $\mathcal{Q}$-property and prove that a Banach space failing it cannot be uniformly or coarsely embedded into a reflexive Banach space. In section 4 it is proved first that a stable Banach space has the $\mathcal{Q}$-property. Then we present a theorem which permits to rule out the $\mathcal{Q}$-property and we use it to prove that the James space J and its dual fail it. To conclude this section, we focus on the space c_0 and prove that it does not have the $\mathcal{Q}$-property. Then we prove a stronger result of Kalton: c_0 cannot be uniformly or coarsely embedded into a Banach space having all its iterated duals separable. Finally in section 5, we compare the structure of the paper [11] with the proof of the fact that $\mathcal{C}[1,\omega_1]$ cannot be uniformly embedded into ℓ_∞ in [12].

2. Preliminaries: Ramsey theory and special graphs

Let $\mathbb{M}$ be an infinite subset of $\mathbb{N}$ and $k \in \mathbb{N}$. The set $G_k(\mathbb{M})$ is the set of all subsets of $\mathbb{M}$ of size k. We will write an element $\overline{n}$ of $G_k(\mathbb{M})$ as follows: $\overline{n} = \{n_1, \dots, n_k\}$, with $n_1 < \dots < n_k$.

First we state Ramsey's theorem (see [18]):

Theorem 2.1. *Let $k, r \in \mathbb{N}$ and $f : G_k(\mathbb{N}) \to \{1, \dots, r\}$ be any map. Then there exists an infinite subset $\mathbb{M}$ of $\mathbb{N}$ and $i \in \{1, \dots, r\}$ such that for every $\overline{n} \in G_k(\mathbb{M})$, $f(\overline{n}) = i$.*

It is not difficult to deduce a topological version of this result.

Corollary 2.2. *Let (K, d) be a compact metric space, $k \in \mathbb{N}$ and $f : G_k(\mathbb{N}) \to K$. Then for every $\varepsilon > 0$, there exists an infinite subset $\mathbb{M}$ of $\mathbb{N}$ such that for every $\overline{n}, \overline{m} \in G_k(\mathbb{M})$, $d(f(\overline{n}), f(\overline{m})) < \varepsilon$.*

We can think about a result as a part of Ramsey theory if for a given coloring of a mathematical object, there exists a sub-object which is monochromatic.

From now we will follow the paper of Kalton [11] (see also [14], [8]). For an infinite subset $\mathbb{M}$ of $\mathbb{N}$, endow the space $G_k(\mathbb{M})$ with the following metric d: two distinct subsets $\overline{n}, \overline{m} \in G_k(\mathbb{M})$ are said to be adjacent ($d(\overline{n}, \overline{m}) = 1$) if

$$n_1 \le m_1 \le n_2 \le \cdots \le n_k \le m_k \text{ or } m_1 \le n_1 \le m_2 \le \cdots \le m_k \le n_k.$$

We will write $\overline{n} < \overline{m}$ when $n_k < m_1$. In this case, $d(\overline{n}, \overline{m}) = k$.

We will start by a Ramsey type result which will be useful to give an obstruction to uniform and coarse embeddability into reflexive Banach spaces. Before to state it we need some tools.

Let X be a Banach space, $k \in \mathbb{N}$, $f : G_k(\mathbb{N}) \to X$ a bounded map and $\mathcal{U}$ a non-principal ultrafilter on $\mathbb{N}$. We define a bounded map $\partial_{\mathcal{U}} f : G_{k-1}(\mathbb{N}) \to X^{**}$ as follows:

$$\forall \overline{n} \in G_{k-1}(\mathbb{N}),\ \partial_{\mathcal{U}} f(\overline{n}) = w^*\text{-}\lim_{n_k \in \mathcal{U}} f(n_1, \ldots, n_{k-1}, n_k).$$

We can iterate this procedure for $1 \le r \le k$: $\partial^r_{\mathcal{U}} f : G_{k-r}(\mathbb{N}) \to X^{(2r)}$, where $X^{(2r)}$ is the $2r$-th dual of X. Then $\partial^k_{\mathcal{U}} f$ is an element of $X^{(2k)}$.

Proposition 2.3. *Let $f : G_k(\mathbb{N}) \to \mathbb{R}$ be a bounded map. Then for every $\varepsilon > 0$, there exists $\mathbb{M}$, an infinite subset of $\mathbb{N}$, such that:*

$$\forall \overline{n} \in G_k(\mathbb{M}), |f(\overline{n}) - \partial^k_{\mathcal{U}} f| < \varepsilon.$$

Proof. Let $\varepsilon > 0$. By induction on $j \in \mathbb{N}$, we will construct $\mathbb{M} = \{m_1, \ldots, m_j, \ldots\}$ such that if $\overline{n} \subset \{m_1, \ldots, m_j\}$ is of size $i \le \min\{j, k\}$, then $|\partial^{k-i}_{\mathcal{U}} f(\overline{n}) - \partial^k_{\mathcal{U}} f| < \varepsilon$:

- Because
$$\partial^k_{\mathcal{U}} f = w^* \text{-} \lim_{n_1 \in \mathcal{U}} \cdots \lim_{n_k \in \mathcal{U}} f(n_1, \ldots, n_k)$$
and for $m \in \mathbb{N}$,
$$\partial^{k-1}_{\mathcal{U}} f(m) = w^* \text{-} \lim_{n_2 \in \mathcal{U}} \cdots \lim_{n_k \in \mathcal{U}} f(m, n_2, \ldots, n_k)$$
we can deduce that there exists $m_1 \in \mathbb{N}$ such that $|\partial^{k-1}_{\mathcal{U}} f(m_1) - \partial^k_{\mathcal{U}} f| < \varepsilon$.
- Assume $m_1 < \cdots < m_j$ chosen.
Let $1 \le i \le \min\{j, k-1\}$ and $\overline{n} = \{n_1, \ldots, n_i\} \subset \{m_1, \ldots, m_j\}$. Then for $m > m_j$,

$$|\partial^{k-(i+1)}_{\mathcal{U}} f(\overline{n} \cup m) - \partial^{k-i}_{\mathcal{U}} f(\overline{n})| \le w^*\text{-} \lim_{n_{i+1} \in \mathcal{U}} \lim_{n_{i+2} \in \mathcal{U}} \cdots \lim_{n_k \in \mathcal{U}} |f(n_1, \ldots, n_i, m, n_{i+2}, \ldots, n_k) \\ - f(n_1, \ldots, n_i, n_{i+1}, n_{i+2}, \ldots, n_k)|$$

Thus there exists $\mathbb{A}_{\overline{n}} \in \mathcal{U}$ such that for every $m \in \mathbb{A}_{\overline{n}}$, $m > m_j$ and

$$w^*\text{-} \lim_{n_{i+1} \in \mathcal{U}} \left(\lim_{n_{i+2} \in \mathcal{U}} \cdots \lim_{n_k \in \mathcal{U}} |f(n_1, \ldots, n_i, m, n_{i+2}, \ldots, n_k) \right. \\ \left. - f(n_1, \ldots, n_i, n_{i+1}, n_{i+2}, \ldots, n_k)| \right) < \varepsilon$$

Moreover the intersection $\mathbb{A}$ of all $\mathbb{A}_{\overline{n}}$ is not empty and belongs to $\mathcal{U}$. Thus pick $m_{j+1} \in \mathbb{A}$.

Then for every $\overline{n} = \{n_1, \ldots, n_i\} \subset \{m_1, \ldots, m_j\}, 1 \le i \le \min\{j, k-1\}$,

$$|\partial^{k-(i+1)}_{\mathcal{U}} f(\overline{n} \cup m_{j+1}) - \partial^k_{\mathcal{U}} f| \le |\partial^{k-(i+1)}_{\mathcal{U}} f(\overline{n} \cup m_{j+1}) - \partial_{\mathcal{U}^{k-i}} f(\overline{n})| + |\partial^{k-i}_{\mathcal{U}} f(\overline{n}) - \partial^k_{\mathcal{U}} f| \\ < 2\varepsilon$$

We deduce the result with $i = k$. □

It is possible to generalize this result to bounded maps which takes values into a Banach space X.

Lemma 2.4. *Let $f : G_k(\mathbb{N}) \to X$ be a bounded map. Then for every $\varepsilon > 0$, there exists $\mathbb{M}$, an infinite subset of $\mathbb{N}$, such that:*

$$\forall \overline{n} \in G_k(\mathbb{M}), \|f(\overline{n})\| < \|\partial_{\mathcal{U}}^k f\| + \omega_f(1) + \varepsilon.$$

Proof. For two bounded maps $f : G_k(\mathbb{N}) \to X$ and $g : G_k(\mathbb{N}) \to X^*$, define $f \otimes g : G_{2k}(\mathbb{N}) \to \mathbb{R}$ by $f \otimes g(\overline{n}) = \langle f(n_2, n_4, \ldots, n_{2k}), g(n_1, n_3, \ldots, n_{2k-1})\rangle$.

Then $\partial_{\mathcal{U}}^2(f \otimes g) = \partial_{\mathcal{U}} f \otimes \partial_{\mathcal{U}} g$. Indeed,

$$\begin{aligned}\partial_{\mathcal{U}}(f \otimes g)(n_1, \ldots, n_{2k-1}) &= \lim_{n_{2k} \in \mathcal{U}} \langle f(n_2, n_4, \ldots, n_{2k}), g(n_1, n_3, \ldots, n_{2k-1})\rangle \\ &= \langle \partial_{\mathcal{U}} f(n_2, \ldots, n_{2k-2}), g(n_1, \ldots, n_{2k-1})\rangle\end{aligned}$$

thus

$$\begin{aligned}\partial_{\mathcal{U}}^2(f \otimes g)(n_1, \ldots, n_{2k-2}) &= \lim_{n_{2k-1} \in \mathcal{U}} \langle \partial_{\mathcal{U}} f(n_2, n_4, \ldots, n_{2k-2}), g(n_1, n_3, \ldots, n_{2k-1})\rangle \\ &= \langle \partial_{\mathcal{U}} f(n_2, \ldots, n_{2k-2}), \partial_{\mathcal{U}} g(n_1, \ldots, n_{2k-3})\rangle \\ &= (\partial_{\mathcal{U}} f \otimes \partial_{\mathcal{U}} g)(n_1, \ldots, n_{2k-2}).\end{aligned}$$

In particular, $\partial_{\mathcal{U}}^{2k}(f \otimes g) = \partial_{\mathcal{U}}^k f \otimes \partial_{\mathcal{U}}^k g$.

Let $f : G_k(\mathbb{N}) \to X$ be a bounded map. Hahn-Banach theorem gives a map g from $G_k(\mathbb{N})$ to X^* such that for every $\overline{n} \in G_k(\mathbb{N})$, $\langle f(\overline{n}), g(\overline{n})\rangle = \|f(\overline{n})\|$ and $\|g(\overline{n})\| = 1$. It follows,

$$|\partial_{\mathcal{U}}^{2k}(f \otimes g)| = |\partial_{\mathcal{U}}^k f \otimes \partial_{\mathcal{U}}^k g| = |\langle \partial_{\mathcal{U}}^k f, \partial_{\mathcal{U}}^k g\rangle| \leq \|\partial_{\mathcal{U}}^k f\| \|\partial_{\mathcal{U}}^k g\| = \|\partial_{\mathcal{U}}^k f\|$$

The map $f \otimes g : G_{2k}(\mathbb{N}) \to \mathbb{R}$ is bounded, then we can apply Proposition 2.3 and for every $\varepsilon > 0$ there exists $\mathbb{A}$ an infinite subset of $\mathbb{N}$ such that for every $\overline{n} \in G_{2k}(\mathbb{A}), |f \otimes g(\overline{n}) - \partial_{\mathcal{U}}^{2k} f \otimes g| < \varepsilon$, hence

$$|f \otimes g(\overline{n})| < \varepsilon + |\partial_{\mathcal{U}}^{2k} f \otimes g| \leq \varepsilon + \|\partial_{\mathcal{U}}^k f\|.$$

Now we enumerate $\mathbb{A} = \{m_1 < n_1 < m_2 < n_2 < \cdots < m_j < n_j < \ldots\}$ and set $\mathbb{M} = \{m_1, \ldots, m_j, \ldots\}$.

Let $\overline{n} \in G_k(\mathbb{M})$, then for any $\overline{p} \in G_k(\mathbb{A})$ which is adjacent to $\overline{n}$ (such a $\overline{p}$ exists by the definitions of $\mathbb{A}$ and $\mathbb{M}$), we have

$$\begin{aligned}\|f(\overline{n})\| &= \langle f(\overline{n}), g(\overline{n})\rangle = \langle f(\overline{p}), g(\overline{n})\rangle + \langle f(\overline{n}) - f(\overline{p}), g(\overline{n})\rangle \\ &\leq f \otimes g(n_1, p_1, \ldots, n_k, p_k) + \|f(\overline{n}) - f(\overline{p})\| \|g(\overline{n})\| \\ &< \varepsilon + \|\partial_{\mathcal{U}}^k f\| + \omega_f(d(\overline{n}, \overline{p})) = \varepsilon + \|\partial_{\mathcal{U}}^k f\| + \omega_f(1)\end{aligned}$$

□

We can now state the result we will use to prove the main theorem:

Corollary 2.5. *Let X be a reflexive Banach space and $f : G_k(\mathbb{N}) \to X$ be a bounded map. Then for every $\varepsilon > 0$, there exists $\mathbb{M}$, an infinite subset of $\mathbb{N}$, and $x \in X$ such that:*

$$\forall \overline{n} \in G_k(\mathbb{M}), \|f(\overline{n}) - x\| \leq \omega_f(1) + \varepsilon.$$

Proof. Since X is reflexive there exists $x \in X$ such that $\partial_{\mathcal{U}}^k f = x$. We define a bounded map $g : G_k(\mathbb{N}) \to X$ by $g(\overline{n}) = f(\overline{n}) - x$, for all $\overline{n} \in G_k(\mathbb{N})$. Clearly $\partial_{\mathcal{U}}^k g = 0$ and $\omega_g(1) = \omega_f(1)$. Finally by a direct application of the previous lemma:

$$\forall \varepsilon > 0, \exists \mathbb{M} \subseteq \mathbb{N} : \forall \overline{n} \in G_k(\mathbb{M}), \|g(\overline{n})\| < \|\partial_{\mathcal{U}}^k g\| + \omega_g(1) + \varepsilon.$$

That is,

$$\forall \varepsilon > 0, \exists \mathbb{M} \subseteq \mathbb{N} : \forall \overline{n} \in G_k(\mathbb{M}), \|f(\overline{n}) - x\| < \omega_f(1) + \varepsilon.$$

□

3. Obstruction to uniform or coarse embeddings into reflexive Banach spaces

Given (M, d) a metric space, $\varepsilon > 0$ and $\delta \geq 0$, we say that M has the $\mathcal{Q}(\varepsilon, \delta)$-property if for every $k \in \mathbb{N}$, for every map $f : G_k(\mathbb{N}) \to M$ with $\omega_f(1) \leq \delta$, there exists an infinite subset $\mathbb{M}$ of $\mathbb{N}$ such that:

$$\forall \overline{n} < \overline{m} \in G_k(\mathbb{M}),\ d(f(\overline{n}), f(\overline{m})) \leq \varepsilon.$$

We define $\Delta_M(\varepsilon)$ as the supremum over all $\delta \geq 0$ such that M has the $\mathcal{Q}(\varepsilon, \delta)$-property.

The key result of this paper is the following:

Theorem 3.1. *Let (M, d) be a metric space.*

(1) *If M uniformly embeds into a reflexive Banach space, then*

$$\forall \varepsilon > 0, \Delta_M(\varepsilon) > 0.$$

(2) *If M coarsely embeds into a reflexive Banach space, then*

$$\lim_{\varepsilon \to +\infty} \Delta_M(\varepsilon) = +\infty.$$

Proof. Let X be a reflexive Banach space and $h : M \to X$ be any map.

We will prove that for every $\delta > 0$ and $f : G_k(\mathbb{N}) \to M$ a map such that $\omega_f(1) \leq \delta$, there exists an infinite subset $\mathbb{M}$ of $\mathbb{N}$ so that for every $\overline{n} < \overline{p} \in G_k(\mathbb{M})$, $\varphi_h(d(f(\overline{n}), f(\overline{p}))) \leq 4\ \omega_h(\delta)$ and conclude.

Let $\delta > 0$ and $f : G_k(\mathbb{N}) \to M$ be a map such that $\omega_f(1) \leq \delta$. We can apply Corollary 2.5 on the map $h \circ f : G_k(\mathbb{N}) \to X$, with $\varepsilon = \omega_{h \circ f}(1)$, to obtain $\mathbb{M}$, an infinite subset of $\mathbb{N}$, and $x \in X$ such that for every $\overline{n}, \overline{p} \in G_k(\mathbb{M})$,

$$\|h \circ f(\overline{n}) - h \circ f(\overline{p})\| \leq \|h \circ f(\overline{n}) - x\| + \|h \circ f(\overline{p}) - x\| \leq 4\ \omega_{h \circ f}(1) \leq 4\ \omega_h(\delta)$$

The last inequality holds because we clearly have $\omega_{h \circ f}(1) \leq \omega_h(\delta)$.

(1) **Uniform embedding.** Let $\varepsilon > 0$, then there exists $\alpha > 0$ such that $\varphi_h(\varepsilon) \geq 4\ \alpha$ and $\delta > 0$ so that $\omega_h(\delta) \leq \alpha$.

For this $\delta > 0$, for every $f : G_k(\mathbb{N}) \to M$ such that $\omega_f(1) \leq \delta$, there exists an infinite subset $\mathbb{M}$ of $\mathbb{N}$ such that $\forall \overline{n} < \overline{p} \in G_k(\mathbb{M})$,

$$\varphi_h(d(f(\overline{n}), f(\overline{p}))) \leq 4\ \omega_h(\delta) \leq 4\ \alpha \leq \varphi_h(\varepsilon).$$

We finally conclude that $d(f(\overline{n}), f(\overline{p})) \leq \varepsilon$, M has the $\mathcal{Q}(\varepsilon, \delta)$-property and $\Delta_M(\varepsilon) > 0$.

(2) **Coarse embedding.** Let $\delta > 0$, then there exist $\beta > 0$ such that $\omega_h(\delta) \leq \beta$ and $t > 0$ such that $\varphi_h(t) \geq 4\,\beta$.

Let ε be greater than t. Then for every $f : G_k(\mathbb{N}) \to M$ such that $\omega_f(1) \leq \delta$, there exists an infinite subset $\mathbb{M}$ of $\mathbb{N}$ such that $\forall \overline{n} < \overline{p} \in G_k(\mathbb{M})$,

$$\varphi_h(d(f(\overline{n}), f(\overline{p}))) \leq 4\,\omega_h(\delta) \leq 4\,\beta \leq \varphi_h(t) \leq \varphi_h(\varepsilon).$$

Then $d(f(\overline{n}), f(\overline{p})) \leq \varepsilon$ and $\Delta_M(\varepsilon) \geq \delta$. To conclude, $\lim_{\varepsilon \to +\infty} \Delta_M(\varepsilon) = +\infty$.

Which completes the proof. □

In the case where X is a Banach space, the function Δ_X has some particular properties:

Lemma 3.2. *Let X be a Banach space.*

(1) *There exists $0 \leq \mathcal{Q}_X \leq 1$ such that for every $\varepsilon > 0$, $\Delta_X(\varepsilon) = \mathcal{Q}_X \cdot \varepsilon$.*
(2) *For every $0 < \varepsilon \leq 1$, we have $\Delta_X(\varepsilon) = \Delta_{B_X}(\varepsilon)$.*

Proof.

(1) To prove that there exists a constant $\mathcal{Q}_X \geq 0$ such that for every $\varepsilon > 0$, $\Delta_X(\varepsilon) = \mathcal{Q}_X \cdot \varepsilon$, it is enough to prove that for every $\lambda > 0$, we have $\Delta_X(\lambda \cdot \varepsilon) = \lambda \cdot \Delta_X(\varepsilon)$. To do so consider $\delta > 0$ and prove that $\delta \leq \Delta_X(\lambda \cdot \varepsilon)$ is equivalent to $\delta \leq \lambda \cdot \Delta_X(\varepsilon)$, exanching the role played by the fonctions f and f/λ.

We will now prove that $\Delta_X(1) \leq 1$ and then conclude that $\mathcal{Q}_X \leq 1$.

Consider $(x_n)_{n\in\mathbb{N}}$ a sequence in X such that for all $m \neq n$, $\|x_n - x_m\| = 1$ and $f : G_1(\mathbb{N}) \to X$ defined by $f(n) = x_n, \forall n \in \mathbb{N}$. In this case $\omega_f(1) = 1$ and for every $n \neq m$, $\|f(n) - f(m)\| = 1$, thus $\mathcal{Q}_X = \Delta_X(1) \leq 1$.

(2) Finally let $0 \leq \varepsilon \leq 1$ and prove $\Delta_{B_X}(\varepsilon) = \Delta_X(\varepsilon)$.
   - Because B_X is a subset of X it is easy to see that $\Delta_{B_X}(\varepsilon) \geq \Delta_X(\varepsilon)$ for all $\varepsilon > 0$.
   - Let $k \in \mathbb{N}$ and $f : G_k(\mathbb{N}) \to X$ be a map.
   Remark that if there exists an infinite subset $\mathbb{M}$ of $\mathbb{N}$ such that for every $\overline{n} < \overline{m} \in G_k(\mathbb{M}), \|f(\overline{n}) - f(\overline{m})\| \leq \varepsilon$, then the image of $G_k(\mathbb{M})$ by f belongs to a ball of radius 1. Indeed if $\mathbb{M} = \{m_1 < \cdots < m_k < \cdots\}$, denote $\overline{m} = (m_1, \ldots, m_k)$ and $\mathbb{M}' = \{m_{k+1} < \cdots < m_j < \ldots\}$. Then for every $\overline{n} \in G_k(\mathbb{M}')$, we have $\|f(\overline{n}) - f(\overline{m})\| \leq \varepsilon \leq 1$, thus $f(G_k(\mathbb{M}')) \subseteq f(\overline{m}) + B_X$.
   So we can consider only $f : G_k(\mathbb{N}) \to X$ so that there exits $\mathbb{M}$ and $x_0 \in X$ such that $f(G_k(\mathbb{M})) \subseteq x_0 + B_X$ and $\omega_f(1) \leq \Delta_{B_X}(\varepsilon)$. Now for $\overline{n} \in G_k(\mathbb{M})$ define $g(\overline{n}) = f(\overline{n}) - x_0$. Because $g : G_k(\mathbb{M}) \to B_X$ and $\omega_g(1) \leq \Delta_{B_X}(\varepsilon)$, there exists $\mathbb{M}'$ an infinite subset of $\mathbb{M}$ such that for every $\overline{n} < \overline{m} \in G_k(\mathbb{M}')$, $\|g(\overline{n}) - g(\overline{m})\| \leq \varepsilon$, that is $\|f(\overline{n}) - f(\overline{m})\| \leq \varepsilon$.
   Finally we can conclude that $\Delta_X(\varepsilon) \geq \Delta_{B_X}(\varepsilon)$.

□

Thanks to this Lemma we are ready to define the so called $\mathcal{Q}$-property:

Definition 3.3. We say that a Banach space X has the $\mathcal{Q}$-property if $\mathcal{Q}_X > 0$.

We can use Theorem 3.1 in order to give an obstruction to uniform or coarse embeddings into reflexive Banach spaces in terms of property $\mathcal{Q}$.

Corollary 3.4. *Let X be a Banach space which fails the $\mathcal{Q}$-property. Then*
1. *B_X cannot be uniformly embedded into a reflexive Banach space.*
2. *X cannot be coarsely embedded into a reflexive Banach space.*

Proof.
1. Assume that B_X uniformly embeds into a reflexive Banach space. Then for every positive ε, $\Delta_{B_X}(\varepsilon) > 0$. But $\Delta_{B_X}(1) = \Delta_X(1) = \mathcal{Q}_X \cdot 1 > 0$, so finally X has the $\mathcal{Q}$-property.
2. Assume that X coarsely embeds into a reflexive Banach space. Then $\lim_{\varepsilon\to+\infty} \mathcal{Q}_X \cdot \varepsilon = \lim_{\varepsilon\to+\infty} \Delta_X(\varepsilon) = +\infty$, hence $\mathcal{Q}_X \neq 0$ and X has the $\mathcal{Q}$-property.

□

4. Examples

4.1. **Reflexive spaces.** It is clear by Corollary 3.4 that a reflexive Banach space has the $\mathcal{Q}$-property.

4.2. **Stable spaces.** Recall that a metric space (M, d) is stable if for every sequences $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ in M, if the following limits exist, then

$$\lim_{m\to+\infty} \lim_{n\to+\infty} d(x_m, y_n) = \lim_{n\to+\infty} \lim_{m\to+\infty} d(x_m, y_n).$$

It is proved in Section 2 of [11] that a stable metric space strongly uniformly embeds into a reflexive Banach space. So we deduce that a stable Banach space has the $\mathcal{Q}$-property. But we will prove this by another way: the next proposition is proved by a Ramsey type argument.

Proposition 4.1. *Let (M, d) be a stable metric space and $f : G_k(\mathbb{N}) \to M$ a bounded map. Then for every $\varepsilon > 0$ there exists $\mathbb{M}$, an infinite subset of $\mathbb{N}$, such that for every $\overline{n} < \overline{m} \in G_k(\mathbb{M})$,*

$$d(f(\overline{n}), f(\overline{m})) < \omega_f(1) + \varepsilon.$$

Proof. Since f is bounded, applying Theorem 2.1, we can find an infinite subset $\mathbb{M}$ of $\mathbb{N}$ and $a > 0$ such that for every $\overline{p}, \overline{q} \in G_k(\mathbb{M})$, $|d(f(\overline{p}), f(\overline{q})) - a| < \frac{\varepsilon}{4}$.

Let $\mathcal{U}$ be a non-principal ultrafilter which contains $\mathbb{M}$. Then,

$$\lim_{m_1\in\mathcal{U}} \lim_{n_1\in\mathcal{U}} \dots \lim_{m_k\in\mathcal{U}} \lim_{n_k\in\mathcal{U}} d(f(\overline{n}), f(\overline{m})) \leq \omega_f(1)$$

and because M is stable (see Lemma 9.19 in [6]),

$$\lim_{m_1\in\mathcal{U}} \dots \lim_{m_k\in\mathcal{U}} \lim_{n_1\in\mathcal{U}} \dots \lim_{n_k\in\mathcal{U}} d(f(\overline{n}), f(\overline{m})) \leq \omega_f(1).$$

Then, one can find $m_1 \leq \cdots \leq m_k \leq n_1 \leq \cdots \leq n_k$ such that

$$d(f(\overline{n}), f(\overline{m})) < \omega_f(1) + \frac{\varepsilon}{4}.$$

Therefore,

$$a < d(f(\overline{n}), f(\overline{m})) + \frac{\varepsilon}{4} < \omega_f(1) + \frac{\varepsilon}{2}.$$

Finally for every $\overline{p}, \overline{q} \in G_k(\mathbb{M})$,

$$d(f(\overline{p}), f(\overline{q})) < \frac{\varepsilon}{4} + a < \omega_f(1) + \varepsilon.$$

□

Corollary 4.2. *A stable Banach space X has the $\mathcal{Q}$-property.*

Proof. Let $\varepsilon > 0$ and $f : G_k(\mathbb{N}) \to X$ be such that $\omega_f(1) \leq \frac{\varepsilon}{2}$. In particular f is bounded and we can use the previous proposition to obtain an infinite subset $\mathbb{M}$ of $\mathbb{N}$ such that for every $\overline{n} < \overline{m} \in G_k(\mathbb{M})$, $\|f(\overline{n}) - f(\overline{m})\| \leq \omega_f(1) + \frac{\varepsilon}{2} \leq \varepsilon$, that is X has the $\mathcal{Q}$-property. □

4.3. **Some Banach spaces failing the $\mathcal{Q}$-property.** The following result will be useful to prove that some spaces do not have the $\mathcal{Q}$-property.

Theorem 4.3. *Let X be a Banach space with the $\mathcal{Q}$-property. Then for every $\varepsilon > 0$ and every $(x_n)_{n\in\mathbb{N}}$ bounded sequence in X with a w^*-cluster point $x^{**} \in X^{**}$, there exists a subsequence $(y_n)_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that*

$$\forall k \in \mathbb{N},\ \forall \overline{n} \in G_{2k}(\mathbb{N}),\ \|\sum_{j=1}^{2k}(-1)^j y_{n_j}\| \geq (1-\varepsilon)\mathcal{Q}_X k d(x^{**}, X).$$

Proof. Let $\varepsilon > 0$ and $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X with a w^*-cluster point $x^{**} \in X^{**}$. We will denote $B := \sup_{n\in\mathbb{N}} \|x_n\|$ and $\theta := d(x^{**}, X)$. We can assume that $\theta > 0$. Let $\lambda > 1$ and $P \in \mathbb{N}$ be such that $\frac{1}{\lambda^2} \geq 1 - \frac{\varepsilon}{2}$ and $\frac{1}{P} \leq \frac{\varepsilon \mathcal{Q}_X \theta}{2(2B+\theta)}$.

First it is possible to extract a subsequence $(v_n)_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that for every $1 \leq m < n$ and every sequence $(a_j)_{j=1}^n$ of positive numbers such that

$$\sum_{j=1}^{m} a_j = \sum_{j=m+1}^{n} a_j = 1,$$

we have

$$\|\sum_{j=1}^{m} a_j v_j - \sum_{j=m+1}^{n} a_j v_j\| > \frac{\theta}{\lambda}.$$

We will prove that one can find a subsequence $(y_n)_{n\in\mathbb{N}}$ of $(v_n)_{n\in\mathbb{N}}$ such that for every $k \geq 1$, there exists $b_k > 0$ such that for every $\overline{n} \in G_{2k}(\mathbb{N})$,

$$b_k - \frac{\theta}{P} \leq \|\sum_{j=1}^{2k}(-1)^j y_{n_j}\| \leq b_k.$$

Consider first $g_1 : \begin{array}{ccc} G_2(\mathbb{N}) & \to & \mathbb{R} \\ \overline{n} & \mapsto & \|v_{n_1} - v_{n_2}\| \end{array}$. Since the sequence $(v_n)_{n\in\mathbb{N}}$ is bounded, using Ramsey's theorem, one can find $b_1 > 0$ and $\varphi_1 : \mathbb{N} \to \mathbb{N}$ an increasing bijection such that:

$$\forall \overline{n} \in G_2(\varphi_1(\mathbb{N})),\ b_1 - \frac{\theta}{P} \leq g_1(\overline{n}) \leq b_1.$$

Now for a fixed $k \in \mathbb{N}$, assume that for every $1 \leq l \leq k-1$, φ_l is constructed such that $\varphi_l(\mathbb{N})$ is extracted from $\varphi_{l-1}(\mathbb{N})$. Consider $g_k : \begin{array}{ccc} G_{2k}(\mathbb{N}) & \to & \mathbb{R} \\ \overline{n} & \mapsto & \|\sum_{j=1}^{2k}(-1)^j v_{n_j}\| \end{array}$.

As previously there exists $b_k > 0$ and $\varphi_k : \mathbb{N} \to \mathbb{N}$ such that:

$$\forall \overline{n} \in G_{2k}(\varphi_1 \circ \cdots \circ \varphi_k(\mathbb{N})),\ b_k - \frac{\theta}{P} \leq g_k(\overline{n}) \leq b_k.$$

If we define $\psi : \begin{array}{ccc} \mathbb{N} & \to & \mathbb{N} \\ n & \mapsto & \varphi_1 \circ \cdots \circ \varphi_{P \cdot n}(n) \end{array}$, we obtain that if $n_1 \geq \frac{k}{P}$, then $\psi(n_1) = \varphi_1 \circ \cdots \circ \varphi_k(n_1)$. Thus the subsequence $(v_{\psi(n)})_{n \in \mathbb{N}}$ verifies: for every $k \in \mathbb{N}$, there exists a constant b_k such that $\forall \overline{n} \in G_{2k}(\mathbb{N})$ verifying $n_1 \geq \frac{k}{P}$, we have

$$b_k - \frac{\theta}{P} \leq \| \sum_{j=1}^{2k} (-1)^j v_{\psi(n_j)} \| \leq b_k.$$

We will denote the subsequence $(v_{\psi(n)})_{n \in \mathbb{N}}$ by $(y_n)_{n \in \mathbb{N}}$.

Fix $k \in \mathbb{N}$ and set $\mathbb{M} = \{n \in \mathbb{N};\ n \geq \frac{k}{P}\}$. Define $f : \begin{array}{ccc} G_k(\mathbb{M}) & \to & X \\ \overline{n} & \mapsto & \sum_{j=1}^{k} y_{n_j} \end{array}$. We have

$$\omega_f(1) = \sup \left\{ \| \sum_{j=1}^{k} y_{n_j} - \sum_{j=1}^{k} y_{m_j} \|;\ \frac{k}{P} \leq m_1 < n_1 < m_2 < \cdots < m_k < n_k \right\}$$
$$= \sup \left\{ \| \sum_{j=1}^{2k} (-1)^j y_{n_j} \|;\ \frac{k}{P} \leq n_1 < \cdots < n_{2k} \right\} \leq b_k.$$

Since X has the $\mathcal{Q}$-property, there exists $\mathbb{M}'$, an infinite subset of $\mathbb{M}$, such that for every $\overline{n} < \overline{m} \in G_k(\mathbb{M}')$, $\|f(\overline{n}) - f(\overline{m})\| \leq \dfrac{b_k}{\mathcal{Q}_X}$. So,

$$k \cdot \frac{\theta}{\lambda} < k \cdot \| \sum_{j=1}^{k} \frac{1}{k} y_{n_j} - \sum_{j=1}^{k} \frac{1}{k} y_{m_j} \| = \|f(\overline{n}) - f(\overline{m})\| \leq \frac{b_k}{\mathcal{Q}_X} < \frac{b_k}{\mathcal{Q}_X} \cdot \lambda$$

that is $b_k \geq \dfrac{\mathcal{Q}_X \cdot k \cdot \theta}{\lambda^2}$.

Now if $\overline{n} \in G_{2k}(\mathbb{N})$, one can find $\overline{m} \in G_{2k}(\mathbb{M})$ such that

$$\| \sum_{j=1}^{2k} (-1)^j y_{n_j} + \sum_{j=1}^{2k} (-1)^j y_{m_j} \| \leq \frac{2Bk}{P} \quad \text{or} \quad \| \sum_{j=1}^{2k} (-1)^j y_{n_j} - \sum_{j=1}^{2k} (-1)^j y_{m_j} \| \leq \frac{2Bk}{P}.$$

Finally,

$$\| \sum_{j=1}^{2k} (-1)^j y_{n_j} \| \geq \| \sum_{j=1}^{2k} (-1)^j y_{m_j} \| - \frac{2Bk}{P} \geq b_k - \frac{\theta}{P} - \frac{2Bk}{P} > b_k - \frac{k\theta}{P} - \frac{2Bk}{P}$$
$$\geq \frac{\mathcal{Q}_X k\theta}{\lambda^2} - \left(\frac{\varepsilon \mathcal{Q}_X k\theta}{2(2B+\theta)} \right) (\theta + 2B) \geq k\theta \mathcal{Q}_X (1 - \varepsilon),$$

which concludes the proof. ☐

Corollary 4.4. *The James space J and its dual J^* fail the $\mathcal{Q}$-property. In particular they cannot be uniformly or coarsely embedded into a reflexive Banach space.*

Proof. Let $(e_n)_{n\in\mathbb{N}}$ be the canonical basis of J and $x_n = \sum_{j=1}^{n} e_j$, $n \in \mathbb{N}$. With the notations of Theorem 4.3, we have $x^{**} = (1, \dots, 1, \dots) \in J^{**}$ and $d(x^{**}, X) = 1$. For every $k \in \mathbb{N}$ and every $\overline{n} \in G_{2k}(\mathbb{N})$,

$$\|\sum_{j=1}^{2k}(-1)^j x_{n_j}\|_J = (2k)^{1/2}$$

Finally assume J has the $\mathcal{Q}$-property, that is $\mathcal{Q}_J > 0$. Then for every $\varepsilon \leq 1$, one can find $k \in \mathbb{N}$ such that $(1-\varepsilon)\mathcal{Q}_J k \geq (2k)^{1/2}$. Thus $(x_n)_{n\in\mathbb{N}}$ does not verify the conclusion of Theorem 4.3.

In the case of J^* we consider the sequence $(e_n^*)_{n\in\mathbb{N}}$ which converges to an element of J^{***} of norm 1. Moreover for every $k \in \mathbb{N}$ and every $\overline{n} \in G_{2k}(\mathbb{N})$, we have $\| \sum_{j=1}^{2k}(-1)^j e_{n_j}^*\|_{J^*} \leq k^{1/2}$ and we conclude as previously.

The second part of the result is just a consequence of Corollary 3.4. □

4.4. **The space c_0.**

Corollary 4.5. *The space c_0 fails the $\mathcal{Q}$-property. In particular c_0 cannot be uniformly or coarsely embedded into a reflexive Banach space.*

Proof. We will prove that the summing bases of c_0 does not verify the conclusion of Theorem 4.3.

Let $(e_n)_{n\in\mathbb{N}}$ be the canonical bases of c_0 and $x_n = \sum_{j=1}^{n} e_j$, $n \in \mathbb{N}$. With the notation of Theorem 4.3, we have $x^{**} = (1, \dots, 1, \dots)$ and $d(x^{**}, X) = 1$. It is clear that for every $k \in \mathbb{N}$ and every $\overline{n} \in G_{2k}(\mathbb{N})$, we have $\| \sum_{j=1}^{2k}(-1)^j x_{n_j}\| = 1$.

Finally assume c_0 has the $\mathcal{Q}$-property, that is $\mathcal{Q}_{c_0} > 0$. Then for every $\varepsilon \leq 1$, one can find $k \in \mathbb{N}$ such that $(1-\varepsilon)\mathcal{Q}_{c_0} k \geq 1$. Thus $(x_n)_{n\in\mathbb{N}}$ does not verify the conclusion of Theorem 4.3.

The second part of the result is a consequence of Corollary 3.4. □

In fact in [11] Kalton proved, before the introduction of the $\mathcal{Q}$-property, that c_0 cannot be uniformly or coarsely embedded into a Banach space such that all its iterated duals are separable. This result is stronger because all iterated duals of J are separable and this space fails the $\mathcal{Q}$-property.

Theorem 4.6. *Let X be a Banach space such that all its duals are separable. Then c_0 cannot be uniformly or coarsely embedded into X.*

Lemma 4.7. *Let X be a Banach space such that for every $k \in \mathbb{N}$, the $2k$-th dual $X^{(2k)}$ of X is separable. Then for every uncountable family $(f_i)_{i\in I}$ of bounded functions $f_i : G_k(\mathbb{N}) \to X$ and for every $\varepsilon > 0$, there exist $i \neq j$ and $\mathbb{M}$, an infinite subset of $\mathbb{N}$, such that*

$$\forall \overline{n} \in G_k(\mathbb{M}), \|f_i(\overline{n}) - f_j(\overline{n})\| < \omega_{f_i}(1) + \omega_{f_j}(1) + \varepsilon.$$

Proof. For every $i \in I$, $\partial_{\mathcal{U}}^k f_i$ belongs to $X^{(2k)}$ and this space is separable thus there exist $i \neq j$ such that $\|\partial_{\mathcal{U}}^k f_i - \partial_{\mathcal{U}}^k f_j\| < \frac{\varepsilon}{2}$.

Now if we apply Lemma 2.4 to $f_i - f_j$, we obtain $\mathbb{M}$ an infinite subset of $\mathbb{N}$ such that for every $\overline{n} \in G_k(\mathbb{M})$,

$$\begin{aligned}\|f_i(\overline{n}) - f_j(\overline{n})\| = \|(f_i - f_j)(\overline{n})\| &< \|\partial_{\mathcal{U}}^k(f_i - f_j)\| + \omega_{f_i - f_j}(1) + \frac{\varepsilon}{2}\\ &< \omega_{f_i - f_j}(1) + \varepsilon \leq \omega_{f_i}(1) + \omega_{f_j}(1) + \varepsilon.\end{aligned}$$

□

Proof of Theorem 4.6: Let X be a Banach space having all its iterated duals separable and $h : c_0 \to X$ be a map. We will prove that h cannot be a coarse or uniform embedding. First, we can assume that h is bounded on bounded sets.

Let $(e_n)_{n\in\mathbb{N}}$ be the canonical basis of c_0 and define, for every $\mathbb{A}$ infinite subset of $\mathbb{N}$,

$$s_{\mathbb{A}}(n) = \sum_{\substack{r\leq n\\ r\in\mathbb{A}}} e_r, n \in \mathbb{N}.$$

Let $k \in \mathbb{N}$ and $0 < t < +\infty$ and define, for every $\mathbb{A}$ infinite subset of $\mathbb{N}$,

$$f_{\mathbb{A}} : \begin{array}{ccc} G_k(\mathbb{N}) & \to & c_0 \\ \overline{n} & \mapsto & t\sum_{j=1}^{k} s_{\mathbb{A}}(n_j) \end{array}$$

We have $\{h \circ f_{\mathbb{A}};\ \mathbb{A}$ infinite subset of $\mathbb{N}\}$ an uncountable family of bounded functions $h \circ f_{\mathbb{A}} : G_k(\mathbb{N}) \to X$, then we can apply Lemma 4.7: for every $\varepsilon > 0$, there exist $\mathbb{A} \neq \mathbb{B}$ and $\mathbb{M}$, infinite subsets of $\mathbb{N}$, such that

$$\forall \overline{n} \in G_k(\mathbb{M}), \|h \circ f_{\mathbb{A}}(\overline{n}) - h \circ f_{\mathbb{B}}(\overline{n})\| < \omega_{h\circ f_{\mathbb{A}}}(1) + \omega_{h\circ f_{\mathbb{B}}}(1) + \varepsilon.$$

Moreover, we have $\omega_{h\circ f_{\mathbb{D}}}(1) \leq \omega_h(t)$, for every $\mathbb{D}$ infinite subset of $\mathbb{N}$. Indeed $\omega_{f_{\mathbb{D}}}(1) \leq t$ and $\omega_{h\circ f_{\mathbb{D}}}(1) = \omega_h(\omega_{f_{\mathbb{D}}}(1)) \leq \omega_h(t)$.

Thus we have $\mathbb{A} \neq \mathbb{B}$ and $\mathbb{M}$, infinite subsets of $\mathbb{N}$, such that for every $\overline{n} \in G_k(\mathbb{N})$,

$$\|h \circ f_{\mathbb{A}}(\overline{n}) - h \circ f_{\mathbb{B}}(\overline{n})\| < 2\omega_h(t) + \varepsilon.$$

Since $\mathbb{A} \neq \mathbb{B}$ are infinite, there exists $\overline{p} \in G_k(\mathbb{M})$ such that $\|f_{\mathbb{A}}(\overline{p}) - f_{\mathbb{B}}(\overline{p})\| = kt$. Hence, $\varphi_h(kt) \leq \|h \circ f_{\mathbb{A}}(\overline{p}) - h \circ f_{\mathbb{B}}(\overline{p})\| < 2\omega_h(t) + \varepsilon$, for every $\varepsilon > 0$. Finally we have

$$\forall k \in \mathbb{N}, \forall t > 0, \varphi_h(kt) < 2\omega_h(t).$$

We will now distinguish two cases to prove that h cannot be a coarse or a uniform embedding:

- **Uniform embedding.** If $\lim_{t\to 0} \omega_h(t) = 0$, we deduce that for every $t > 0$, $\varphi_h(t) = 0$ and conclude that h cannot be a uniform embedding.
- **Coarse embedding.** If for every $t > 0$, $\omega_h(t)$ is finite, we can deduce that $\lim_{t\to+\infty} \varphi_h(t)$ is finite, that is h is not a coarse embedding.

□

5. Lipschitz and uniform embeddings into ℓ_∞

To conclude we mention that in [12] Kalton follows the same ideas to prove that $\mathcal{C}[1,\omega_1]$ cannot be uniformly embedded into ℓ_∞, where ω_1 is the first uncountable ordinal.

For every $k \in \mathbb{N}$ we define $G_k(\omega_1)$ the set of all subsets of ω_1 of size k. We keep the same notations as previously and define a distance d over $G_k(\omega_1)$ in the same way. Kalton proved the following results:

Theorem 5.1 (To compare to Corollary 2.5)**.** *Let $f : G_k(\omega_1) \to \ell_\infty$ be a Lipschitz mapping with Lipschitz constant L. Then there exist $x \in \ell_\infty$ and $\Omega \subset \omega_1$ such that for every $\overline{\alpha} \in G_k(\Omega)$,*

$$\|f(\overline{\alpha}) - x\| \leq \frac{L}{2}.$$

As a corollary (to compare to Corollary 4.5) it is proved:

Corollary 5.2. *The Banach space and $\mathcal{C}[1,\omega_1]$ cannot be uniformly embedded into ℓ_∞.*

References

1. I. Aharoni, *Every separable space is Lipschitz equivalent to a subset of c_0^+*, Israel J. Math. 19 (1974), 284-291.
2. I. Aharoni and J. Lindenstrauss, *Uniform equivalence between Banach spaces*, Bull. Amer. Math. Soc. 84 (1978), 281-283.
3. N. Aronszajn, *Differentiability of Lipschitz functions in Banach spaces*, Studia Math. 57 (1976), 147-160.
4. P. Assouad, *Remarques sur un article de Israel Aharoni sur les plongements Lipschitziens dans c_0*, Israel J. Math. 31 (1978), 97-100.
5. F. Baudier, *Embeddings of proper metric spaces into Banach spaces*, Houston J. Math 38 (2012), 209-224.
6. Y. Benyamini and J. Lindenstrauss, *Geometric nonlinear functional analysis*, A.M.S. Colloquium publications, vol 48, American Mathematical Society, Providence, RI, (2000).
7. J.P.R. Christensen, *Measure theoritic zero sets in infinite dimensional spaces and applications to differentiability of Lipschitz mappings*, in "Actes du Deuxième Colloque d'Analyse Fonctionnelle de Bordeaux, 1973", Vol. 2, 29-39.
8. G. Godefroy, G. Lancien and V. Zizler, *The non-linear geometry of Banach spaces after Nigel Kalton*, Rocky Mountain J. of Math., *to appear*.
9. G. Godefroy and N.J. Kalton, *Lipschitz-free spaces*, Studia Math. 159 (2003), 121-141.
10. N.J. Kalton, *Spaces of Lipschitz and Hölder functions and their applications*, Collect. Math. 55 (2004), 171-217.
11. N.J. Kalton, *Coarse and uniform embeddings into reflexive spaces*, Quart. J. Math. (Oxford) 58 (2007), 393-414.
12. N.J. Kalton, *Lipschitz and uniform embeddings into ℓ_∞*, Fund. Math. 212 (2011), 53-69.
13. N.J. Kalton and G. Lancien, *Best constants for Lipschitz embeddings of metric spaces into c_0*, Fund. Math. 199 (2008), 249-272.
14. G. Lancien, *A short course on nonlinear geometry of Banach spaces*, Topics in Functional and Harmonic Analysis, 77-102, Theta Series in Advanced Mathematics, Bucharest 2012, Ed. by C. Badea, D. Li and V. Petkova.
15. P. Mankiewicz, *On the differentiability of Lipschitz mappings in Frechet spaces*, Studia Math. 45 (1973), 15-29.
16. M. Mendel and A. Naor, *Metric cotype*, Ann. of Math. (2) 168(1) (2008), 247-298.
17. J. Pelant, *Embeddings into c0*, Topology Appl. 57 (1994), 259269.
18. W. T. Gowers, *Ramsey methods in Banach spaces*, Handbook of the geometry of Banach spaces, vol 2, North-Holland, Amsterdam (2003), 10711097.

Laboratoire de Mathématiques de Besançon, CNRS UMR 6623,, Université de Franche-Comté, 16 Route de Gray, 25030 Besançon Cedex, FRANCE.
E-mail address: aude.dalet@univ-fcomte.fr

Influence of uniform asymptotic smoothness on small compression embeddability.

Abstract

Following the exposition of N.J.Kalton and N.L.Randrianarivony in [KR], we will introduce some techniques related to metric midpoints and the notion of asymptotic uniform smoothness in order to study some results concerning the uniform structure of Banach spaces. In particular, we will show that $\ell_p \oplus \ell_2$ is not uniformly homeomorphic to L_p for $1 \leq p < \infty$. As in the recent paper of F. Baudier [B], we also will use the techniques from [KR] to compute the compression exponent $\alpha_{\ell_q}(\ell_p)$ of the embeddings of ℓ_p into ℓ_q, when $1 \leq p < q < \infty$.

1 Introduction

In this note, we will discuss useful tools for the study of uniform embeddings of metric spaces. These methods are particularly relevant for Lebesgue sequence spaces ℓ_p or more generally ℓ_p-sums of Banach spaces.
A Banach space B has unique uniform structure if whenever E is a Banach space uniformly homeomorphic to B, then E is linearly isomorphic to B. The space ℓ_p is known to have unique uniform structure for $1 < p < \infty$, by a result of W.B. Johnson, J. Lindenstrauss and G. Schechtman in [JLS]. It was also asked in [JLS] if $\ell_p \oplus \ell_2$ has unique uniform structure : the answer is still unknown. The uniqueness of the uniform structure of L_p is still an open question as well. As $\ell_p \oplus \ell_2$ is a complemented subspace of L_p, it is natural to ask if the two latter spaces are uniformly homeomorphic. The authors of [KR] proved the following theorem.

Theorem 1 (Theorems 5.2 and 5.6 in [KR]) *Let $1 \leq p < \infty$, $p \neq 2$. The space $\ell_p \oplus \ell_2$ is not uniformly homeomorphic to L_p.*

We will show that this result follows from two other interesting facts : for $p < r < 2$, the space $\ell_p \oplus \ell_2$ is not uniformly homeomorphic to any Banach space containing a copy of ℓ_r; and for $1 \leq p < \infty$, $p \neq 2$, $\ell_p \oplus \ell_2$ is

not uniformly homeomorphic to any Banach space containing $(\sum \ell_2)_{\ell_p}$.
The proofs of the results in [KR] rely on the combination of two main ideas. The first idea is the use of metric midpoints in the study of coarse Lipschitz embeddings, which yields some restrictions for embeddings between ℓ_p-spaces or ℓ_p-sums of Banach spaces. The second idea is the notion of asymptotic uniform smoothness which was used in the form given in the next theorem. We denote by $G_k(\mathbb{M})$ the set of k-tuples of elements of a subset $\mathbb{M} \subset \mathbb{N}$, equipped with the distance

$$d((n_1, ...n_k), (m_1, ...m_k)) = \frac{1}{2}|A \triangle B|$$

where $A = \{n_1, ...n_k\}$ and $B = \{m_1, ...m_k\}$.

Theorem 2 (Theorem 4.2 in [KR]) *Let $1 < p < \infty$. Let X be a reflexive Banach space such that*

$$\limsup ||x + x_n||^p \leq ||x||^p + \limsup ||x_n||^p \quad (*)$$

for all $x \in X$, and all weakly null sequence $(x_n)_n$ in X. Assume that $\mathbb{M}$ is an infinite subset of $\mathbb{N}$, and that $f : G_k(\mathbb{M}) \to X$ is a Lipschitz map. Then for any $\epsilon > 0$, there exists an infinite subset $\mathbb{M}'$ of $\mathbb{M}$ such that

$$\mathrm{diam}(f(G_k(\mathbb{M}'))) \leq 2\mathrm{Lip}(f)k^{1/p} + \epsilon.$$

G. lancien noticed in [L] that the tools used in [KR] can be used to give a very simple proof of the fact that ℓ_p is not uniformly homeomorphic to ℓ_q when $1 \leq p, q < \infty$, $p \neq q$(recall that a more general fact is known for $1 < p < \infty$ since ℓ_p has unique uniform structure). More precisely, he proved the following result.

Theorem 3 (Corollaries 4.8 and 4.10 in [L]) *Let $1 \leq p, q < \infty$, $p \neq q$. Then ℓ_p does not coarsely Lipschitz embed into ℓ_q.*

Theorem 2 was also used by F. Baudier in [B] to compute the ℓ_q-compression of ℓ_p, denoted by $\alpha_{\ell_q}(\ell_p)$.

Theorem 4 (Corollary 2.19 in [B]) *Let $1 \leq p < q < \infty$. Then*

$$\alpha_{\ell_q}(\ell_p) = \frac{p}{q}.$$

In the sequel, we assume the reader familiar with the following notions related to metric embeddings : uniform (coarse, Lipschitz, or coarse Lipschitz) embeddings (homeomorphisms), Lipschitz (coarse Lipschitz) constants, the *Lipschitz for large distances* principle, compression exponent of an embedding.
This note is organized as follows. In section 2, we discuss the metric midpoints technique. Section 3 is devoted to asymptotic uniform smoothness and the proof of Theorem 2. In section 4, we show Theorem 1 and Theorem 3 with the tools introduced in the two previous sections. In section 5, we give the proof of Theorem 4.
Sections 2,3 and 4 follow the exposition of [KR]. This note was also inspired by the course [L] where much more can be found on related subjects, and in particular on the use of the uniform asymptotic smoothness modulus.

Notations :

B_X (resp. S_X) : the closed unit ball (resp. sphere) of the space X

$[v_i]_i$: the closed linear span of the elements $(v_i)_i$

$A \lesssim B$: there exists a constant $C > 0$ such that $A \leq CB$

$$\mathrm{Lip}_s(f) = \sup_{d(x,y)\geq s} \frac{d(f(x), f(y))}{d(x, y)} \quad \text{and} \quad \mathrm{Lip}_\infty = \inf_{s>0} \mathrm{Lip}_s(f)$$

2 Metric midpoints

The aim of this section is to show the following proposition.

Proposition 5 (Proposition 3.5 in [KR]) *Let $(X_j)_j$ be a sequence of Banach spaces, and $1 \leq p < r < \infty$. Let $f : (\sum_j X_j)_{\ell_r} \to \ell_p$ be any coarse Lipschitz map. Then, for any $t > 0$, $\delta > 0$, there exist $x \in (\sum_j X_j)_{\ell_r}$, $\tau > t$ and a subspace $E \subset (\sum_j X_j)_{\ell_r}$ of the form $E = \{\ w = (w_j)_{j=1}^\infty \in (\sum_j X_j)_{\ell_r} \mid w_1 = ... = w_N = 0\ \}$ for some N, so that for some compact set $K \subset \ell_p$ we have $f(x + \tau B_E) \subset K + \delta\tau B_{\ell_p}$.*

Remark 6 (i) The same result holds for any equivalent norm to the usual norm on the domain space $(\sum_j X_j)_{\ell_p}$ (this is clear from the proof below).
(ii) One can generalize the previous proposition by replacing the target space ℓ_p by any finite direct sum of the form $(\sum_{j=1}^n \ell_{p_j})_{\ell_{p_n}}$ for $1 \leq p_1, ...p_n < r < \infty$ (see Proposition 3.6 in [KR]).

To show Proposition 5, we will use the by now well-known metric midpoint technique. The latter notion was introduced by Enflo in an unpublished

paper, to show that ℓ_1 and L_1 are not uniformly homeomorphic.
Let X be a metric space. For $x, y \in X$, and $\delta > 0$, the approximate metric midpoint of x, y with error δ is the set

$$\mathrm{Mid}(x,y,\delta) = \{\ z \in X \mid \max(d(x,z), d(y,z)) \leq (1+\delta)\frac{d(x,y)}{2}\ \}.$$

Before proving Proposition 5, we prove three lemmas.

Lemma 7 *Let X be a Banach space, and Y a metric space. Let $f : X \to Y$ be a coarse Lipschitz map. If $\mathrm{Lip}_\infty(f) > 0$, then for any $t, \epsilon > 0$ and any $0 < \delta < 1$ there exist $x, y \in X$ with $||x-y|| > t$, and*

$$f(\mathrm{Mid}(x,y,\delta)) \subset \mathrm{Mid}(f(x), f(y), (1+\epsilon)\delta).$$

Proof Fix t, ϵ, δ as in the statement of the lemma. Let $\nu > 0$.
Recall that for $s > 0$, we have $\mathrm{Lip}_s(f) = \sup_{d(x,y)\geq s} \frac{d(f(x),f(y))}{d(x,y)}$. We have also

$$\mathrm{Lip}_\infty(f) = \inf_{s>0} \mathrm{Lip}_s(f) = \lim_{s\to\infty} \mathrm{Lip}_s(f).$$

Then there exists $s > t$ such that : $\mathrm{Lip}_s(f) < (1+\nu)\mathrm{Lip}_\infty(f)$ (1).
On the other hand, for all $s > 0$ we have $\mathrm{Lip}_{2s(1-\delta)}(f) > \mathrm{Lip}_\infty(f)$. Hence we can find $x, y \in X$ satisfying

$$||x-y|| \geq 2s(1-\delta)^{-1} \quad (2)\ , \text{ and}$$
$$d(f(x), f(y)) > (1-\nu)\mathrm{Lip}_\infty(f)||x-y|| \quad (3).$$

Now let $u \in \mathrm{Mid}(x,y,\delta)$. By inequality (2) above, and a triangle inequality, it is clear that $||x-u|| > s$. So we obtain

$$\begin{aligned} d(f(x), f(u)) &\leq (1+\nu)\mathrm{Lip}_\infty(f)||x-u|| \text{ (by (1) and } ||x-u|| \geq s) \\ &\leq \frac{1}{2}(1+\nu)(1+\delta)\mathrm{Lip}_\infty(f)||x-y|| \text{ (since } u \in \mathrm{Mid}(x,y,\delta)) \\ &\leq \frac{1}{2}\frac{1+\nu}{1-\nu}(1+\delta)(f)d(f(x), f(y)) \text{ (by inequality (3)).} \end{aligned}$$

Repeating the inequalities above with y and u, we deduce that

$$\max(d(f(x), f(u)), d(f(y), f(u))) \leq \frac{1}{2}\frac{1+\nu}{1-\nu}(1+\delta)(f)d(f(x), f(y)).$$

The lemma follows if we choose ν sufficiently close to 0. ■

Lemma 8 *Let $1 \le p < \infty$, and let $(X_j)_j$ be a sequence of Banach spaces. Let $x, y \in (\sum_j X_j)_{\ell_p}$, and define $u = \frac{1}{2}(x+y)$, $v = \frac{1}{2}(x-y)$. then for any $0 < \delta < 1$, there is a closed subspace $E = \{ \ w = (w_j)_{j=1}^{\infty} \mid w_1 = ...w_N = 0 \ \}$ for some N, so that*

$$u + \delta^{1/p}||v||B_E \subset \mathrm{Mid}(x, y, \delta).$$

Proof For $p = 1$, this is easily checked with $N = 0$.
Let $p > 1$, and let $0 < \nu < (((1+\delta)^p - 1)^{1/p} - \delta^{1/p})||v||$. Take N such that $\sum_{j>N} ||v_j||^p < \nu^p$. Define $E := (\sum_{j>N} X_j)_{\ell_p} \subset (\sum_j X_j)_{\ell_p}$.
Let $z = \delta^{1/p}||v||z'$ for some $z' \in B_E$. By the choice of ν, it is clear that $||z|| < ((1+\delta)^p - 1)^{1/p}||v|| - \nu$. We now check that $u + z \in \mathrm{Mid}(x, y, \delta)$. Notice that $x - u - z = v - z$, so we have

$$\begin{aligned} ||x-u-z||^p &\le \sum_{j\le N} |v_j|^p + \sum_{j>N} |v_j - z_j|^p \\ &\le ||v||^p + ((\sum_{j>N} |v_j|^p)^{1/p} + (\sum_{j>N} |z_j|^p)^{1/p})^p \\ &\le ||v||^p + (\nu + ((1+\delta)^p - 1)^{1/p}||v|| - \nu)^p \\ &\le (1+\delta)^p ||v||^p. \end{aligned}$$

Hence we have $||x-u-z|| \le (1+\delta)||v||$. Since $y - u - z = -v - z$, we have also $||y-u-z|| \le (1+\delta)||v||$, and the lemma is proved. ■

Lemma 9 *Let $1 \le p < \infty$, $x, y \in \ell_p$, and define $u = \frac{1}{2}(x+y)$, $v = \frac{1}{2}(x-y)$. Then for any $0 < \delta < 1$, there is a compact set K such that*

$$\mathrm{Mid}(x, y, \delta) \subset K + 2\delta^{1/p}||v||B_{\ell_p}.$$

Proof Let $\nu > 0$, and write $v = (v_j)_{j=1}^{\infty} \in \ell_p$. Take N such that $\sum_{j>N} |v_j|^p < \nu^p$. Let $u + z \in \mathrm{Mid}(x, y, \delta)$, and write $z = z' + z''$ where $z' \in E_0 = [e_j]_{j\le N}$ and $z'' \in E = [e_j]_{j>N}$.
Since we have $2||z|| \le ||z-v|| + ||z+v|| \le 2(1+\delta)||v||$, we have $||z'|| \le (1+\delta)||v||$. Hence $u + z' \in K := u + (1+\delta)||v||E_0$.
Now from the convexity inequalities

$$|a|^p \le \frac{1}{2}(|a+b|^p + |a-b|^p) \text{ for all } a, b \in \mathbb{C},$$

we obtain

$$||v||^p - \nu^p + ||z''||^p \le \frac{1}{2}(||v+z||^p + ||v-z||^p),$$

and

$$\begin{aligned}||z''||^p &\le ((1+\delta)^p - 1)||v||^p + \nu^p \\ &\le 2^p\delta||v||^p\end{aligned}$$

for ν sufficiently small, since $((1+\delta)^p - 1) < 2^p$. This shows that $z'' \in 2\delta^{1/p}||v||B_{\ell_p}$. ■

Now we are able to give the proof of Proposition 5.

Proof of Proposition 5 : If $\mathrm{Lip}_\infty(f) = 0$, then for any t, there exists $\tau > t$ such that $\mathrm{Lip}_\tau(f) < \delta$. Then the conclusion of the proposition holds with $x = 0$ and $K = \{f(0)\}$.
Now we assume that $0 < \mathrm{Lip}_s(f) = C < \infty$ for some s. Take $0 < \nu < 1$ such that $4C\nu^{1/p-1/r} < \delta$. By Lemma 7, we can find $u, v \in (\sum_j X_j)_{\ell_r}$ such that $||u-v|| > \max(s, 2t\nu^{-1/r})$ and $f(\mathrm{Mid}(u,v,\nu)) \subset \mathrm{Mid}(f(u), f(v), 2\nu)$.
Let $x = \frac{1}{2}(u+v)$, and define $\tau = \nu^{1/r}||\frac{1}{2}(u-v)||$. By Lemma 8, there is a closed subspace $E = \{\ w = (w_j)_{j=1}^\infty \in (\sum_j X_j)_{\ell_r} \mid w_1 = ...w_N = 0\ \}$ for some N, so that $x+\tau B_E \subset \mathrm{Mid}(u,v,\nu)$. So $f(x+\tau B_E) \subset \mathrm{Mid}(f(u), f(v), 2\nu)$.
By Lemma 9, $\mathrm{Mid}(f(u), f(v), 2\nu) \subset K + 2\nu^{1/p}||f(u) - f(v)||B_{\ell_p}$ for some compact subset K.
Since $||u-v|| \ge s$ and $C = \mathrm{Lip}_s(f)$, we have

$$\begin{aligned}2\nu^{1/p}||f(u) - f(v)|| &\le 2\nu^{1/p}C||u-v|| \\ &= 4\nu^{1/p-1/r}C\tau \\ &\le \delta\tau.\end{aligned}$$

So proposition 5 is proved. ■

3 Asymptotic uniform smoothness

The aim of this section is to prove Theorem 2. First of all, we discuss briefly the assumptions of Theorem 2, in particular the asymptotic smoothness condition $(*)$. We end the section by two remarks concerning the study of the asymptotic uniform smoothness in the literature.

The reflexivity assumption on the space X is necessary : the non-reflexive space c_0 satisfies condition $(*)$ for any p, but every separable metric space can be Lipschitz embedded into c_0 by a result of Aharoni [A].

The ℓ_p-spaces satisfy condition $(*)$. Let us check this fact now. Take $x, x_n \in \ell_p(X)$ (for X an infinite countable set) such that $(x_n)_n$ is a weakly null sequence. Let $\epsilon > 0$, and let $\epsilon_k > 0$ be positive numbers, $I \subset X$ a finite subset such that

$$\sum_{k \notin I} |x_k|^p < \epsilon^p \quad \text{and} \quad \sum_{k \in X} |\epsilon_k|^p < \epsilon^p.$$

Denote $x_n = (x_{n,k})_k \in \ell_p(X)$. There exists N such that for all $n \geq N$, and all $k \in I$, we have

$$|x_k + x_{n,k}|^p \leq |x_k|^p + \epsilon_k^p.$$

Moreover, by Minkowski inequality, we have

$$\begin{aligned}\sum_{k \notin I} |x_k + x_{n,k}|^p &\leq ((\sum_{k \notin I} |x_k|^p)^{1/p} + (\sum_{k \notin I} |x_{n,k}|^p)^{1/p})^p \\ &\leq (\epsilon + ||x_n||)^p.\end{aligned}$$

Hence it follows, for $n \geq N$,

$$\begin{aligned}||x + x_n||^p &= \sum_{k \in I} |x_k + x_{n,k}|^p + \sum_{k \notin I} |x_k + x_{n,k}|^p \\ &\leq ||x||^p + \epsilon^p + (\epsilon + ||x_n||)^p.\end{aligned}$$

By passing to the $\limsup$ and letting ϵ tend to 0, this shows that x and $(x_n)_n$ satisfy inequality $(*)$.

Now we give the proof of Theorem 2.

Proof of Theorem 2 : The theorem is a straightforward consequence of the following statement : for any $k \in \mathbb{N}\backslash\{0\}$, any Lipschitz map $f : G_k(\mathbb{M}) \to X$ and any $\epsilon > 0$, there exists an infinite subset $\mathbb{M}' \subset \mathbb{M}$ and $u \in X$ such that

$$||f(n_1, ...n_k) - u|| < \text{Lip}(f)k^{1/p} + \epsilon/2 \text{ for all } (n_1, ...n_k) \in G_k(\mathbb{M}').$$

We show this statement by induction on k. For $k = 1$, there exists a subset $\mathbb{M}_0$ such that $(f(n))_{n \in \mathbb{M}_0}$ converges weakly (since X is reflexive and the sequence $(f(n))_n$ is bounded in X), and we denote $u = \lim_{n \in \mathbb{M}_0} f(n)$ its limit. Then for all $n \in \mathbb{M}$, we have

$$\begin{aligned}||f(n) - u|| &\leq \lim_{m \in \mathbb{M}_0} ||f(n) - f(m)|| \\ &\leq \text{Lip}(f).\end{aligned}$$

The statement for $k = 1$ follows immediately.
Now we assume that the inductive statement holds for $k-1$, and let $f : G_k(\mathbb{M}) \to X$ be a Lipschitz map, and $\epsilon > 0$.
By weak compactness, there exists an infinte subset $\mathbb{M}_0 \subset \mathbb{M}$ such that for all $\overline{n} = (n_1, ...n_{k-1}) \in G_{k-1}(\mathbb{M})$, the sequence $(f(\overline{n}, n_k))_{n_k}$ converges weakly along $\mathbb{M}_0$, and we denote $\tilde{f}(\overline{n})$ its limit. The map $\tilde{f} : G_{k-1}(\mathbb{M}) \to X$ is bounded and satisfies $\mathrm{Lip}(\tilde{f}) \leq \mathrm{Lip}(f)$.
Let $\epsilon' > 0$. By the inductive assumption, there exists an infinite subset $\mathbb{M}_1 \subset \mathbb{M}_0$ and $u \in X$ such that

$$||\tilde{f}(\overline{n}) - u|| < \mathrm{Lip}(\tilde{f})(k-1)^{1/p} + \epsilon'.$$

By assumption $(*)$, we have

$$\begin{aligned}\limsup_{n_k \in \mathbb{M}_1} ||f(\overline{n}, n_k) - u||^p \leq \; & (\mathrm{Lip}(\tilde{f})(k-1)^{1/p} + \epsilon')^p \\ & + \limsup_{n_k \in \mathbb{M}_1} ||f(\overline{n}, n_k) - \tilde{f}(\overline{n})||^p.\end{aligned}$$

As for the case $k = 1$, we have

$$\begin{aligned}\limsup_{n_k \in \mathbb{M}_1} ||f(\overline{n}, n_k) - \tilde{f}(\overline{n})||^p &\leq \limsup_{n_k \in \mathbb{M}_1} \limsup_{n'_k \in \mathbb{M}_1} ||f(\overline{n}, n_k) - f(\overline{n}, n'_k)||^p \\ &\leq \mathrm{Lip}(f)^p.\end{aligned}$$

Hence we obtain

$$\begin{aligned}\limsup_{n_k \in \mathbb{M}_1} ||f(\overline{n}, n_k) - u|| &\leq ((\mathrm{Lip}(f)(k-1)^{1/p} + \epsilon')^p + \mathrm{Lip}(f)^p)^{1/p} \\ &\leq (\mathrm{Lip}(f)^p k + f_1(\epsilon'))^{1/p} \\ &\leq \mathrm{Lip}(f)k^{1/p} + f_2(\epsilon')\end{aligned}$$

for some functions f_1, f_2 (depending on L, k, p) which tend to 0 as ϵ' tends to 0. Then we choose ϵ' so that $f_2(\epsilon') < \epsilon/4$. By the inequality above, we can find an infinite subset $\mathbb{M}' \subset \mathbb{M}_1$ such that the following holds

$$||f(\overline{n}) - u|| \leq \mathrm{Lip}(f)k^{1/p} + \epsilon/2.$$

This completes the proof of Theorem 2. ∎

Remark 10 In section 4 of [KW], the authors characterize Banach spaces with property (m_p), that is Banach spaces satisfying :

$\limsup ||x+x_n|| = (||x||^p + \limsup ||x_n||^p)^{1/p}$ for all weakly null sequence $(x_n)_n$.

Property (m_p) is obviously a strenghtening of assumption $(*)$. It is shown in [KW] that the spaces ℓ_p and the Bergman spaces on the unit disk have property (m_p) for $1 < p < \infty$ and $p \neq 2$, whereas L_p and the Schatten p-ideals S_p don't have property (m_p).

Remark 11 The following modulus of asymptotic uniform smoothness for a Banach space X, was introduced by V. Milman in [M] :

$$\overline{\rho}_X(t) = \sup_{x \in S_X} \inf_{\dim(X/Y)<\infty} \sup_{y \in S_Y} (||x + ty|| - 1).$$

For instance, we have :
- $\overline{\rho}_X(t) = (1 + t^p)^{1/p} - 1$ if $X = (\sum_n F_n)_{\ell_p}$, where $1 \leq p < \infty$ and the F_n's are finite dimensional;
- $\overline{\rho}_X(t) = 0$, for $0 < t < 1$ and $X = c_0$.
For all $x \in X \backslash \{0\}$, and all weakly null sequence $(x_n)_n$ in X, we have the following generalization of assumption $(*)$:

$$\limsup ||x + x_n|| \leq ||x||(1 + \overline{\rho}_X(\frac{\limsup ||x_n||}{||x||})).$$

This was used in [KR] to prove a more general version of Theorem 2 (see Theorem 6.1 in [KR], and section 4.4 in [L]).

4 About uniform structure of $\ell_p \oplus \ell_2$

In this section, we first prove Theorem 1, and then Theorem 3. We will need the following Ramsey-type argument in our proofs.

Lemma 12 *Let X be a Banach space, $\mathbb{M}$ be an infinite subset of $\mathbb{N}$, and $f : G_k(\mathbb{M}) \to X$ be any map with the property that for some compact set K and some $\delta > 0$, we have $f(G_k(\mathbb{M})) \subset K + \delta B_X$. Then for any $\epsilon > 0$, there is an infinite subset $\mathbb{M}' \subset \mathbb{M}$ such that* $\text{diam}(f(G_k(\mathbb{M}'))) \leq 2\delta + \epsilon$.

Proof Decompose f as $f = g + h$ where $g : G_k(\mathbb{M}) \to K$ and $h : G_k(\mathbb{M}) \to \delta B_X$. By Ramsey's theorem (applied to a finite covering of K by balls of radius $\epsilon/2$), there exists an infinite subset $\mathbb{M}' \subset \mathbb{M}$ such that $\text{diam}(g(G_k(\mathbb{M}))) \leq \epsilon$. ■

For $p < 2$, Theorem 1 will be an easy consequence of the following proposition.

Proposition 13 (see Theorem 5.1 in [KR]) *Let $1 \leq p < r < 2$. Then ℓ_r does not coarse Lipschitz embed into $\ell_p \oplus \ell_2$.*

Proof Let $f : \ell_r \to \ell_p \oplus \ell_2$ be a coarse Lipschitz embedding. Consider f as a map into $\ell_p \oplus_\infty \ell_2$ and assume (after a rescaling of f) that

$$||x-y|| \leq ||f(x)-f(y)|| \leq C||x-y|| \text{ whenever } ||x-y|| \geq 1.$$

Denote $f(x) = (g(x), h(x))$. Let $k \in \mathbb{N}$, and $\delta > 0$. By Proposition 5, there exist $\tau > k, x \in \ell_r$, and $N \in \mathbb{N}$ such that $g(x + \tau B_E) \subset K + \delta\tau B_{\ell_p}$ for some compact subset $K \subset \ell_p$, and where $E = [e_j]_{j>N}$.
Let $\mathbb{M} = \{\ n \in \mathbb{N} \mid n > N\ \}$, and define $\varphi : G_k(\mathbb{M}) \to \ell_r$ by

$$\varphi(n_1, ...n_k) = x + \tau k^{-1/r}(e_{n_1} + ... + e_{n_k}).$$

It is clear that $k^{-1/r}(e_{n_1} + ... + e_{n_k}) \in B_E$ for $(n_1, ...n_k) \in \mathbb{M}$, hence we have $g \circ \varphi(G_k(\mathbb{M})) \subset K + \delta\tau B_{\ell_p}$. By Lemma 12, there exists an infinite subset $\mathbb{M}_0 \subset \mathbb{M}$ such that $\mathrm{diam}(g \circ \varphi(G_k(\mathbb{M}_0))) \leq 3\delta\tau$.
Moreover, we have $\mathrm{Lip}(\varphi) \leq 2^{1/r}\tau k^{-1/r}$. Indeed, take $\overline{n} = (n_1, ...n_k)$ and $\overline{n}' = (n'_1, ..., n'_k)$ in $\mathbb{M}$. After reordering (this operation clearly does not change the computation below), we can assume that $n_i = n'_i$ for $i \leq s$, and that $n_i \neq n'_i$ for $i > s$. Notice that in such a case we have $d(\overline{n}, \overline{n}') = k - s$. Then the following equalities hold :

$$\begin{aligned} ||\varphi(\overline{n}) - \varphi(\overline{n}')||_r &= \tau k^{1/r}||(e_{n_1} + ... + e_{n_k}) - (e_{n'_1} + ... + e_{n'_k})||_r \\ &= \tau k^{-1/r}(\sum_{i>s} |e_{n_i}|^r + |e_{n'_i}|^r)^{1/r} \\ &\leq \tau k^{-1/r} 2^{1/r} d(\overline{n}, \overline{n}'). \end{aligned}$$

Then it follows that $\mathrm{Lip}(h \circ \varphi) \leq 2^{1/r}C\tau k^{-1/r}$. By Theorem 2 (with $p = 2$, and $\epsilon = 2^{1/r}C\tau k^{1/2-1/r}$), we have $\mathrm{diam}(h \circ \varphi(G_k(\mathbb{M}'))) \leq 3 \times 2^{1/r}C\tau k^{1/2-1/r}$. Thus

$$\mathrm{diam}(f \circ \varphi(G_k(\mathbb{M}'))) \leq 3 \times 2^{1/r}\tau(Ck^{1/2-1/r} + \delta).$$

On the other hand, it is clear that $\mathrm{diam}(\varphi(G_k(\mathbb{M}'))) > \tau$, so that

$$\mathrm{diam}(f \circ \varphi(G_k(\mathbb{M}'))) > \tau.$$

Then the following inequality holds :

$$1 < 3 \times 2^{1/r}(Ck^{1/2-1/r} + \delta).$$

For k large enough, and δ close enough to 0, this gives a contradiction. ■

Corollary 14 (Theorem 5.2 in [KR]) *Let $1 \leq p < 2$. Then L_p is not uniformly homeomorphic to $\ell_p \oplus \ell_2$.*

Proof Because of the *Lipschitz for large distances* principle, it is sufficient to prove that L_p does not coarsely Lipschitz embed into $\ell_p \oplus \ell_2$. But this is a well-known fact that ℓ_r isometrically embeds into L_p for $1 \leq p < r < 2$. Then the result follows from the previous Proposition 13. ■

Remark 15 In [KR], a more general version of Proposition 13 is proved (with an analog proof) in Theorem 5.1. From this version, the authors show that $\ell_p \oplus \ell_q$ has unique uniform structure when $1 < p < 2 < q < \infty$. The uniqueness of the uniform structure of $\ell_p \oplus \ell_q$ was proved in [JLS] for the cases $1 < p < q < 2$ and $2 < p < q < \infty$.

Now we deal with the case $p > 2$. The second half of Theorem 1 is a consequence of the following obstruction for coarse Lipschitz embeddings into $\ell_p \oplus \ell_2$.

Proposition 16 (Theorem 5.5 in [KR]) *Let $2 < p < \infty$. Then there is no coarse Lipschitz embedding of $(\sum \ell_2)_{\ell_p}$ into $\ell_p \oplus \ell_2$.*

Proof The proof is a slight modification of the proof of Proposition 13. Take $f = (g, h) : (\sum \ell_2)_{\ell_p} \to \ell_2 \oplus_\infty \ell_p$ satisfying the Lipschitz condition with constant 1 and C for distances ≥ 1, as in Proposition 13. Let $k \in \mathbb{N}$ and $\delta > 0$. For every i, let $(e_{ij})_j$ be the canonical basis of the i-th coordinate space ℓ_2 in $(\sum \ell_2)_{\ell_p}$.
By Proposition 5, there exist $\tau > k$, $x \in (\sum \ell_2)_{\ell_p}$ and N such that $g(x + \tau B_E) \subset K + \delta\tau B_{\ell_2}$ for some compact subset $K \subset \ell_2$, and where $E = [e_{ij}]_{i>N, j\geq 1}$.
Define $\varphi : G_k(\mathbb{N}) \to (\sum \ell_2)_{\ell_p}$ by

$$\varphi(n_1, ...n_k) = x + \tau k^{-1/2}(e_{N+1,n_1}, ..., e_{N+1,n_k}).$$

Since $g\circ\varphi(G_k(\mathbb{N})) \subset K + \delta\tau B_{\ell_2}$, Lemma 12 implies that $\text{diam}(g\circ\varphi(G_k(\mathbb{M}_0))) \leq 3\delta\tau$ for some infinite subset $\mathbb{M}_0 \subset \mathbb{N}$.
Moreover we have $\text{Lip}(h \circ \varphi) \leq C\sqrt{2}\tau k^{-1/2}$, so by Theorem 2 there exists an infinite subset $\mathbb{M} \subset \mathbb{M}_0$ such that $\text{diam}(G_k(\mathbb{M})) \leq 3\sqrt{2}C\tau k^{1/p-1/2}$.
Thus $\text{diam}(f \circ \varphi(G_k(\mathbb{M}))) \leq 3\sqrt{2}\tau(Ck^{1/p-1/2} + \delta)$. On the other hand, we have $\text{diam}(f \circ \varphi(G_k(\mathbb{M}))) > \tau$. Hence

$$1 < 3\sqrt{2}(Ck^{1/p-1/2} + \delta),$$

which is a contradiction for large k and small δ. ■

The following corollary completes the proof of Theorem 1.

Corollary 17 (Theorem 5.6 in [KR]) *Let $2 < p < \infty$. Then L_p is not uniformly homeomorphic to $\ell_p \oplus \ell_2$.*

Proof Since $2 < p < \infty$, ℓ_2 embeds isometrically in L_p. Hence $(\sum \ell_2)_{\ell_p}$ embeds isometrically in $(\sum L_p)_{\ell_p} \simeq L_p$. Then the result follows from Proposition 16. ■

Remark 18 The authors of [KR] also prove (with the same idea but again with some modification on the embedding of the discrete sets $G_k(\mathbb{M})$) that there is no coarse Lipschitz embedding of $(\sum \ell_2)_{\ell_p}$ into $\ell_p \oplus \ell_2$ when $1 \leq p < 2$ (Theorem 5.7).

Now using the same tools as in the previous proofs, we prove Theorem 3.

Proof of Theorem 3 : First let $1 \leq p < q < \infty$, and let $f : \ell_q \to \ell_p$ be a coarse Lipschitz map. Let $\delta > 0$. By Proposition 5, there exists $x \in \ell_q$, $N \in \mathbb{N}$ and $\tau > 0$ (which can be chosen arbitrary large) such that

$$f(x + \tau B_{E_N}) \subset K + \delta\tau B_{\ell_p}$$

for some compact subset $K \subset \ell_p$, and $E_N = [e_j]_{j>N}$. For $n \geq 1$, define $x_n = x + \tau e_{N+n}$. Then $||x_n - x_m|| \geq \tau$ whenever $n \neq m$. Moreover for all $n \geq 1$, we have $f(x_n) = k_n + \delta\tau v_n$ for some $k_n \in K$ and $v_n \in B_{\ell_p}$. By passing to a subsequence still denoted by $(x_n)_n$, we have $||f(x_n) - f(x_m)||_p \leq 3\delta\tau$ for all $n, m \in \mathbb{N}$. Since δ can be chosen arbitrary small and τ arbitrary large, inequalities for the sequence $(x_n)_n$ contradicts the fact that f is a coarse Lipschitz embedding. Hence ℓ_q does not coarsely Lipschitz embed into ℓ_p when $1 \leq p < q < \infty$.
For the second half of the proof, let $1 \leq q < p < \infty$ and let $f : \ell_q \to \ell_p$ be a coarse Lipschitz map such that

$$||x - y||_q \leq ||f(x) - f(y)||_p \leq C||x - y||_q \text{ whenever } ||x - y||_q \geq 1.$$

Define $\varphi : G_k(\mathbb{N}) \to \ell_q$ by $\varphi(\overline{n}) = e_{n_1} + ... + e_{n_k}$. A computation as before gives $\mathrm{Lip}(f \circ \varphi) \leq 2C$. By Theorem 2, there exists an infinite subset $\mathbb{M} \subset \mathbb{N}$ such that $\mathrm{diam}(f \circ \varphi)(G_k(\mathbb{M})) \leq 6Ck^{1/p}$. On the other hand, $\mathrm{diam}(f \circ \varphi)(G_k(\mathbb{M})) \geq (2k)^{1/q}$ since $\mathbb{M}$ is infinite. Since $q < p$, we have a contradiction for large k, and the theorem is proved. ■

5 Compression exponent $\alpha_{\ell_q}(\ell_p)$

In this section, we show how Theorem 2 was used in [B] to compute the compression exponent $\alpha_{\ell_q}(\ell_p)$ for $1 \leq p < q < \infty$. First we recall the definition of the Y-compression exponent of X, for X, Y Banach spaces.

Definition 19 Let X, Y be Banach spaces. The Y-compression exponent of X, denoted by $\alpha_Y(X)$ is the supremum of all numbers $0 \leq \alpha \leq 1$ over all embeddings $f : X \to Y$ such that

$$||x-y||^\alpha \lesssim ||f(x)-f(y)|| \leq ||x-y|| \text{ whenever } ||x-y|| \geq 1.$$

Theorem 4 asserts that $\alpha_{\ell_q}(\ell_p) = \frac{p}{q}$ when $1 \leq p < q < \infty$. One half of the result is a consequence of the following result obtained in [AB].

Proposition 20 (Proposition 5.2 in [AB]) *Let $1 \leq p < q < \infty$. Then there exists a Lipschitz embedding of $(\ell_p, ||.||_p^{p/q})$ into $(\ell_q, ||.||_q)$. In particular, we have the inequality $\alpha_{\ell_q}(\ell_p) \geq \frac{p}{q}$.*

The proof of the previous proposition uses a construction of specific maps to define a Lipschitz embedding from $\ell_p(\mathbb{N}, \mathbb{R})$ into $\ell_q(\mathbb{N} \times \mathbb{Z} \times \mathbb{Z}, \mathbb{R})$. More precisely, the authors of [AB] prove (see Theorem 3.4 in [AB]) that there exist real-valued functions $(\psi_{j,k})_{(j,k)\in\mathbb{Z}}$ and positive constants $A_{p,q}, B_{p,q}$ such that

$$A_{p,q}|x-y|^p \leq \sum_{k\in\mathbb{Z}} \sum_{j\in\mathbb{Z}} |\psi_{j,k}(x) - \psi_{j,k}(x)|^q \leq B_{p,q}|x-y|^p \text{ for all } x, y \in \mathbb{R}.$$

Then we define

$$\begin{aligned} f : \ell_p(\mathbb{N}, \mathbb{R}) &\to \ell_q(\mathbb{N} \times \mathbb{Z} \times \mathbb{Z}, \mathbb{R}) \\ (x_i)_{i\in\mathbb{N}} &\mapsto (\psi_{j,k}(x_i) - \psi_{j,k}(0))_{i,j,k\in\mathbb{N}\times\mathbb{Z}\times\mathbb{Z}}. \end{aligned}$$

By the inequalities above, it is easily checked that the map f is a Lipschitz embedding of $(\ell_p, ||.||_p^{p/q})$ into $(\ell_q, ||.||_q)$.

Now we show how the uniform asymptotic smoothness argument was used in [B] to give an upper-bound to the compression exponent $\alpha_{\ell_q}(\ell_p)$.

Proof of Theorem 4 : In view of Proposition 20, we are left to show that $\alpha_{\ell_q}(\ell_p) \leq \frac{p}{q}$. Let $k \in \mathbb{N}$, and $0 \leq \alpha \leq 1$. Let $f : \ell_p \to \ell_q$ be a map such that

$$||x-y||^\alpha \lesssim ||f(x)-f(y)|| \leq ||x-y|| \text{ whenever } ||x-y|| \geq 1.$$

Define $\varphi : G_k(\mathbb{N}) \to \ell_p$ by

$$\varphi(n_1, ...n_k) = e_{n_1} + ... + e_{n_k} \text{ for all } (n_1, ...n_k) \in G_k(\mathbb{N}).$$

It is clear that φ is $2^{1/p}$-Lipschitz, so $f \circ \varphi$ is Lipschitz as well. Then by Theorem 2, there exists an infinite subset $\mathbb{M} \subset \mathbb{N}$ such that

$$\text{diam}(f \circ \varphi(G_k(\mathbb{M}))) \lesssim k^{1/q}.$$

On the other hand, we have $\text{diam}(\varphi(G_k(\mathbb{M}))) = (2k)^{1/p}$. It follows that

$$\text{diam}(f \circ \varphi(G_k(\mathbb{M}))) \gtrsim \text{diam}(\varphi(G_k(\mathbb{M})))^\alpha = k^{\alpha/p}.$$

The condition $k^{\alpha/p} \lesssim k^{1/q}$ for all $k \in \mathbb{N}$, implies that $\alpha \leq \frac{p}{q}$. Hence the theorem is proved. ■

References

[A] I. Aharoni. Every separable metric space is Lipschitz equivalent to a subset of c_0^+. *Isr. J. Math.*, 19, 284–291, 1974.

[AB] F. Albiac and F. Baudier. Embeddability of snowflaked metrics with applications to the nonlinear geometry of the spaces L_p and ℓ_p for $0 < p < \infty$. *arxiv*, 1206.3774, 2012.

[B] F. Baudier. Quantitative nonlinear embeddings into Lebesgue sequence spaces. *J. Top. Anal.* (to appear).

[JLS] W.B. Johnson, J. Lindenstrauss and G. Schechtman. Banach spaces determined by their uniform structures. *Geom. Funct. Anal.* 6, 430–470, 1996.

[JR] W.B. Johnson and N.L. Randrianarivony. ℓ_p $(p > 2)$ does not coarsely embed into a Hilbert space. *Proc. Amer. Math. Soc.*, 134, no. 4: 1045–1050, 2006.

[KR] N.J. Kalton and N.L. Randrianarivony. The coarse Lipschitz geometry of $\ell_p \oplus \ell_q$. *Math. Ann.* 341, 223–237, 2008.

[KW] N.J. Kalton and D. Werner. Property (M), M-ideals, and almost isometric structure of Banach spaces. *J. Reine Angew. Math.*, 461, 137–178, 1995.

[L] G. Lancien. A short course in non-linear geometry of Banach spaces. *Topics in Functional and Harmonic Analyis*, Theta Series in Advanced Mathematics, 77–102, 2012.

[M] V.D. Milman. Geometric theory of Banach spaces. II. Geometry of the unit ball (Russian). *Uspehi Mat. Nauk*, 26, 73–149, 1971. English version : Russian Math. Surveys, 26, 79–163, 1971.

Address

Baptiste Olivier.
E-mail : baptiste.olivier100@gmail.com

ESTIMATING DISTORTION VIA METRIC INVARIANTS

SEAN LI

1. Introduction

In the latter half of the 20th century, researchers started realizing the importance of understanding how well certain metric spaces can quantitatively embed into other metric spaces. Here "well" depends on the context at hand. In essence, the better a metric space X embeds into Y, the closer geometric properties of X correspond to that of Y. If one continues on this train of thought, then one can try to embed a relatively not well understood metric space X into another more well understood space Y. If such a good embedding exists, then one can try to use the properties and techniques developed for Y to understand the geometry of X. If no such good embedding exists, then one can try to deduce the obstruction of such an embedding, which would give more information for both X and Y.

While part of the motivation came from purely mathematical considerations, the same philosophy also found use in the development of approximation algorithms in theoretical computer science. Indeed, one may be asked to solve a computational problem on a data sets that comes with a natural metric. There are some untractable problems that become much easier to solve when the data set contains some further structure—being Euclidean for example. Thus, be embedding the data set into this easier to solve space, one can speed up the computation at the loss of accuracy that can be bound by the fidelity of the embedding. While the topic of such applications are intresting on their own and could easily (and have) fill books, they are beyond the scope of these notes, and we will not pursue this line any further. See [18] for more information on approximation algorithms.

We now introduce the the quantity to which we measure how well a metric space embeds into one another. Recall that $f : (X, d_X) \to (Y, d_Y)$ is called a biLipschitz embedding if there exists some $D \geq 1$ so that there exists some $s \in \mathbb{R}$ for which

$$sd_X(x, y) \leq d_Y(f(x), f(y)) \leq Dsd_X(x, y).$$

Thus, up to rescalings of the metric, f preserves the metric of X in Y up to a multiplicative factor of D. Here, D is called the biLipschitz distortion (or just distortion) of f. Two metric spaces are said to be biLipschitz equivalent if there exists a surjective biLipschitz embedding between them. We will typically be calculating distortions of embeddings of metric spaces in Banach spaces and so by rescaling the function, we can usually suppose that $s = 1$ or $1/D$.

Given metric spaces X and Y, we let

$$c_Y(X) := \inf\{D : \exists f : X \to Y \text{ with distortion } D\},$$

with the understanding that $c_Y(X) = \infty$ if no such biLipschitz embedding exist. We can then say that $c_Y(X)$ is the distortion of X into Y without referencing any map.

Upper bounding $c_Y(X)$ typically entails constructing an explicit embedding for which you bound the distortion. We will be more interested in lower bounding $c_Y(X)$, for which one has to show that *all* embeddings must have biLipschitz distortion greater than our lower bound. There are many ways to achieve this. In these notes, we will use biLipschitz metric invariants (or just metric invariants) to accomplish such a task. We now introduce our first metric invariant, the Enflo type, which will give a nice simple example to show how one can use such properties to estimate distortion lower bounds.

2. Enflo type

We define the quantity

$$c_2(n) = \sup\{c_{\ell_2}(X) : X \text{ is an } n\text{-point metric space}\}.$$

Thus, every n-point metric space can embed into Hilbert space with distortion no more than $c_2(n)$.

Amazingly, it wasn't until 1970 that it was shown that $\sup_{n\in\mathbb{N}} c_2(n) = \infty$. The first proof of this fact was given by Enflo in [6]. Nowadays, it is known that $c_2(n) \asymp \log n$.[1] The upper bound was found by an explicit embedding by Bourgain [4] in 1985 and the lower bound was first matched by expander graphs as shown in [9] in 1995. The lower bound established in [6] is the following.

Theorem 2.1.

$$c_2(n) \geq \sqrt{\log n}.$$

We now describe the metric space used by Enflo. The Hamming cubes are the metric spaces

$$D_n = (\{0,1\}^n, \|\cdot\|_1).$$

Thus, elements of D_n are strings 0 and 1 of length n. These are just the corners of a cube of the ℓ_1^n normed space. We call two pairs of points in D_n an edge if they are of distance 1 (*i.e.* if their strings differ in only one place) and a diagonal if they are of distance n (*i.e.* if their strings differ at every place). Note that the metric of D_n can also be viewed as a graph path metric based on the set of edges. Each point x has n other points that form edges with x and 1 other point forms a diagonal with x.

To establish our lower bound, we need to show that $c_2(D_n) \geq \sqrt{n}$. One can easily verify that D_n also embed into ℓ_2 with distortion no more than $\sqrt{n}$ if one just embeds the points to the corresponding points of the unit cube in $\mathbb{R}^n$ so our lower bound will actually tight for this specific example.

Enflo proved Theorem 2.1 using the following proposition.

Proposition 2.2. *Let $f : D_n \to \ell_2$ be any map. Then,*

$$\sum_{\{x,y\}\in\textbf{diags}} \|f(x) - f(y)\|^2 \leq \sum_{\{u,v\}\in\textbf{edges}} \|f(u) - f(v)\|^2. \tag{1}$$

Note that we are not really using the metric structure of D_n here, just the graph structure. We will first need the following lemma.

Lemma 2.3 (Short diagonals lemma)**.** *Let x, y, z, w be arbitrary points in ℓ_2. Then*

$$\|x - z\|^2 + \|y - w\|^2 \leq \|x - y\|^2 + \|y - z\|^2 + \|z - w\|^2 + \|w - x\|^2.$$

[1] In these notes, we will say $a \lesssim b$ (resp. $a \gtrsim b$) if there exists some absolute constant $C > 0$ so that $a \leq Cb$ (resp. $a \geq Cb$). We write $a \asymp b$ if $a \lesssim b \lesssim a$.

Proof. As the norms are raised to the power 2, they break apart according to their coordinates. Thus, it suffices to prove that

$$(x-z)^2 + (y-w)^2 \leq (x-y)^2 + (y-z)^2 + (z-w)^2 + (w-x)^2.$$

But

$$(x-y)^2 + (y-z)^2 + (z-w)^2 + (w-x)^2 - (x-z)^2 - (y-w)^2 = (x-y+z-w)^2 \geq 0.$$

□

Proof of Proposition 2.2. We will induct on n. For the base case when $n = 2$, we simply set x, y, z, w to be the images of the points D_n so that $\{x, z\}$ and $\{y, w\}$ correspond to the images of diagonals. Lemma 2.3 then gives our needed inequality.

Now suppose we have shown the statement for D_{n-1} and consider D_n. Note that D_n can be viewed as two separate copies of D_{n-1}. Indeed, the set of points of D_{n-1} that correspond to strings all beginning with 0 form one such D_{n-1} and the subset corresponding to strings all beginning with 1 form the other. Let $D^{(0)}$ and $D^{(1)}$ denote these two subsets in D_{n-1}. It easily follows that, for each $v \in D^{(0)}$, there exists a unique $w \in D^{(1)}$ for which $\{v, w\}$ form an edge and vice versa. Let $\mathbf{edges}' \subset \mathbf{edges}$ denote this collection of edges. Let $\mathbf{edges}_0$ and $\mathbf{edges}_1$ denote the edges of $D^{(0)}$ and $D^{(1)}$, respectively. Note that these are still edges of D_n and

$$\mathbf{edges} = \mathbf{edges}_0 \cup \mathbf{edges}_1 \cup \mathbf{edges}', \tag{2}$$

where the union above is disjoint. Let $\mathbf{diags}_0$ and $\mathbf{diags}_1$ denote the diagonals of $D^{(0)}$ and $D^{(1)}$. Note that these are *not* diagonals of D_n as their distances are only $n-1$. However, if $\{u, v\}$ is a diagonal of $\mathbf{diags}_0$ and $u', v' \in D^{(1)}$ so that $\{u, u'\}$ and $\{v, v'\}$ are edges of $\mathbf{edges}'$, then $\{u', v'\} \in \mathbf{diags}_1$ and $\{u, v'\}, \{u', v\} \in \mathbf{diags}$.

By the induction hypothesis, we have that

$$\sum_{\{x,y\} \in \mathbf{diags}_0} \|f(x) - f(y)\|^2 \leq \sum_{\{u,v\} \in \mathbf{edges}_0} \|f(u) - f(v)\|^2, \tag{3}$$

$$\sum_{\{x,y\} \in \mathbf{diags}_1} \|f(x) - f(y)\|^2 \leq \sum_{\{u,v\} \in \mathbf{edges}_1} \|f(u) - f(v)\|^2. \tag{4}$$

Let $\{u, v\} \in \mathbf{diags}_0$. As was stated above, there exists a unique $\{u', v'\} \in \mathbf{diags}_1$ so that $\{u, v'\}, \{u', v\} \in \mathbf{diags}$. Using Lemma 2.3, we get that

$$\begin{aligned} &\|f(u) - f(v')\|^2 + \|f(u') - f(v)\|^2 \\ &\qquad \leq \|f(u) - f(v)\|^2 + \|f(v) - f(v')\|^2 + \|f(v') - f(u')\|^2 + \|f(u') - f(v)\|^2. \end{aligned}$$

As all diagonals of $\mathbf{diags}$ can be expressed in such manner, we get that

$$\begin{aligned} &\sum_{\{x,y\} \in \mathbf{diags}} \|f(x) - f(y)\|^2 \\ &\qquad \leq \sum_{\{u,v\} \in \mathbf{edges}'} \|f(u) - f(v)\|^2 + \sum_{i=0}^{1} \sum_{\{w,z\} \in \mathbf{diags}_i} \|f(w) - f(z)\|^2. \end{aligned} \tag{5}$$

The proposition now follows immediately from (2), (3), (4), and (5). □

We can now prove our main theorem.

Proof of Theorem 2.1. Let $f : D_n \to \ell_2$ be any embedding satisfying the biLipschitz bounds

$$sd(x,y) \leq \|f(x) - f(y)\| \leq Ds \cdot d(x,y), \tag{6}$$

for some $s \in \mathbb{R}$. We then get from Proposition (2.2) and the biLipschitz bounds of f that

$$\begin{aligned} s^2 n^2 |\mathbf{diags}| = s^2 \sum_{\{x,y\}\in\mathbf{diags}} d(x,y)^2 &\overset{(6)}{\leq} \sum_{\{x,y\}\in\mathbf{diags}} \|f(x)-f(y)\|^2 \\ &\overset{(1)}{\leq} \sum_{\{u,v\}\in\mathbf{edges}} \|f(u)-f(v)\|^2 \overset{(6)}{\leq} D^2 s^2 \sum_{\{u,v\}\in\mathbf{edges}} d(u,v)^2 = D^2 s^2 |\mathbf{edges}|. \end{aligned}$$

In the first and last equalities, we used the fact that edges and diagonals have distance 1 and n, respectively. One easily calculates that $|\mathbf{diags}| = 2^{n-1}$ and $|\mathbf{edges}| = n2^{n-1}$. This gives that

$$D \geq \sqrt{\frac{n^2|\mathbf{diags}|}{|\mathbf{edges}|}} = \sqrt{\frac{n^2 2^{n-1}}{n2^{n-1}}} = \sqrt{n}.$$

This shows that

$$c_2(D_n) \geq \sqrt{n},$$

which finishes the proof as $|D_n| = 2^n$. □

Looking back at the proof of Theorem (2.1), we see that the crucial property that allowed everything to work was the fact that ℓ_2 satisfied (1) for every embedding $f : D_n \to \ell_2$. Thus, any metric space (X, d) satisfying (1) for every embedding $f : D_n \to X$ satisfies

$$c_X(D_n) \geq \sqrt{n}.$$

For any $p > 1$, we say that a metric space has *Enflo type* p if there exists some $T > 0$ so that for every $f : D_n \to X$,

$$\sum_{\{x,y\}\in\mathbf{diags}} d(f(x), f(y))^p \leq T^p \sum_{\{u,v\}\in\mathbf{edges}} d(f(u), f(v))^p. \tag{7}$$

We let $T_p(X)$ be the best possible T such that (7) is satisfied is called the Enflo type p constant. We usually do not care about its specific value other than the fact that it exists. A superficial modification to the proof of Theorem (2.1) immediately shows that there exists some $C > 0$ depending on T and p so that

$$c_X(D_n) \geq Cn^{1-\frac{1}{p}}$$

and so $c_X(n) \geq C(\log n)^{1-\frac{1}{p}}$ also.

We make a few important observations before moving on from Enflo type.

Observe that having Enflo type p is a biLipschitz invariant, that is, if $f : (X, d_X) \to (Y, d_Y)$ is a bijective biLipschitz map between two metric spaces and one has Enflo type p, then so does the other. Letting $D \geq 1$ be the distortion of f, one can further bound the Enflo constants

$$\frac{1}{D} T_p(X) \leq T_p(Y) \leq DT_p(X). \tag{8}$$

Also, if a metric space X biLipschitz embeds into another metric space Y that has Enflo type p, then X also has Enflo type p.

Note also how the proof of the distortion lower bound comes from the statement of the property. The property gives us a ratio bound of distances in the image. The first step then is to apply the biLipschitz bounds of the embedding to translate that into a ratio bound of distances in the domain along with the distortion constant. The distortion lower bound then follows from using the geometry of the domain to estimate showing its ratio bound of distances. A more succinct way of expressing this comes from (8). One gets from Proposition 2.2 that $T_2(\ell_2) \leq 1$. One can also calculate (as we did) that $T_2(D_n) \geq \sqrt{n}$. Thus, if D is the distortion of $F : D_n \to \ell_2$, one gets

$$D \overset{(8)}{\geq} \frac{T_2(D_n)}{T_2(\ell_2)} \geq \sqrt{n}.$$

Thus, we see that this method follows the philosophy of metric embeddings described in the introduction as getting a good distortion lower bound will come from the fact that the domain's geometry does not allow for the distance ratio to be as good as that in the image. In the case of Enflo type, diagonals in Hilbert space can be much shorter than they are in D_n.

The distortion bound for Hamming cubes is one of the simplest and straightforward bounds one can derive from metric invariants. There are other metric invariant that allow you to calculate distortion bounds of other spaces, but they may not always follow so quickly and easily. The next metric invariant we discuss will also give a simple distortion bound for a different family of metric spaces. But before we introduce it this new metric invariant, we make a brief detour into nonlinear functional analysis to show how certain linear invariants can give rise to metric invariants.

3. The Ribe program

Let X be an infinite dimensional Banach space. Recall that X has Enflo type p if there exists some $T > 0$ so that for all $n \in \mathbb{N}$ and all embeddings $f : D_n \to X$, we have

$$\sum_{\{x,y\} \in \mathbf{diags}} \|f(x) - f(y)\|^p \leq T^p \sum_{\{u,w\} \in \mathbf{edges}} \|f(u) - f(v)\|^p.$$

The reason that this is called Enflo type is because it is a generalization of the linear property Rademacher type, which says that for some $T' > 0$, for any $n \in \mathbb{N}$ and $x_1, ..., x_n \in X$, we have that

$$\mathbb{E}_\varepsilon \left\| \sum_{i=1}^n \varepsilon_i x_i \right\|^p \leq T'^p \sum_{i=1}^n \|x_i\|^p. \tag{9}$$

Here, $\mathbb{E}$ is taking the expectation with respect to uniformly chosen $\varepsilon \in \{-1, 1\}^n$. To see how Enflo type p implies Rademacher type p, simply take f to be the linear function

$$f : D_n \to X$$
$$\{a_1, ..., a_n\} \mapsto \sum_{i=1}^n (2a_i - 1)x_i.$$

Pisier proved a partial converse [16] which says that if X is a Banach space with Rademacher type $p > 1$, then X has Enflo type p' for any $p' < p$. Thus, the Banach spaces with Rademacher type $p > p' > 1$ give us a rich class of Enflo type p' metric spaces. Whether Rademacher type p implies Enflo type p is still an open question.

Note that Rademacher type is a linear isomorphic invariant much like Enflo type is a metric biLipschitz invariant. One can also observe that Rademacher type only depends on the finite dimensional linear substructure of X. Indeed, the defining inequality (9) only needs to be verified on finite subsets of vectors. Such isomorphic properties that depend only on the finite dimensional substructure of a Banach space are called *local* properties.

We recall that a Banach space X is said to be finitely representable in another Banach space Y if there exists some $K < \infty$ so that for each finite dimensional subspace $Z \subset X$, there exists some subspace $Z' \subset X$ so that $d_{BM}(Z, Z') \leq K$. Here, d_{BM} is the Banach-Mazur distance. We have thus shown that local properties are invariant under finitely representability. Examples of local properties include Rademacher type, Rademacher cotype, superreflexivity, uniform convexity, and uniform smoothness.

Ribe proved in [17] the following theorem, which gives a sufficient condition for mutual finite representability.

Theorem 3.1. *Let X and Y be infinite dimensional separable Banach spaces that are uniformly homeomorphic. Then X and Y are mutually finitely representable.*

Ribe's theorem should be compared to the Mazur-Ulam theorem [12], which shows that any bijective isometry between Banach spaces is affine, and Kadets's theorem [7], which states that any two separable Banach spaces are homeomorphic. These two theorems state that the super-rigid world of isometries and the super-relaxed world of homeomorphisms are completely trivial for completely opposite reasons when applied to Banach spaces. Thus, Ribe's theorem states that interesting phenomena exist inbetween these two extremes.

Thus, if two Banach spaces are equivalent in some metrically quantitative way (as expressed by the modulus of continuity for the uniform homeomorphism), then their finite dimensional linear substructures are isomorphically equivalent. In particular, uniform homeomorphisms preserve local properties.

Thus, as uniform homeomorphisms only deal with a Banach space's metric structure, Ribe's theorem suggests that local properties may be recast in purely metric terms. This endeavor to do so is called the Ribe program and has produced a great number of metric invariants, including Enflo type (although Enflo type predates Ribe's theorem). The next metric invariant we will cover is Markov p-convexity, which characterized the local property of having a modulus of uniform convexity of power type p. We will not go into any more details about the rest of the Ribe program, but we refer the interested reader to the surveys [1, 14] and the references that lie therein.

4. Markov convexity

Recall that a Banach space X is said to be uniformly convex if for every $\varepsilon > 0$, there exists some $\delta = \delta(\varepsilon) > 0$ so that for any $x, y \in X$ such that $\|x\|, \|y\| \leq 1$ and $\|x - y\| < \delta$, we have

$$1 - \left\| \frac{x+y}{2} \right\| < \varepsilon.$$

As the name suggests, the unit ball of a uniformly convex Banach space is convex in a uniform fashion. Here, $\delta(\varepsilon)$ is called the modulus of uniform convexity. A Banach space is said to be uniformly convex of power type $p > 1$ (or just p-convex) if there is some $C > 0$ so that the modulus satisfies

$$\delta(\varepsilon) \geq C\varepsilon^p.$$

It easily follows that a p-convex Banach space is p'-convex for all $p' > p$. It was proven in [2] that a Banach space X is p-convex if and only if there exists some $K > 0$ and an equivalent norm $|\cdot|$ so that

$$|x|^p + |y|^p \geq 2\left|\frac{x-y}{2}\right|^p + 2\left|\frac{x+y}{2K}\right|^p, \qquad \forall x, y \in X. \tag{10}$$

This is a one-sided parallelogram inequality with power p. We will use this characterization of p-convex Banach spaces from now on.

Pisier prove in [15] the striking fact that any uniformly convex Banach spaces can be renormed to be p-convex for some $p \in [2, \infty)$. He also proved that p-convexity is preserved by isomorphism and so is actually a local property as it only depends on 2-subspaces of X. Thus, the Ribe program suggests that there is a metric invariant characterizing p-convexity. In [10], the authors introduced a metric invariant known as Markov p-convexity that was shown to be implied by p-convexity. In [13], the authors completed the characterization by showing that any Banach space that was Markov p-convex had an equivalent norm that was p-convex. Before we describe Markov p-convexity, we first must establish some notation.

Given some Markov chain $\{X_t\}_{t\in\mathbb{Z}}$ on a state space Ω and some integer $k \in \mathbb{Z}$, we let $\{\widetilde{X}_t(k)\}_{k\in\mathbb{Z}}$ denote the Markov chain on Ω so that for $t \leq k$, $\widetilde{X}_t(k) = X_t$ and for $t > k$, $\widetilde{X}_t(k)$ evolves independently (but with respect to the same transition probabilities) to X_t. We never specify that the Markov chain has to be time homogeneous. We can now describe Markov p-convexity.

Let $p > 0$. We say that a metric space (X, d) is Markov p-convex if there exists some $\Pi > 0$ so that for every $f : \Omega \to (X, d)$ and every Markov chain $\{X_t\}_{t\in\mathbb{Z}}$ on Ω,

$$\sum_{k=0}^{\infty}\sum_{t\in\mathbb{Z}} \frac{\mathbb{E}\left[d\left(f(X_t), f\left(\widetilde{X}_t(t-2^k)\right)\right)^p\right]}{2^{kp}} \leq \Pi^p \sum_{t\in\mathbb{Z}} \mathbb{E}\left[d(f(X_t), f(X_{t-1})^p\right]. \tag{11}$$

It follows immediately that this is indeed a biLipschitz metric invariant.

The full proof of the equivalence of Markov p-convexity with p-convexity is beyond the scope of these notes. We will just prove the easy direction of p-convexity implying Markov p-convexity later. First, we will try to make sense of exactly what Markov convexity is saying. For this, it will be more illuminating to see what spaces are *not* Markov convex.

Markov p-convexity says in essence that independent Markov chains do not drift too far apart compared to how far they travel at all places and all scales. An example of a simple metric space that does not satisfy this property—and the one that motivated the definition of Markov convexity—are complete binary trees. Indeed, the branching nature of trees allow for Markov chains to diverge linearly.

Let $\{X_t\}_{t\in\mathbb{Z}}$ be the standard downward random walk on B_n, the complete binary tree of depth n, where each branching is taken independently with probability $1/2$ and the walk stops completely after it reaches a leaf (thus, B_n is our state space). We can set X_t to be at the root for $t \leq 0$. Here, we are using time inhomogeneity.

Proposition 4.1. *Let X_t be the random walk as described above on B_n. Then there exists some constant $C > 0$ depending only on p so that*

$$\sum_{k=0}^{\infty} \sum_{t\in\mathbb{Z}} \frac{\mathbb{E}\left[d\left(X_t, \widetilde{X}_t(t-2^k)\right)^p\right]}{2^{kp}} \geq C \log n \sum_{t\in\mathbb{Z}} \mathbb{E}\left[d(X_t, X_{t-1})^p\right].$$

Proof. From the description of the random walk, we can easily compute

$$\sum_{t\in\mathbb{Z}} \mathbb{E}\left[d(X_t, X_{t+1})^p\right] = n. \tag{12}$$

To compute the left hand side of (11), we have when $k \in \left\{0, ..., \left\lfloor \frac{1}{2}\log n \right\rfloor\right\}$ and $t \in \{2^k, ..., n\}$ that

$$\mathbb{E}\left[d\left(X_t, \widetilde{X}_t(t-2^k)\right)^p\right] = 2^{p-1}2^{kp}.$$

Indeed, this is simply because there is a 1/2 chance that X_{t-2^k+1} and $X_{t-2^k+1}(t-2^k)$ are different in which case X_t and $X_t(t-2^k)$ would differ by 2^{k+1}. Thus, we have the lower bound

$$\sum_{k=0}^{\left\lfloor \frac{1}{2}\log n \right\rfloor} \sum_{t=2^k}^{n} \frac{\mathbb{E}\left[d\left(X_t, \widetilde{X}_t(t-2^k)\right)^p\right]}{2^{kp}} = \sum_{k=0}^{\left\lfloor \frac{1}{2}\log n \right\rfloor} \sum_{t=2^k}^{n} 2^{p-1} \geq Cn \log n, \tag{13}$$

where $C > 0$ is some constant depending only on n. By (12) and (13), we have finish the proof. ☐

We can now prove the following theorem.

Theorem 4.2. *Let $p > 1$ and suppose (X, d) is a metric space that is Markov p-convex. Then there exists some $C > 0$ depending only on X so that $c_X(B_n) \geq C(\log\log|B_n|)^{1/p}$.*

Proof. Let X_t be the random walk on B_n as described above. Let $f : B_n \to X$ be a Lipschitz map with distortion D. Then we have by definition of Markov p-convexity and distortion that there exists some $\Pi > 0$ so that

$$s^p \sum_{k=0}^{\infty} \sum_{t\in\mathbb{Z}} \frac{\mathbb{E}\left[d\left(X_t, \widetilde{X}_t(t-2^k)\right)^p\right]}{2^{kp}} \leq \sum_{k=0}^{\infty} \sum_{t\in\mathbb{Z}} \frac{\mathbb{E}\left[d\left(f(X_t), f\left(\widetilde{X}_t(t-2^k)\right)\right)^p\right]}{2^{kp}}$$

$$\overset{(11)}{\leq} 2^{p-1}\Pi^p \sum_{t\in\mathbb{Z}} \mathbb{E}\left[d(f(X_t), f(X_{t-1}))^p\right] \leq 2^{p-1}D^p s^p \Pi^p \sum_{t\in\mathbb{Z}} \mathbb{E}\left[d(X_t, X_{t-1})^p\right].$$

Appealing to Proposition 4.1, we see that we must have $D \geq \frac{(2\log n)^{1/p}}{2\Pi}$, which establishes the claim once we remember that $|B_n| = 2^{n+1} - 1$. ☐

We now prove the following theorem.

Theorem 4.3. *Let $p \in [2, \infty)$ and let X be a p-convex Banach space. Then X is Markov p-convex.*

As an immediate corollary of Theorems 4.2 and 4.3, we get

Corollary 4.4. *Let X be a p-convex Banach space. Then there exists some constant $C > 0$ depending only on X so that $c_X(B_n) \geq C(\log n)^{1/p}$.*

We will follow the proof of Theorem 4.3 as done in [13]. We first need the following fork lemma.

Lemma 4.5. *Let X be a Banach space whose norm $|\cdot|$ satisfies* (10). *Then for every $x, y, z, w \in X$,*

$$\frac{|x-w|^p + |x-z|^p}{2^{p-1}} + \frac{|z-w|^p}{4^{p-1}K^p} \leq |y-w|^p + |z-y|^p + 2|y-x|^p. \tag{14}$$

Proof. We have by (10) that for every $x, y, z, w \in X$ that

$$|y-x|^p + |y-w|^p \geq \frac{|x-w|^p}{2^{p-1}} + \frac{2}{K^p}\left|y - \frac{x+w}{2}\right|^p,$$

$$|y-x|^p + |y-z|^p \geq \frac{|x-z|^p}{2^{p-1}} + \frac{2}{K^p}\left|y - \frac{x+z}{2}\right|^p.$$

Thus, adding these two inequalities together and using convexity to $|\cdot|^p$, we get

$$\begin{aligned} 2|y-x|^p + |y-z|^p + |y-w|^p &\geq \frac{|x-w|^p + |x-z|^p}{2^{p-1}} + \frac{2}{K^p}\left|y - \frac{x+w}{2}\right|^p + \frac{2}{K^p}\left|y - \frac{x+z}{2}\right|^p \\ &\geq \frac{|x-w|^p + |x-z|^p}{2^{p-1}} + \frac{|z-w|^p}{4^{p-1}K^p}. \end{aligned}$$

□

This lemma says that the tips of the fork z, w cannot be too far apart if $\{x, y, z\}$ and $\{x, y, w\}$ are almost geodesic. Thus, if z and w are independently evolved Markov chains, this will property essentially tells us then that they cannot diverge far.

We can now prove Theorem 4.3. The only property concerning p-convex Banach spaces we will use in the following proof is (14). However, (14) is a purely metric statement (although it is not biLipschitz invariant). Thus, the following proof shows that any metric space satisfying (14) is Markov p-convex.

Proof of Theorem 4.3. We get from (14) that for every Markov chain $\{X_t\}_{t\in\mathbb{Z}}$, $f : \Omega \to X$, $t \in \mathbb{Z}$, and $k \geq 0$ that

$$\begin{aligned} &\frac{|f(X_t) - f(X_{t-2^k})|^p + |f(\widetilde{X}_t(t-2^{k-1})) - f(X_{t-2^k})|^p}{2^{p-1}} + \frac{|f(X_t) - f(\widetilde{X}_t(t-2^{k-1})|^p}{4^{p-1}K^p} \\ &\leq |f(X_{t-2^{k-1}}) - f(X_t)|^p + |f(X_{t-2^{k-1}}) - f(\widetilde{X}_t(t-2^{k-1}))|^p + 2|f(X_{t-2^{k-1}}) - f(X_{t-2^k})|^p. \end{aligned}$$

Note that $(X_{t-2^k}, \widetilde{X}_t(t-2^{k-1}))$ and (X_{t-2^k}, X_t) have the same distribution by definition of $\widetilde{X}_t(t-2^{k-1})$. Thus, taking expectation, we get that

$$\begin{aligned} &\frac{\mathbb{E}\left[|f(X_t) - f(X_{t-2^k})|^p\right]}{2^{p-2}} + \frac{\mathbb{E}\left[|f(X_t) - f(\widetilde{X}_t(t-2^{k-1})|^p\right]}{4^{p-1}K^p} \\ &\qquad \leq 2\mathbb{E}\left[|f(X_{t-2^{k-1}}) - f(X_t)|^p\right] + 2\mathbb{E}\left[|f(X_{t-2^{k-1}}) - f(X_{t-2^k})|^p\right]. \end{aligned}$$

We divide this inequality by $2^{(k-1)p+2}$ to get

$$\frac{\mathbb{E}\left[|f(X_t)-f(X_{t-2^k})|^p\right]}{2^{kp}}+\frac{\mathbb{E}\left[|f(X_t)-f(\widetilde{X}_t(t-2^{k-1})|^p\right]}{2^{(k+1)p}K^p} \leq \frac{\mathbb{E}\left[|f(X_{t-2^{k-1}})-f(X_t)|^p\right]}{2^{(k-1)p+1}}+\frac{\mathbb{E}\left[|f(X_{t-2^{k-1}})-f(X_{t-2^k})|^p\right]}{2^{(k-1)p+1}}.$$

Sum this inequality over $k=1,...,m$ and $t\in\mathbb{Z}$ to get

$$\begin{aligned}\sum_{k=1}^{m}\sum_{t\in\mathbb{Z}}\frac{\mathbb{E}\left[|f(X_t)-f(X_{t-2^k})|^p\right]}{2^{kp}}&+\sum_{k=1}^{m}\sum_{t\in\mathbb{Z}}\frac{\mathbb{E}\left[|f(X_t)-f(\widetilde{X}_t(t-2^{k-1})|^p\right]}{2^{(k+1)p}K^p}\\ &\leq \sum_{k=1}^{m}\sum_{t\in\mathbb{Z}}\frac{\mathbb{E}\left[|f(X_{t-2^{k-1}})-f(X_t)|^p\right]}{2^{(k-1)p+1}}+\sum_{k=1}^{m}\sum_{t\in\mathbb{Z}}\frac{\mathbb{E}\left[|f(X_{t-2^{k-1}})-f(X_{t-2^k})|^p\right]}{2^{(k-1)p+1}}\\ &=\sum_{j=0}^{m-1}\sum_{t\in\mathbb{Z}}\frac{\mathbb{E}\left[|f(X_t)-f(X_{t-2^j})|^p\right]}{2^{jp}}. \end{aligned} \tag{15}$$

By the triangle inequality, we have that

$$\sum_{t\in\mathbb{Z}}\frac{\mathbb{E}\left[|f(X_t)-f(X_{t-2^j})|^p\right]}{2^{jp}}\leq\sum_{t\in\mathbb{Z}}\mathbb{E}\left[|f(X_t)-f(X_{t+1})|^p\right].$$

We can clearly assume that $\sum_{t\in\mathbb{Z}}\mathbb{E}\left[|f(X_t)-f(X_{t+1})|^p\right]<\infty$ as otherwise the statement of the proposition is trivial. Thus, we have that the summation on the right hand side of (15) is finite for every $m\geq 1$. We can thus subtract the left hand side from the right hand side in (15) to get

$$\begin{aligned}&\sum_{k=1}^{m}\sum_{t\in\mathbb{Z}}\frac{\mathbb{E}\left[|f(X_t)-f(\widetilde{X}_t(t-2^{k-1})|^p\right]}{2^{(k+1)p}K^p}\\ &\leq\sum_{t\in\mathbb{Z}}\mathbb{E}\left[|f(X_t)-f(X_{t+1})|^p\right]-\sum_{t\in\mathbb{Z}}\frac{\mathbb{E}\left[|f(X_t)-f(X_{t-2^m})|^p\right]}{2^{mp}}\leq\sum_{t\in\mathbb{Z}}\mathbb{E}\left[|f(X_t)-f(X_{t+1})|^p\right].\end{aligned}$$

This is the same as the following inequality

$$\sum_{k=0}^{m-1}\sum_{t\in\mathbb{Z}}\frac{\mathbb{E}\left[|f(X_t)-f(\widetilde{X}_t(t-2^{k})|^p\right]}{2^{kp}}\leq(4K)^p\sum_{t\in\mathbb{Z}}\mathbb{E}\left[|f(X_t)-f(X_{t+1})|^p\right].$$

Taking $m\to\infty$ then finishes the proof. □

Corollary 4.4 was first proven by Matoušek in [11] using a metric differentiation argument. The result of [11] was itself a sharpening of a result of Bourgain in [5] which says that the finite complete binary trees embed with uniformly bounded distortion into a Banach space X if and only if X is *not* isomorphic to any uniformly convex space. This is actually the first result of the Ribe program giving a metrical characterization of the local property of a space being isomorphic to a uniformly convex space (also called superreflexivity, although this was not the original formulation). It is now known that the statement of Bourgain also holds with the infinite complete binary tree [3].

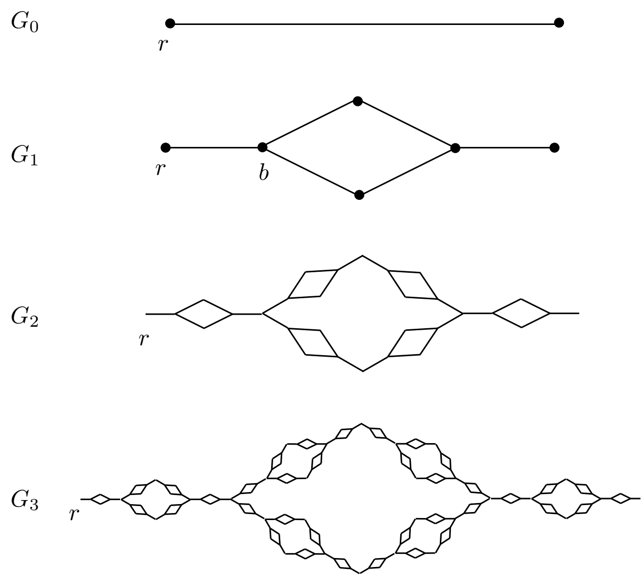


FIGURE 1. The first four Laakso graphs

Another class of metric spaces that lend well for using Markov convexity to estimate distortion bounds are the Laakso graphs. Laakso graphs were described in [8]. We define the graphs $\{G_k\}_{k=0}^{\infty}$ as follows. The first stage $G_0 = 0$ is simply an edge and G_1 is as pictured in Figure 1. To get G_k from G_{k-1}, one replaces all the edges of G_{k-1} with a copy of G_1. The metric is the shortest path metric. For each G_k, let r be the left-most vertex as shown in Figure 1. We can define the random walk $\{X_t\}_{t\in\mathbb{Z}}$ on each G_k where $X_t = r$ for $t \le 0$ and for $t > 0$, X_t is the standard rightward random walk along the graph of G_k where each branch is taken independently with probability 1/2. Once X_t hits the right-most vertex (at $t = 6^n$), it stays there forever.

We have the following proposition.

Proposition 4.6. *Let G_n be the Laakso graphs of stage n and let X_t be the random walk on G_n as described above. Then there exists some constant $C > 0$ depending only on p so that*

$$\sum_{k=0}^{\infty}\sum_{t\in\mathbb{Z}} \frac{\mathbb{E}\left[d\left(X_t, \widetilde{X}_t(t-2^k)\right)^p\right]}{2^{kp}} \ge Cn \sum_{t\in\mathbb{Z}} \mathbb{E}\left[d(X_t, X_{t-1})^p\right].$$

The proof is similar to the proof of Proposition 4.1 although it does require a little more work. The reader can either attempt to prove it as an exercise or consult Proposition 3.1 of [13].

Analogous to Theorem 4.2, we get the following distortion bounds for embeddings of G_n into Markov p-convex metric spaces:

Theorem 4.7. *Let $p > 1$ and suppose (X, d) is a metric space that is Markov p-convex. Then there exists some $C > 0$ depending only on X so that $c_X(G_n) \ge C(\log|G_n|)^{1/p}$.*

REFERENCES

[1] K. Ball, *The Ribe programme*, Séminaire Bourbaki, 2012. exposé 1047.

[2] K. Ball, E. A. Carlen, and E. H. Lieb, *Sharp uniform convexity and smoothness inequalities for trace norms*, Invent. Math. **115** (1994), no. 3, 463-482.
[3] F. Baudier, *Metrical characterization of super-reflexivity and linear type of Banach spaces*, Archiv Math. **89** (2007), 419-429.
[4] J. Bourgain, *On Lipschitz embedding of finite metric spaces in Hilbert spaces*, Israel J. Math **52** (1985), no. 1-2, 46-52.
[5] ———, *The metrical interpretation of superreflexivity in Banach spaces*, Israel J. Math **56** (1986), no. 2, 222-230.
[6] P. Enflo, *On the nonexistence of uniform homeomorphisms between L_p-spaces*, Ark. Mat. **8** (1970), no. 2, 103-105.
[7] M. I. Kadets, *Proof of the topological equivalence of all separable infinite-dimensional Banach spaces*, Funct. Anal. Appl. **1** (1967), no. 1, 53-62.
[8] T. Laakso, *Plane with A_∞-weighted metric not biLipschitz embeddable to $\mathbb{R}^n$*, Bull. London Math. Soc. **34** (2002), no. 6, 667-676.
[9] N. Linial, E. London, and Y. Rabinovich, *The geometry of graphs and some of its algorithmic applications*, Combinatorica **15** (1995), 215-245.
[10] J. Lee, A. Naor, and Y. Peres, *Trees and Markov convexity*, Geom. Funct. Anal. **18** (2009), no. 5, 1609-1659.
[11] J. Matoušek, *On embedding trees into uniformly convex Banach spaces*, Israel J. Math. **114** (1999), 221-237.
[12] S. Mazur and S. Ulam, *Sur les transformations isométriques d'espaces vectoriels normés*, C. R. Math. Acad. Sci. Paris **194** (1932), 946-948.
[13] M. Mendel and A. Naor, *Markov convexity and local rigidity of distorted metrics*, J. Eur. Math. Soc. **15** (2013), no. 1, 287-337.
[14] A. Naor, *An introduction to the Ribe program*, Jpn. J. Math. **7** (2012), no. 2, 167-233.
[15] G. Pisier, *Martingales with values in uniformly convex spaces*, Israel J. Math. (1975), 326-350.
[16] ———, *Probabilistic methods in the geometry of Banach spaces*, Probability and analysis (Varenna, 1985), Lecture Notes in Math., Springer, 1986, pp. 167-241.
[17] M. Ribe, *On uniformly homeomorphic normed spaces*, Ark. Mat. **14** (1976), no. 2, 237-244.
[18] D. P. Williamson and D. B. Shmoys, *The design of approximation algorithms*, Cambridge University Press, 2011.

POINCARÉ TYPE INEQUALITIES AND NON-EMBEDDABILITIES: GROSS TRICK AND SPHERE EQUIVALENCE

MASATO MIMURA

Abstract. This report describes a rough sketch of proofs and explains the motivation of main results in the paper "Sphere equivalence, Banach expanders, and extrapolation" (in Int. Math. Res. Notices) [Mim14] by the author. Specially, we indicate some potentially use of group theory, which we call "the Gross trick", to study metric embeddings of general graphs.

1. Motivations

First we give our notation. Unless stating, we always assume the following:

- $\Gamma = (V, E)$ is a finite connected undirected graph, possibly with multiple edges and self-loops (here E is the set of oriented edges). Γ is a metric space with the path metric d_Γ (namely, $d_\Gamma(v, w)$ is the shortest length of a path connecting v and w, and set $d_\Gamma(v, v) = 0$), and $\mathrm{diam}(\Gamma)$ means the diameter (the length of largest distance).
- For $v \in V$, $\deg(v)$, the degree of v, is the number of edges which starts at v. Note that a *self-loop contributes twice* to the degree of the vertex. $\Delta(\Gamma)$ is the maximal degree $\max_{v\in V} \deg(v)$ of Γ.
- $\{\Gamma_n = (V_n, E_n)\}_n$ is a sequence of finite graphs.
- (X, p) is a pair of a Banach space X and an exponent p. *We always assume that* $p \in [1, \infty)$ (in particular, *p is always assumed to be finite.*)
- Y is also used for a Banach space. q is also used for an exponent in $[1, \infty)$.
- For $r \in [1, \infty]$ and $k \geq 1$, ℓ_r^k stands for the real ℓ_r-space of dimension k. ℓ_r means the real ℓ_r-space over an infinite countable set.
- In this report, $\tilde{X}_{(p)}$ means $\ell_p(\mathbb{N}, X)$.
- For X, $S(X)$ is the unit sphere of X.
- In a metric space L and $A, B \subseteq L$, $\mathrm{dist}(A, B)$ means the distance, namely, $\inf\{d_L(a, b) : a \in A, b \in B\}$.
- $a \precsim b$ for two nonnegative functions from the same parameter set $\mathcal{T}$ means that there exists $C > 0$ *independent of* $t \in \mathcal{T}$ such that for any $t \in \mathcal{T}$, $a(t) \leq Cb(t)$. $a \asymp b$ means both $a \precsim b$ and $a \succsim b$ hold. $a \precsim_q b$ if parameter set $\mathcal{T}$ has variable q and $C = C_q$ may depend on q.
- We write $a \precnsim b$ if $a \precsim b$ holds but $a \succsim b$ fails to be true.

1.1. Classical spectral gaps.

1.1. Classical spectral gaps. Here assume that Γ is k-regular (that means, $\deg(v) = k$ for all $v \in V$). Then the (nonnormalized) Laplacian $L(\Gamma) := kI_V - A(\Gamma)$, $A(\Gamma)$ being the adjacency matrix (the matrix $(a_{v,w})_{v,w}$ where $a_{v,w}$ is the number of edges connecting v

Date: October 31, 2014.

Key words and phrases. Expanders; Banach spectral gaps; Matoušek's extrapolation; coarse embeddings; distortions.

Supported in part by the Grant-in-Aid for Young Scientists (B), no.25800033 from the JSPS.

and w, counting self-loop twice), is a positive operator and has eigenvalues $0 = \lambda_0(\Gamma) < \lambda_1(\Gamma) \leq \lambda_2(\Gamma) \leq \cdots \leq \lambda_{|V|}(\Gamma)$. This $\lambda_1(\Gamma)$ is the *classical spectral gap* of Γ. This has a *Rayleigh quotient formula*:

$$\lambda_1(\Gamma) = \frac{1}{2} \inf_{f\colon V\to\mathbb{R}} \frac{\sum_{v\in V}\sum_{e=(v,w)\in E}|f(w)-f(v)|^2}{\sum_{v\in V}|f(v)-m(f)|^2}. \quad \cdots\ (*)$$

Here $m(f) := \sum_{v\in V} f(v)/|V|$ and f runs over all nonconstant maps.

1.2. **Banach spectral gaps.** The point in $(*)$ is that $\mathbb{R}$ has a metric and a mean structures.

Definition 1.1. For (X,p), define the the (X,p)-*spectral gap* of Γ by

$$\lambda_1(\Gamma;X,p) := \frac{1}{2} \inf_{f\colon V\to X} \frac{\sum_{v\in V}\sum_{e=(v,w)\in E}\|f(w)-f(v)\|^p}{\sum_{v\in V}\|f(v)-m(f)\|^p}. \quad \cdots\ (**)$$

Here $m(f) := \sum_{v\in V} f(v)/|V|$ and f runs over all nonconstant maps.

Example 1.2. $\lambda_1(\Gamma) = \lambda_1(\Gamma;\mathbb{R},2) = \lambda_1(\Gamma;\ell_2,2)$ (the latter equality is by Lemma 1.3). It is known that $\lambda_1(\Gamma;\mathbb{R},1)$ is proportional to $h(\Gamma)$, the (edge-)isoperimetric constant (also known as (nonnormalized) Cheeger constant) of Γ, see [Chu97, Theorem 2.5]. Here $h(\Gamma)$ is defined as $\inf\{|E(A,V\setminus A)|/|A| : 0 < |A| \leq |V|/2\}$, where $E(A,V\setminus A) := \{e=(v,w)\in E : v\in A, w\in V\setminus A\}$.

We note that Mendel and Naor [MN12] have explicitly introduced the notion of *nonlinear spectral gaps* (for the more general case where X is a metric space) and studied that in detail.

1.3. **Poincaré-type inequality.** $(**)$ is equivalent to saying the following:

$$\forall f\colon V\to X,\quad \sum_{v\in V}\|f(v)-m(f)\|^p \leq \frac{1}{\lambda_1(\Gamma;X,p)}\frac{1}{2}\sum_{v\in V}\sum_{e=(v,w)\in E}\|f(w)-f(v)\|^p. \quad \cdots\ (***)$$

This bounds the "p-variance" from below by the "p-energy" in a rough sense.

Lemma 1.3. (1) *If Y is a subspace of X, then $\lambda_1(\Gamma;Y,p) \geq \lambda_1(\Gamma;X,p)$.*
(2) $\lambda_1(\Gamma;X,p) = \lambda_1(\Gamma;\tilde{X}_{(p)},p)$.
In particular, $\lambda_1(\Gamma;\mathbb{R},p) = \lambda_1(\Gamma;\ell_p,p)$.

Proof. (1) is trivial. For (2), $\geq$ is from (1). To get $\leq$, integrate $(***)$ over $\mathbb{N}$. □

1.4. **Banach expanders.**

Definition 1.4. A sequence $\{\Gamma_n\}_{n\in\mathbb{N}}$ is called (X,p)-*anders* if the following three conditions are satisfied:

(*i*) $\sup_n \Delta(\Gamma_n) < \infty$;
(*ii*) $\lim_{n\to\infty}\operatorname{diam}(\Gamma_n) = \infty$;
(*iii*) There exists $\epsilon > 0$ such that $\inf_n \lambda_1(\Gamma_n;X,p) \geq \epsilon$.

(Classical) expanders equal $(\mathbb{R},2)$-anders, which also equal $(\mathbb{R},p)$-anders for all p by Matoušek's extrapolation (Theorem 1.16). By Lemma 1.3, they are also equal to (ℓ_p,p)-anders.

1.5. Who cares? 1: coarse embeddings.

Definition 1.5. Let (Λ, d_Λ) be a metric space. We say $f\colon \Lambda \to X$ is a *coarse embedding* if there exist a nondecreasing $\rho_-\rho_+\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with $\lim_{t\to+\infty} \rho_-(t) = +\infty$ such that for any $v, w \in \Lambda$,

$$\rho_-(d_\Lambda(v, w)) \leq \|f(v) - f(w)\|_X \leq \rho_+(d_\Lambda(v, w)).$$

This (ρ_-, ρ_+) is called a *control pair.*

For $\{\Gamma_n\}_n$ with $\lim_{n\to\infty} \operatorname{diam}(\Gamma_n) = \infty$, define a *coarse disjoint union* $\coprod_n \Gamma_n$ to be an (infinite) metric space $(\coprod_n \Gamma_n, d)$ whose point set is $\bigsqcup_n V_n$ and whose metric satisfies:

- For every n, $d\,|_{V_n\times V_n}= d_n$, where d_n denotes the original metric on Γ_n.
- For $n \neq m$, $\operatorname{dist}(V_n, V_m) \geq \operatorname{diam}(\Gamma_n) + \operatorname{diam}(\Gamma_m)$.

Theorem 1.6 (Matoušek, Gromov, Higson, et al.)**.** *Let $\{\Gamma_n\}_n$ be (X, p)-anders for some p. Then $\coprod_n \Gamma_n$ does not admit coarse embeddings into X.*

Proof. Take $\epsilon > 0$ in Definition 1.4 and $K := \epsilon^{-1}$. Suppose, in contrary, that $f\colon \coprod_n \Gamma_n \to X$ be a coarse embedding with control pair (ρ_-, ρ_+). Set $f_n := f|_{V_n}$. For considering each f_n, we may assume $m(f_n) = 0$. Then by $(***)$,

$$\begin{aligned}\frac{1}{|V_n|}\sum_{v\in V_n} \|f_n(v)\|^p &\leq \frac{1}{2|V_n|}K \sum_{v\in V_n}\sum_{e=(v,w)\in E_n} \|f_n(w) - f_n(v)\|^p\\ &\leq K\Delta(\Gamma_n)\rho_+(1)^p.\end{aligned}$$

Therefore, by letting $M = (2K \sup_n \Delta(\Gamma_n))^{1/p}\rho_+(1)$ (independent on n), we have that at least half of $v \in V_n$ satisfies $\|f_n(v)\| \leq M$. Because $\operatorname{diam}(\Gamma_n) \to \infty$, this contradicts that $\lim_{t\to+\infty} \rho_-(t) = +\infty$. □

Remark 1.7. Recently Arzhantseva and Tessera [AT14] prove the following:

Theorem 1.8 ([AT14])**.** *There exists $\{\Gamma_n\}_n$ such that*

(i) $\sup_n \Delta(\Gamma_n) < \infty$*;*
(ii) $\coprod_n \Gamma_n$ *does not admit coarse embeddings into* ℓ_2*;*
(iii) but $\coprod_n \Gamma_n$ *does not admit weak embeddings of any expanders into itself.*

Here a sequence $\{\Lambda_m\}_m$ of finite graphs is said to admit a weak embedding into a metric space Z if there exist $K > 0$ and K-Lipschitz maps $f_m\colon \Lambda_m \to Z$ such that $\lim_{m\to\infty} \sup_{v\in V(\Lambda_m)} |f_m^{-1}(f_m(v))|/|\Lambda_m| = 0.$

This shows that expanders are not the only obstruction to admitting coarse embeddings into ℓ_2. Their proof of *(ii)* employs some sorts of *relative Poincaré-type inequalities.*

1.6. Who cares? 2: distortions.

Definition 1.9. The *distortion* of Γ into X, denoted by $c_X(\Gamma)$ is defined by

$$c_X(\Gamma) := \inf\left\{C > 0 : \begin{array}{r}\exists f\colon V \to X, \exists r > 0 \text{ such that } \forall v, w \in V,\\ rd(v, w) \leq \|f(v) - f(w)\| \leq Crd(v, w)\end{array}\right\}.$$

We have $1 \leq c_{\ell_2^{|V|}}(\Gamma) \leq \operatorname{diam}(\Gamma)$. The latter estimate is obtained by the trivial embedding: $\Gamma \ni v \mapsto \delta_v \in \ell_2(V)$. Hence, by the Dvoretzy theorem, for infinite dimensional X, we have

$$1 \leq c_X(\Gamma) \precsim_X \operatorname{diam}(\Gamma).$$

Theorem 1.10 (Generalized Grigorchuk–Nowak inequality, see [GN12] and Theorem 2.3 of [Mim14])**.** *For any* $\epsilon \in (0,1)$*,*

$$c_X(\Gamma) \geq \frac{(1-\epsilon)^{1/p} r_\epsilon(\Gamma)}{2} \mathrm{diam}(\Gamma) \left(\frac{\lambda_1(\Gamma; X, p)}{\Delta(\Gamma)} \right)^{1/p}.$$

Here $r_\epsilon(\Gamma)$ *is defined as* $\inf\{\mathrm{diam}(A)/\mathrm{diam}(\Gamma) : |A| \geq \epsilon |V|\}$.

Theorem 1.11 (Special case of a generalized Jolissaint–Valette inequality, see [JV14] and Theorem 2.3 of [Mim14])**.** *Let* Γ *be a vertex-transitive graph (this means that the graph automorphism group acts* V *transitively). Then*

$$c_X(\Gamma) \geq 2^{-(p-1)/p} \mathrm{diam}(\Gamma) \left(\frac{\lambda_1(\Gamma; X, p)}{\Delta(\Gamma)} \right)^{1/p}.$$

Note that, as we will recall in Section 3, all Cayley graphs are vertex-transitive.

Corollary 1.12. *For infinite dimensional* X*, assume* $\{\Gamma_n\}_n$ *be* (X,p)*-anders for some* p*. Then* $c_X(\Gamma_n) \asymp_X \mathrm{diam}(\Gamma_n)$*.*

Proof. Note that $\{\Gamma_n\}_n$ is in particular a family of expanders (see (3) of Corollary 1.17) and is of (uniformly) exponential growth. If you do not know this fact, then this is deduced from the Matoušek extrapolation (Theorem 1.16) and Example 1.2 on isoperimetric constants.

Hence the conclusion follows from Theorem 1.10 and the discussion above. □

Lemma 1.13 (Austin's lemma [Aus11], see also in Lemma 2.7 in [Mim14])**.** *Let* $\{\Gamma_n\}_n$ *satisfy* $\mathrm{diam}(\Gamma_n) \nearrow \infty$ *(possibly with* $\sup_n \Delta(\Gamma_n) = \infty$*). Let* $\rho \colon \mathbb{R}_+ \nearrow \mathbb{R}_+$ *be a map with* $\lim_{t\to+\infty} \rho(t) = +\infty$ *which satisfies that* $\rho(t)/t$ *is nonincreasing for* t *large enough. Assume that for* n *large enough* $\frac{\mathrm{diam}(\Gamma_n)}{\rho(\mathrm{diam}(\Gamma_n))} \precnsim c_X(\Gamma_n)$ *hold. Then for any* $C > 0$*,* (ρ, Ct) *is* not *a control pair of* $\coprod_n \Gamma_n$ *into* X*.*

Proof. Assume, in the contrary, that there exists a coarse embedding $f \colon \coprod_n \Gamma_n \to X$ such that

$$\rho(d(v,w)) \leq \|f(v) - f(w)\| C d(v,w), \quad v, w \in \coprod_n \Gamma_n$$

holds. Set $f_n := f|_{\Gamma_n} \colon V_n \to X$. We may assume, by rescaling, that f is a 1-Lipschitz map and that each f_n is biLipschitz. Then we have the following order inequalities.

$$\frac{\mathrm{diam}(\Gamma_n)}{\rho(\mathrm{diam}(\Gamma_n))} \precnsim c_X(\Gamma_n) \leq \|f_n^{-1}\|_{\mathrm{Lip}} \leq \max_{v \neq w \in V_n} \frac{d(v,w)}{\|f_n(v) - f_n(w)\|}$$
$$\precsim \max_{v \neq w \in V_n} \frac{d(v,w)}{\rho(d(v,w))} \precsim \frac{\mathrm{diam}(\Gamma_n)}{\rho(\mathrm{diam}(\Gamma_n))}.$$

This is a contradiction. □

Lemma 1.13, together with Corollary 1.12, gives an alternative proof of Theorem 1.6. Indeed, suppose, in contrary, that there exists a coarse embedding f of (X,p)-anders into X. By rescaling, we may assume that the control pair for f is (ρ, t) for some ρ (note that because $\coprod_n \Gamma_n$ is uniformly discrete, ρ_+ may be taken as linear function). By replacing ρ with a smaller proper function if necessary, we may also assume that $\rho(t)/t$ is nonincreasing for t large enough. Then Lemma 1.13 and Corollary 1.12 give the desired contradiction.

1.7. **Motivating problem.** A naive question on (X,p)-anders might be: *"Are any expanders are automatically (X,p)-anders for all (X,p)?"* The answer is *no.* Indeed, by the *Fréchet embedding*:

$$V_n \ni v \mapsto (d(v,w))_{w\in V_n},$$

Γ_n embeds isometrically into $\ell_\infty^{|V_n|}$. Thus if X has *trivial cotype*, then there exists a biLipschitz embedding of any $\coprod_n \Gamma_n$ into X. Here X is said to have *trivial cotype* if X contains uniformly isomorphic (in particular uniformly biLipschitz) copies of $\{\ell_\infty^n\}_n$.

The following question is a big open problem in this field:

Problem 1.14. *Are any expanders are automatically (X,p)-anders for all X of nontrivial cotype and for all p?*

In this report, we study the following two questions:

Problem 1.15. *For arbitrarily taken Γ,*
(a) estimate $\lambda_1(\Gamma;Y,p)$ from $\lambda_1(\Gamma;X,p)$;
(b) estimate $\lambda_1(\Gamma;X,q)$ from $\lambda_1(\Gamma;X,p)$.
In both cases, estimates may depend on $\Delta(\Gamma)$, but not *on $|\Gamma|$ itself.*

1.8. **previously known results.**
(b): Matoušek extrapolation

Theorem 1.16 ([Mat97])**.** (1) *For $p\in[1,2)$, $\lambda_1(\Gamma;\mathbb{R},2)^{p/2} \gtrsim_{\Delta(\Gamma),p} \lambda_1(\Gamma;\mathbb{R},p) \gtrsim_{\Delta(\Gamma),p} \lambda_1(\Gamma;\mathbb{R},2)$.*
(2) *For $p\in[2,\infty)$, $\lambda_1(\Gamma;\mathbb{R},p) \asymp_{\Delta(\Gamma),p} \lambda_1(\Gamma;\mathbb{R},2)^{p/2}$.*

Corollary 1.17. (1) *For any p, $\{\Gamma_n\}_n$ are expanders if and only if they are $(\mathbb{R},p)$-anders.*
(2) *Expanders do not admit coarse embeddings into ℓ_p for any p.*
(3) *For any (X,p), (X,p)-anders are (classical) expanders.*

Proof. (1) immediately follows. (2) is from Theorem 1.6. (3) follows from $X\supseteq\mathbb{R}$. □

(a): Pisier [Pis10]

The following definition is in [Pis10], which uses some idea of V. Lafforgue: X is said to be *uniformly curved* if $\lim_{\epsilon\to+0} D_X(\epsilon) = 0$ holds. Here $D_X(\epsilon)$ denote the infimum over those $D\in(0,\infty)$ such that for every $n\in\mathbb{N}$, every matrix $T=(t_{ij})_{i,j}\in \mathrm{M}_n(\mathbb{R})$ with

$$\|T\|_{\ell_2^n\to\ell_2^n}\le\epsilon \quad\text{and}\quad \|\mathrm{abs}(T)\|_{\ell_2^n\to\ell_2^n}\le 1,$$

where $\mathrm{abs}(T)=(|t_{ij}|)_{i,j}$ is the entrywise absolute value of T, satisfies that

$$\|T\otimes I_X\|_{\ell_2(n,X)\to\ell_2(n,X)}\le D.$$

Theorem 1.18 ([Pis10])**.** *Expanders are automatically $(X,2)$-anders for any uniformly curved Banach space X.*

Expamles of uniformly curved Banach spaces are ℓ_p, L_p, noncommutative L_p spaces, for $p\in(1,\infty)$, and more generally are given by complex interpolation theorey.

Remark 1.19. Pisier also showed in [Pis10] that uniformly curved Banah spaces are *superreflexive*, which is equivalent to admitting equivalent and uniformly convex norms. Recall that X is said to be *uniformly convex* if for any $\epsilon\in(0,2]$,

$$\sup\{\|x+y\|/2 : x,y\in S(X),\ \|x-y\|\ge\epsilon\}<1.$$

We also mention that the existence of "special expanders", which have the expander property for a wider class of Banach spaces, is known independently by V. Lafforgue [Laf08] and Mendel–Naor [MN12]:

Theorem 1.20 ([Laf08], [MN12]). *There exist (explicitly constructed) expanders* $\{\Gamma_n\}_n$ *which are* $(X, 2)$*-anders for any* X *of nontrivial type.*

Here recall that X has *trivial type* if and only if X contains uniformly isomorphic copies of $\{\ell_1^n\}_n$.

2. Main results

2.1. Sphere equivalence and Ozawa's result.

In [Mim14], the author call the following equivalence the *sphere equivalence*. This has been intensively studied for several years, and we refer the reader to Chapter 9 of [BL00].

Definition 2.1. X and Y are said to be *sphere equivalent*, written as $X \sim_S Y$, if there exists a uniform homeomorphism (, namely, a biuniformly continuous map) between $S(X)$ and $S(Y)$. We write $[Y]_S$ for the sphere equivalence class of Y.

If X and Y are isomorphic (in other words, if Y has an equivalent norm to that of X), then clearly $X \sim_S Y$. There, however, exist many nonisomorphic Banach spaces which are sphere equivalent.

Example 2.2. The sphere equivalence class of Hilbert spaces for instance contains the following:

- ℓ_p, L_p for any p: a uniform homeomophism is given by the *Mazur map*. For ℓ_p, the Mazur map is
$$M_{p,2}\colon S(\ell_p) \to S(\ell_2); \quad (a_i)_i \mapsto (\mathrm{sign}(a_i)|a_i|^{p/2})_i.$$
- Noncommutative L_p spaces associated with arbitrary von Neumann algebras [Ray02].
- Any Banach space of nontrivial cotype with unconditional basis [OS94].

Note that this sphere equivalence may go beyond superreflexivity; and moreover having nontrivial type. Indeed, the results mentioned above on (noncommutative) L_p spaces hold even for $p = 1$.

Example 2.3. Another example is given by complex interpolations (for a comprehensive treatise of complex interpolation, see a book [BL76]). Theorem 9.12 in [BL00] states that for a complex interpolation pair (X_0, X_1), if either X_0 or X_1 is uniformly convex, then any $0 < \theta < \theta' < 1$, $X_\theta \sim_S X_{\theta'}$. This result will be used for the proof of our main results.

On (a) of Problem 1.15, Ozawa [Oza04] made the first contribution.

Theorem 2.4 ([Oza04]). *If* $X \sim_S \ell_2$*, then expanders do not admit coarse embeddings into* X*. In fact, any expanders satisfy a weak form of* $(X, 1)$*-ander condition for such* X*.*

2.2. Main results.

Here we exhibit main results in this report, extracted from [Mim14].

Theorem A (For more precise statement, see Theorem 4.1 in [Mim14]). *Assume* $X \sim_S Y$*. Then for any* $p \in [1, \infty)$*, and a sequence* $\{\Gamma_n\}_n$*,* $\{\Gamma_n\}_n$ *are* (X, p)*-anders if and only if they are* (Y, p)*-anders.*

More precisely, for a uniform homeomorphism $\phi\colon S(X) \to S(Y)$*, we may bound* $\lambda_1(\Gamma; X, p)$ *from below in terms of*

- $\lambda_1(\Gamma; Y, p)$;
- *the modulus of continuity of* ϕ;
- *and some constants depending on* p, $\Delta(\Gamma)$, *and the modulus of continuity of* ϕ^{-1}.

For instance, if ϕ *is* α*-Hölder continuous for some* $\alpha \in (0, 1]$, *then we have*

$$\lambda_1(\Gamma; X, p) \gtrsim_{p,\Delta(\Gamma),M} \lambda_1(\Gamma; Y, p)^{1/\alpha}.$$

Here M *is a constant only depend on the modulus of continuity of* ϕ^{-1}.

Note that on the estimation above, the order of the estimate (for instance, the Hölder exponent if we have an estimate of such type) depends *only on* the modulus of continuity of ϕ. The modulus of continuity of the inverse map ϕ^{-1} appears only on postive scalar constant in our estimate.

Theorem B (Generalization of Matoušek's extrapolation)**.** *Let* $(\infty >)p, q > 1$. *Then for any* X *sphere equivalent to a uniformly convex Banach space, and a sequence* $\{\Gamma_n\}_n$, $\{\Gamma_n\}_n$ *are* (X, p)*-anders if and only if they are* (X, q)*-anders.*

Remark 2.5. We note that recently Naor, in Theorem 1.10 and Theorem 4.15 in [Nao14], has independently established similar results. Our approach is *group theoretic*, and different from his. In our proof, we introduce the *"Gross trick"*, see Section 6.

As byproducts to above Theorems A and B, and aforementioned works of Ozawa and Pisier; and Lafforgue and Mendel–Naor, we have the following corollaries.

Corollary C. *Any expanders are automatically* (X, p)*-anders for an* X *sphere equivalent to uniformly curved Banach space and for* $p \in (1, \infty)$. *If, moreover,* $X \in [\ell_2]_S$, *then the assertion above holds even for* $p = 1$.

In particular, for expanders $\{\Gamma_n\}_n$, *we have for such* X *of infinite dimension that*

$$c_X(\Gamma_n) \asymp_X \mathrm{diam}(\Gamma_n).$$

Corollary D. *The expanders constructed in Theorem* 1.20 *are* $(Y, 2)$*-anders for any* Y *sphere equivalent to a Banach space with nontrivial type.*

In particular, they do not admit coarse embedding into any such Y.

Note that, for instance, noncommutative L_1 spaces are examples of such Y with trivial type (though all expanders do not admit coarse embeddings to them by Theorem 2.4).

In the view of the results above, the following questions might be of importance.

Problem 2.6. (1) *Does the class of Banach spaces sphere equivalent to uniformly curved Banach spaces contain all superreflexive Banach spaces? Does it contain all Banach spaces of nontrivial type/nontrivial cotype?*
(2) *Does the class of Banach spaces sphere equivalent to Banach spaces of nontrivial type coincide with the class of all Banach spaces of nontrivial cotype?*

To the best of my knowledge, all of the problems above may be open.

Remark 2.7. On (2), one inclusion is verified from Corollary D and Subsection 1.7 (also, in [BL00], the authors of the book announced a result that the class of Banach spaces with trivial cotype is closed under the sphere equivalence). Hence, the true question in (2) is whether the sphere equivalence class above contain all Banach spaces of nontrivial cotype.

Remark 2.8. There is also a notion of "*ball equivalence*" (namely, the unit balls are uniformly homeomorphic). In [BL00, Chapter 9], it is shown that if X and Y are ball equivalent, then $X \oplus \mathbb{R} \sim_S Y \oplus \mathbb{R}$ (the other direction: "$X \sim_S Y$ implies that X and Y are ball equivalent" is easy). Therefore, if we consider Banach spectral gap, then there is no serious difference between the sphere equivalence and the ball equivalence.

3. Representation theoretic constants for Cayley graphs

We first give the proof of Theorem A for *Cayley* graph of (finite) groups, and explain where group theory can contribute to this problem. In this section, let G be a finite group, $S \not\ni e$ be a symmetric (finite) generating set of G. Recall the definition of Cayley graphs. The *Cayley graph* of (G, S), written as $\mathrm{Cay}(G, S)$, is constructed as

- the vertex set $V = G$;
- and the edge set $E = \{(g, sg) : g \in G, s \in S\}$.

Example 3.1. $\mathrm{Cay}(\mathbb{Z}/n\mathbb{Z}, \{\pm 1\})$, $n \geq 3$, is the cycle of lenth n. Although we do not treat in this report, Cayley graphs are also defined for G infinite. In that case, $\mathrm{Cay}(\mathbb{Z}^2, \{\pm(1,0), \pm(0,1)\})$ is the $\mathbb{Z}^2$-lattice in $\mathbb{R}^2$. For a free group F_2 with 2 free generators a, b, $\mathrm{Cay}(F_2, \{a^{\pm 1}, b^{\pm 1}\})$ is the 4-regular tree.

Remark 3.2. Recall that a group G has two natural action on itself: the left multiplication and the right one. We have employed the left multiplication to connect edges in $\mathrm{Cay}(G, S)$, and the right one is left. In fact, this right multiplication becomes a graph automorphsim (in other words, for every $g \in G$, $(v, w) \in E$ iff $(vg, wg) \in E$). Since this right action of G on itself is transitive, $\mathrm{Cay}(G, S)$ is a *vertex-transitive* graph (it means that the automorphism group of the graph acts transitively on the vertex set). Hence, (finite) Cayley graphs are special among all (finite) graphs.

Also recall that in our notaion, we allow graphs to have self-loops and multiple edges. However, if we consider only Cayley graphs, then they do not show up.

3.1. isometric linear representations and displacement constant.

Definition 3.3. We take (G, S) and (X, p).

(1) Define $\pi_{G;X,p} = \pi_{X,p}$ as the left-regular representation of G on $\ell_p(G, \tilde{X}_{(p)})$, namely, for $g \in G$ and $\xi \in \ell_p(G, \tilde{X}_{(p)})$, $\pi_{X,p}(g)\xi(v) := \xi(g^{-1}v)$. Then $\ell_p(G, \tilde{X}_{(p)})$ decomposes as G-representation spaces: $\ell_p(G, \tilde{X}_{(p)}) = \ell_p(G, \tilde{X}_{(p)})^{\pi_{X,p}(G)} \oplus \ell_{p,0}(G, \tilde{X}_{(p)})$. Here the first space is the space of $\pi_{X,p}(G)$-invariant vectors (which consists of "constant functions" form G to $\tilde{X}_{(p)}$); and the second space is the space of "zero-sum" functions, namely,

$$\ell_{p,0}(G, \tilde{X}_{(p)}) := \{\xi \in \ell_p(G, \tilde{X}_{(p)}) : \sum_{v \in G} \xi(v) = 0\}.$$

We omit writting G in $\pi_{G;X,p}$ if G is fixed. We use the same symbol $\pi_{X,p}$ for the restricted representation on $\ell_{p,0}(G, \tilde{X}_{(p)})$.

(2) (p-displacement constant) The *p-displacement constant* of (G, S) on X, written as $\kappa_{X,p}(G, S)$, is defined as

$$\kappa_{X,p}(G, S) := \inf_{0 \neq \xi \in \ell_{p,0}(G, \tilde{X}_{(p)})} \sup_{s \in S} \frac{\|\pi_{X,p}(s)\xi - \xi\|}{\|\xi\|}.$$

Remark 3.4. We will use in the proof of Proposition 5.1 the following norm inequality: for $\xi \in \ell_{p,0}(G, \tilde{X}_{(p)})$, we have

$$\mathrm{dist}(\xi, \ell_p(G, \tilde{X}_{(p)})^{\pi_{X,p}(G)}) \geq \frac{1}{2}\|\xi\|.$$

Indeed, set $\eta = \eta_1 + \eta_0$ for any $\eta \in \ell_p(G, \tilde{X}_{(p)})$ according to the decomposition $\ell_p(G, \tilde{X}_{(p)}) = \ell_p(G, \tilde{X}_{(p)})^{\pi_{X,p}(G)} \oplus \ell_{p,0}(G, \tilde{X}_{(p)})$. Then the map $\eta \mapsto \eta_1$ is given by taking the mean of η. Because the p-mean of the norm is at least the norm of the mean, we have that $\|\eta\| \geq \|\eta_1\|$. Hence for any $\zeta \in \ell_p(G, \tilde{X}_{(p)})^{\pi_{X,p}(G)}$, we have that $\|\xi - \zeta\| \geq \|\zeta\|$ (set $\eta := \xi - \zeta$). Therefore

$$2 \cdot \inf_{\zeta \in \ell_p(G, \tilde{X}_{(p)})^{\pi_{X,p}(G)}} \|\xi - \zeta\| \geq \inf_{\zeta \in \ell_p(G, \tilde{X}_{(p)})^{\pi_{X,p}(G)}} (\|\xi - \zeta\| + \|\zeta\|) \geq \|\xi\|,$$

and we are done.

3.2. **Fundamental lemma for Banach spectral gaps of Cayley graphs.** The following lemma plays a fundamental rôle, which relates p-displacement constant on X to (X, p)-spectral gap for a Cayley graph.

Lemma 3.5. *For a Cayley graph $\Gamma = \mathrm{Cay}(G, S)$ and a pair (X, p), we have that*

$$\kappa_{X,p}(G, S)^p \leq \lambda_1(\Gamma; X, p) \leq \frac{|S|}{2}\kappa_{X,p}(G, S)^p.$$

Proof. First note that by Lemma 1.3, $\lambda_1(\Gamma; X, p) = \lambda_1(\Gamma; \tilde{X}_{(p)}, p)$. Take a nonconstant map $f\colon V \to \tilde{X}_{(p)}$ and by replacing f with $f - m(f)$ we may assume $m(f) = 0$. Then we may regard f as a nonzero vector $\xi \in \ell_{p,0}(G, \tilde{X}_{(p)})$. Therefore

$$\begin{aligned}\lambda_1(\Gamma; X, p) &= \frac{1}{2} \inf_{0 \neq \xi \in \ell_{p,0}(G, \tilde{X}_{(p)})} \frac{\sum_{v \in G} \sum_{s^{-1} \in S} \|\pi_{X,p}(s)\xi(v) - \xi(v)\|^p_{\tilde{X}_{(p)}}}{\|\xi\|^p} \\ &= \frac{1}{2} \inf_{0 \neq \xi \in \ell_{p,0}(G, \tilde{X}_{(p)})} \sum_{s \in S} \left(\frac{\|\pi_{X,p}(s)\xi - \xi\|}{\|\xi\|}\right)^p.\end{aligned}$$

This ends our proof (note that $\|\pi_{X,p}(s)\xi - \xi\| = \|\pi_{X,p}(s^{-1})\xi - \xi\|$ because $\pi_{X,p}(s)$ is an isometric operator). $\square$

Remark 3.6. If we consider $\{(G_n, S_n)\}$ where $\sup_n |S_n| < \infty$, then Lemma 3.5 gives the optimal order estimate between $\kappa_{X,p}(G_n, S_n)$ and $\lambda_1(\mathrm{Cay}(G_n, S_n); X, p)$. However if $\sup_n |S_n| = \infty$, then Lemma 3.5 may not give the precise order.

Nevertheless, if S_n's have "high symmetry", then we have more accurate inequalities. For more precise meaning, we refer the reader to [Mim14, Theorem 3.4], which is based on the work of Pak and Żuk [PZ02].

4. Key propositions on sphere equivalence

4.1. **upper moduli and** $\mathrm{Sym}(F)$ **equivariant homeomorphisms.**

Definition 4.1. Let $X \sim_S Y$, and $\phi\colon S(X) \to S(Y)$ be a uniformly continuous map.

(*i*) Define $\mathcal{M}_\phi$ to be the class of all functions $\delta\colon [0, 2] \to \mathbb{R}_{\geq 0}$ which satisfy the following three conditions:
- δ is nondecreasing;
- $\lim_{\epsilon \to +0} \delta(\epsilon) = 0$;

- and, for any $x_1, x_2 \in S(X)$ with $\|x_1 - x_2\|_X \leq \epsilon$, we have $\|\phi(x_1) - \phi(x_2)\|_Y \leq \delta(\epsilon)$.

We call an element δ in $\mathcal{M}_\phi$ an *upper modulus of continuity* of ϕ.

(*ii*) Define $\overline{\phi}\colon X \to Y$ to be the extension of ϕ by homogeneity, namely, $\overline{\phi}(x) := \|x\|_X \phi(x/\|x\|_X)$ for $0 \neq x \in X$ and $\overline{\phi}(0) := 0$. We call $\overline{\phi}$ the *natural extension* of ϕ.

Note that $\overline{\phi}$ is uniformly continuous if we restrict it on a bounded set of X; but that itself is in general *not*.

Example 4.2. In Example 2.2, we have seen the definition of the Mazur map $M_{p,2}\colon S(\ell_p) \to S(\ell_2)$. This map (and also the inverse map) is known to be uniformly continuous, more precisely,

- If $p \geq 2$, then the function $\delta\colon [0, 2] \to \mathbb{R}_{\geq 0}$; $\delta(\epsilon) := (p/2)\delta$ is in $\mathcal{M}_{M_{p,2}}$ ($M_{p,2}$ is Lipschitz).
- If $p < 2$, then the function $\delta\colon [0, 2] \to \mathbb{R}_{\geq 0}$; $\delta(\epsilon) := 4\delta^{p/2}$ is in $\mathcal{M}_{M_{p,2}}$ ($M_{p,2}$ is $p/2$-Hölder).

Surprislingly, these estimations of Hölder exponents remain to be optimal even when we consider the "noncommutative Mazur map" from noncommutative L_p spaces associated with any von Neumann algebra. This assertion has been recently showed by Ricard [Ric14].

Definition 4.3. Let F be an at most countable set. For a map $\phi\colon S(\ell_p(F, X)) \to S(\ell_q(F, Y))$, we say that ϕ is $\mathrm{Sym}(F)$-*equivariant* if for any $\sigma \in \mathrm{Sym}(F)$, $\phi \circ \sigma_{X,p} = \sigma_{Y,q} \circ \phi$ holds true. Here a Banach space Z and $r \in [1, \infty)$, the symbol $\sigma_{Z,r}$ denotes the isometry $\sigma_{Z,r}$ on $\ell_r(F, Z)$ induced by σ, namely, $(\sigma_{Z,r}\xi)(a) := \xi(\sigma^{-1}(a))$ for $\xi \in \ell_r(F, Z)$ and $a \in F$. Here by $\mathrm{Sym}(F)$ we mean the group of all permutations on F, *including ones of infinite supports.*

For instance, if we consider the Mazur map $M_{p,2}$ as a map from $\ell_p(\mathbb{N}, \mathbb{R})$ to $\ell_2(\mathbb{N}, \mathbb{R})$, then $M_{p,2}$ is $\mathrm{Sym}(\mathbb{N})$-equivariant. This is because $M_{p,2}$ is coordinatewise.

4.2. **Key proposition for Theorem A.**

Proposition 4.4. *Assune that $\phi\colon S(X) \to S(Y)$ is a uniformly continuous map for two Banach spaces X and Y. Then for any $p \in [1, \infty)$, the map*

$$\Phi = \Phi_p\colon S(\tilde{X}_{(p)}) \to S(\tilde{Y}_{(p)}); \quad (x_i)_i \mapsto (\overline{\phi}(x_i))_i$$

is again a uniformly continuous map that is $\mathrm{Sym}(\mathbb{N})$*-equivariant. Here $\overline{\phi}$ is the natural extension of ϕ and we see $\tilde{X}_{(p)}$ and $\tilde{Y}_{(p)}$, respectively, as $\ell_p(\mathbb{N}, X)$ and $\ell_p(\mathbb{N}, Y)$.*

Furthermore, if ϕ is α-Hölder, then so is Φ_p. More precisely. if $\delta(t) := Ct^\alpha \in \mathcal{M}_\phi$ for some $C > 0$ and some $\alpha \in (0, 1]$, then $\delta'(t) := (2C + 2)t^\alpha$ belongs to $\mathcal{M}_{\Phi_p}$.

Proof. By construction, this Φ_p is coordinatewise and hence in particular $\mathrm{Sym}(\mathbb{N})$-equivariant. Our proof of the uniform continuity of Φ_p consists of two cases. Here we only prove the case where $Ct^\alpha \in \mathcal{M}_\phi$ (for general case, we may need to replace δ with larger upper modulus).

Case 1 : for $p = 1$. Let $(x_i)_i$ and $(y_i)_i$ be in $S(\tilde{X}_{(1)})$.

First we consider the case where for all $i \in \mathbb{N}$ $\|x_i\|_X = \|y_i\|_X$. Set $r_i := \|x_i\|_X$ and $\epsilon_i r_i = \|x_i - y_i\|_X$. Observe that δ is concave in $[0,2]$. By the Jensen inequality, we have the following:

$$\|\Phi((x_i)_i) - \Phi((y_i)_i)\|_{\tilde{Y}_{(1)}} \leq \sum_i r_i \delta(\epsilon_i) \leq \delta(\sum_i r_i \epsilon_i) = \delta(\|(x_i)_i - (y_i)_i\|_{\tilde{X}_{(1)}}).$$

Secondly we deal with the general case. For $(x_i)_i, (y_i)_i \in \tilde{X}_{(1)}$, define $z_i := \frac{\|x_i\|_X}{\|y_i\|_X} y_i$ ($z_i := x_i$ if $y_i = 0$). Suppose $\|(x_i)_i - (y_i)_i\|_{\tilde{X}_{(1)}} \leq \epsilon$. Because for any i, $\|x_i - y_i\|_X \geq \|z_i - y_i\|_X$, we have that $\|(z_i)_i - (y_i)_i\|_{\tilde{X}_{(1)}} \leq \epsilon$. Hence we obtain that $\|(x_i)_i - (z_i)_i\|_{\tilde{X}_{(1)}} \leq 2\epsilon$. Therefore in the first argument, we have that $\|\Phi((x_i)_i) - \Phi((z_i)_i)\|_{\tilde{Y}_{(1)}} \leq \delta(2\epsilon)$. Since $\|\Phi((y_i)_i) - \Phi((z_i)_i)\|_{\tilde{Y}_{(1)}} \leq \epsilon$ by homogeneity, we conclude that $\delta'(t) := \delta(2t) + t = 2^\alpha C t^\alpha + t$ belongs to $\mathcal{M}_{\Phi_1}$.

Case 2 : for general $p > 1$. First observe that $t \in [0, 2^{1/p}]$, we have that $\delta(t)^p \leq C^{p-1}\delta(t^p)$. Then the remaining argument goes along a similar line to one in Case 1. Thus we can show that $\delta'(t) := (C^{p-1}\delta((2t)^p))^{1/p} + t = 2^\alpha C t^\alpha + t$ belongs to $\mathcal{M}_{\Phi_p}$.

In each case, finally observe that for $t \in [0,2]$, $(2C+2)t^\alpha \geq 2^\alpha C t^\alpha + t$. ☐

Lemma 9.9 in [BL00] showed the first assertion above. However, the estimation of upper moduli is worse than in this proposition, and did not verify the latter assertions.

4.3. Generalized Mazur map: key proposition for Theorem B.

Theorem 4.5. *For any uniformly convex Banach space X and $p, q \in (1, \infty)$, we have that $\tilde{X}_{(p)} \sim_S \tilde{X}_{(q)}$. Furthermore, we may have a uniform homeomorphism $\phi\colon S(\ell_p(\mathbb{N}, X)) \to S(\ell_q(\mathbb{N}, X))$ which is* Sym($\mathbb{N}$)*-equivariant.*

Proof. Choose $1 < p_0 < \min\{p, q\}$ and $\infty > p_1 > \max\{p, q\}$. Then [BL76, Theorem 5.1.2] applies to the case where $\Omega = \mathbb{N}$ and $A_0 = A_1 = X$. This tells us that both of $\tilde{X}_{(p)}$ and $\tilde{X}_{(q)}$ are, respectively, isometrically isomorphic to some intermediate points of a complex interpolation pair $(\tilde{X}_{(p_0)}, \tilde{X}_{(p_1)})$. Because $\tilde{X}_{(p_0)}$ and $\tilde{X}_{(p_1)}$ are uniformly convex, the result mentioned in Example 2.3 applies.

The last assertion follows from the proof of [BL00, Theorem 9.12]. Indeed, the definition of f_x for $x \in \ell_p(\mathbb{N}, X)$, as the minimizer of a certain norm, in Proposition I.3 in [BL00] is Sym($\mathbb{N}$)-equivariant in the current setting. ☐

This map may be regareded as a *generalized Mazur map* because it coincide with the usual Mazur map if we consider the complex interpolation pair (ℓ_{p_0}, ℓ_{p_1}) in the proof (for $X = \mathbb{R}$). However, note that we are only able to define it for $p, q > 1$, as long as we employ the complex interpolation.

5. Proof of Theorem A for Cayley graphs

This part is based on a work of Bader–Furman–Gelander–Monod [BFGM07]. See Section 4.a in [BFGM07] for the original idea of them. We will show the following proposition concerning the p-displacement constants.

Proposition 5.1. *Let $X \sim_S Y$ and $\phi\colon S(X) \to S(Y)$ be a uniform homeomorphism. Let G be a finite group, $S \not\ni e$ be a symmetric (finite) subset. Then for any $p \in [1,\infty)$, we have the following inequality:*

$$\kappa_{X,p}(G,S) \geq \delta_1^{-1}\left(\frac{1}{2}\delta_2^{-1}\left(\frac{1}{2}\right)\kappa_{Y,p}(G,S)\right).$$

Here $\delta_1 \in \mathcal{M}_{\Phi_p}$ and $\delta_2 \in \mathcal{M}_{\Phi_p^{-1}}$.

Proof. By Proposition 4.4, $\Phi_p\colon \tilde{X}_{(p)} \to \tilde{Y}_{(p)}$ is a uniform homeomorphism that is $\mathrm{Sym}(\mathbb{N})$-equivariant. By coordinate transformation, we may regard Φ_p as

$$\Phi_p\colon S(\ell_p(G,\tilde{X}_{(p)})) \to S(\ell_p(G,\tilde{Y}_{(p)}))$$

(note that $\ell_p(G,\tilde{X}_{(p)}) \simeq \tilde{X}_{(p)}$), which is $\mathrm{Sym}(G)$-equivariant. We thus have that $\Phi_p \circ \pi_{X,p} = \pi_{Y,p}\circ\Phi_p$. Note that we consider $\pi_{X,p}$ and $\pi_{Y,p}$ as G-representations, respectively, on $\ell_p(G,\tilde{X}_p)$ and $\ell_p(G,\tilde{Y}_p)$, *not* on $\ell_{p,0}$.

Choose any $\xi \in S(\ell_{p,0}(G,\tilde{X}_{(p)})) \subseteq S(\ell_p(G,\tilde{X}_{(p)}))$ and set $\eta := \Phi_p(\xi) \in S(\ell_p(G,\tilde{Y}_{(p)}))$. We warn that η does *not* belong to $S(\ell_{p,0}(G,\tilde{Y}_{(p)}))$ in general. We however overcome this difficulty in the following argument. Recall that $\ell_p(G,\tilde{X}_{(p)})$ is decomposed as the direct sum of $\ell_p(G,\tilde{X}_{(p)})^{\pi_{X,p}(\Gamma)}$ and $\ell_{p,0}(G,\tilde{X}_{(p)})$. Note that the former subspace is sent to $\ell_p(G,\tilde{Y}_{(p)})^{\pi_{Y,p}(G)}$ by Φ_p (again because Φ_p is $\mathrm{Sym}(G)$-equivariant). Recall the inequality in Remark 3.4 and get that $\mathrm{dist}(\xi,\ell_p(G,\tilde{X}_{(p)})^{\pi_{X,p}(G)}) \geq \frac{1}{2}$.

In particular, from this we have that $\mathrm{dist}(\xi, S(\ell_p(G,\tilde{X}_{(p)})^{\pi_{X,p}(\Gamma)})) \geq \frac{1}{2}$. Therefore, by the uniform continuity of Φ_p^{-1}, we have that $\mathrm{dist}(\eta, S(\ell_p(G,\tilde{Y}_{(p)})^{\pi_{Y,p}(\Gamma)})) \geq \delta_2^{-1}\left(\frac{1}{2}\right)$.

Decompose η as $\eta = \eta_1 + \eta_0$ where $\eta_1 \in \ell_p(G,\tilde{Y}_{(p)})^{\pi_{Y,p}(G)}$ and $\eta_0 \in \ell_{p,0}(G,\tilde{Y}_{(p)})$. We claim that

$$\|\eta_0\| \geq \frac{1}{2}\delta_2^{-1}\left(\frac{1}{2}\right).$$

Indeed, let $\eta_1' := \eta_1'/\|\eta_1'\|$ (if $\eta_1 = 0$, then set η_1' as any vector in $S(\ell_p(G,\tilde{Y}_{(p)})^{\pi_{Y,p}(G)})$). Then by the inequality in the paragraph above, we have that $\|\eta-\eta_1'\| \geq \delta_2^{-1}\left(\frac{1}{2}\right)$. Because $\|\eta_1\| \geq 1-\|\eta_0\|$, we also have that $\|\eta_1-\eta_1'\| \leq \|\eta_0\|$ and that $\|\eta-\eta_1'\| \leq \|\eta-\eta_1\|+\|\eta_1-\eta_1'\| \leq 2\|\eta_0\|$. By combining these inequalities, we prove the claim.

By the definition of $\kappa_{Y,p}(G,S)$, we have that

$$\sup_{s\in S}\|\pi_{Y,p}(s)\eta-\eta\| = \sup_{s\in S}\|\pi_{Y,p}(s)\eta_0-\eta_0\| \geq \|\eta_0\|\kappa_{Y,p}(G,S) \geq \frac{1}{2}\delta_2^{-1}\left(\frac{1}{2}\right)\kappa_{Y,p}(G,S).$$

Finally, because $\Phi_p\circ\pi_{X,p} = \pi_{Y,p}\circ\Phi_p$, we conclude by the uniform continuity of Φ_p that

$$\sup_{s\in S}\|\pi_{X,p}(s)\xi-\xi\| \geq \delta_1^{-1}\left(\frac{1}{2}\delta_2^{-1}\left(\frac{1}{2}\right)\kappa_{Y,p}(G,S)\right).$$

By taking the infimum over $\xi \in S(\ell_{p,0}(G,\tilde{X}_{(p)}))$, we obtain the desired assertion. □

By combining the proposition above, Proposition 4.4, and Lemma 3.5, we obtain the conclusion in Theorem A for Γ a Cayley graph.

6. The Gross trick

In this section, we give the proof of Theorem A for Γ arbitrary finite graph. To do this, our idea is to consider *Schreier coset graphs* and to reduce all cases to these ones. The Gross theorem, which we will mention later, enables us to perform the latter procedure. The author call this trick the *Gross trick*.

6.1. **Schreier coset graph.** In the proof of Lemma 3.5 and Proposition 5.1, it may be noticed that we have *never* employed the right regular representation. This means, we only need group multiplication only on one side, which was used to connect the edges. From this observation, we encounter with the conception of *Schreier coset grpahs.*

Definition 6.1. Let G be a finitely generated group, S be a symmetric finite generating set, and H be a subgroup of G *of finite index*. By $\mathrm{Sch}(G,H,S)$ we mean the *Schreier coset graph*, that is

- the vertex set is the left cosets: $V = G/H$;
- the edge set $E := \{(gH, sgH) : gH \in G/H, s \in S\}$.

Remark 6.2. One remark is that we may take G as a finite group in the definition.

The other remark is that, unlike Cayley grpahs, Schreier coset graphs in general have *no* symmetry at all (note that only possible muliplication on G/H is from the left, but this is used for connecting edges). Moreover, in general $\mathrm{Sch}(G,H,S)$ may have self-loops and multiple edges.

Once we employ the concept of Schreier coset graphs, we have a similar definition of *p-displacement constants* for the triple (G,H,S) in terms of the quasi-regular representation of G on $\ell_{p,0}(G/H, \tilde{X}_{(p)})$. Furthermore, we have exactly the same inequalities as ones in Lemma 3.5 and Proposition 5.1 for Schreier coset graphs. In this report, we omit the precise forms. Instead, we refer the reader to Definition 3.1, Lemma 3.3, and Proposition 4.2 in [Mim14].

Thus we ends the proof of Theorem A for the case where Γ is a Shreier coset graph.

6.2. **the Gross trick.** Now we explain the main trick on the proof. This employs the following result of Gross.

Theorem 6.3 ([Gro77])**.** *Any finite connected and* regular *graph* (*possibly with multiple edges and self-loops*) *with* even *degree can be realized as a Schreier coset graph.*

Remark 6.4. The proof of Gross's theorem is based on the "2-factorization" of such a graph (Petersen). This means, for such a graph, we can decompose the (undirected) edge set as the disjoint union of 2-regular graphs (cycles). From these cycles, we can endow Γ with the structure of a Schreier coset graph. Hence this realization is not just the existence, but not sufficiently concrete or handlable in general setting.

Also, by passing to apropriate limits, the Gross theorem can be extended to infinite regular conncected graphs of even degree.

This theorem of Gross roughly asserts that Schreier coset graphs are "more or less universal" among graphs of uniformly bounded degree (compare with speciality of Cayley graphs!). More precise meaing of "universal" will be explained in the usage of "Gross trick", as below.

The following argument is the *Gross trick*: Let $\Gamma = (V,E)$ be a finite connected graph. Then we take the *even regularization* of Γ in the following sense: we let V unchanged. We first double each edge in E. Note that then for any $v, w \in V$, $\deg(v) - \deg(w) \in 2\mathbb{Z}$ and that the maximum degree is $2\Delta(\Gamma)$. Finally, we let a vertex v whose degree is $2\Delta(\Gamma)$ unchanged, and for all the other vertices add, respectively, appropriate numbers of self-loops to have the resulting degree $= 2\Delta(\Gamma)$ for each vertex. We write the resulting graph as $\Gamma' = (V, E')$. Then by the Gross theorem, Γ' can be realized as a Schreier coset graph and thus the argument in Subsection 6.1 applies to Γ'. Finally observe that $\lambda_1(\Gamma'; Z, p) = 2\lambda_1(\Gamma; Z, p)$ for any Banach space Z because self-loops do not affect the spectral gap.

This completes our proof of Theorem A for general graphs Γ.

7. Proofs of Theorem B and Corollary C

Proof of Theorem B. Let $X \sim_S Y$, where Y is uniformly convex and let $p, q \in (1, \infty)$. By Theorem 4.5, there exists an $\mathrm{Sym}(\mathbb{N})$-equivariant uniform homeomorphism $\Phi := \Phi_{p,q} \colon S(\tilde{Y}_{(p)}) \to S(\tilde{Y}_{(q)})$. First we start from the case where Γ is of the form $\mathrm{Sch}(G, H, S)$. Then we regard Φ as an $\mathrm{Sym}(G/H)$-equivariant uniform homeomorphism

$$\Phi \colon S(\ell_p(G/H, \tilde{Y}_{(p)})) \to S(\ell_q(G/H, \tilde{Y}_{(q)})).$$

We thus may apply a similar argument to Proposition 5.1 to the pair $((Y, p); (Y, q))$. Because Proposition 5.1 works for the pairs $((X, p); (Y, p))$ and $((Y, q); (X, q))$, we are done.

For general cases, apply the Gross trick. □

Proof of Corollary C. The first assertion holds true by Theorem A, Theorem B, and the fact of that uniformly curved Banach spaces are isomorphic (and in particular sphere equivalent) to some uniformly convex Banach spaces, see Remark 1.19. The second assertion holds true for the following reason: if $X \in [\ell_2]_S$, then by Theorem A and Lemma 1.3, the (X, p)-ander property is equivalent to the $(\mathbb{R}, p)$-ander property. The original Matoušek extrapolation enables us to extend our results even for $p = 1$. □

8. Application: embeddings of Hamming graphs into noncommutative L_p spaces

As an application of our main results, we consider embeddings of Hamming graphs into noncommutative L_p spaces associated with arbitrary von Neumann algebras. For $d \geq 1$ and $k \geq 2$, the *Hamming graph* $H(d, k)$ is defined as the following:

- the vertex set V is the set of the ordered d-tuples of T, $|T| = k$;
- the edge set E consists of all pairs which diffres in precisely one coordinate.

In other words, $H(d, k)$ is the product of d copies of the complete graph K_k on k vertices. It is easy to see that $H(d, k)$ is $d(k-1)$-regular and $\mathrm{diam}(H(d, k)) = d$. As a byproduct of Theorem A, we have the following:

Theorem 8.1. *Let $\mathcal{M}$ be a von Neumann algebra. By $L_p(\mathcal{M})$, we denote the noncommutative L_p space associated with $\mathcal{M}$.*

(1) *For $p \in [1, 2)$, then we have that $\lambda_1(H(d, k); L_p(\mathcal{M}), p) \asymp_p k$.*
(2) *For $p \in [2, \infty)$, then we have that $\lambda_1(H(d, k); L_p(\mathcal{M}), 2) \asymp_p k$.*

Note that the multiplicative constants in these estimation do *not* depend on d, k, and $\mathcal{M}$; and only depend on p.

Proof. We only prove the case where k is a prime number. For other cases, we use a similar technique to the Gross trick (namely, we add multiple edges and self-loops to have better graph) in order to apply [Mim14, Theorem 3.4].

Note that $H(d,k) = \mathrm{Cay}(G_{d,k}, S_{d,k})$, where $G_{d,k} = (\mathbb{Z}_k)^d$ and $S_{d,k}$ consists of vectors whose exactly one coordinate is non-zero. Then we can apply [Mim14, Theorem 3.4] (see also Remark 3.6) with $\nu = 1$ and we have that

$$\lambda_1(H(d,k); X, q) = \frac{d(k-1)}{2}\kappa_{X,q}(G_{d,k}, S_{d,k})^q.$$

Recall that by the result of Ricard [Ric14] (see also Example 4.2) the noncommutative Mazur map, which we also write $M_{p,2}$, is

- $p/2$-Hölder if $p \in [1,2]$;
- and Lipschitz if $p > 2$.

(Note that multiplicative constants do not depend on $\mathcal{M}$ in direct sumargument.) By spectral calculus, it is not difficult to show that $\lambda_1(H(d,k); \mathbb{R}, 2) = k$, and so

$$\kappa_{X,2}(G_{d,k}, S_{d,k}) = \left(\frac{2k}{d(k-1)}\right)^{1/2}.$$

Therefore by Proposition 5.1, we have that

- $\lambda_1(H(d,k); L_p(\mathcal{M}), p) \gtrsim_p k$ for $p \in [1,2]$;
- and $\lambda_1(H(d,k); L_p(\mathcal{M}), 2) \gtrsim_p k$ for $p > 2$.

(For the former inequalities, see that $\ell_p(\mathbb{N}, L_p(\mathcal{M}))$ is again a noncommutative L_p space.) Finally, we will prove the converse order inequalities. For $p \in [1,2]$, consider the following mapping

$$f_p \colon H(d,k) \to \ell_p(d, \ell_p(T, \mathbb{R})); \quad (a_1, \ldots a_d) \mapsto (\chi_{\{a_1\}}, \ldots, \chi_{\{a_d\}}).$$

Here T is the base set ($|T| = k$) of $H(d,k)$, and χ stands for the characteristic function. Then simple calculation shows that

$$\frac{1}{2}\frac{\sum_{v\in V_{d,k}}\sum_{e=(v,w)\in E_{d,k}}\|f_p(w) - f_p(v)\|^p}{\sum_{v\in V_{d,k}}\|f_p(v) - m(f_p)\|^p} = \frac{k^p}{(k-1)^{p-1}+1} \asymp_p k$$

(note that $k \geq 2$). Because $\ell_p(\mathbb{N}, L_p(\mathcal{M}))$ contains ℓ_p, this shows that $\lambda_1(H(d,k); L_p(\mathcal{M}), p) \asymp_p k$ for $p \in [1,2]$. For $p > 2$, because ℓ_2 is an isometric subspace of $L_p((0,1))$, we can approximately embed $H(d,k)$ into $\ell_2(\mathbb{N}, L_p(\mathcal{M}))$ by using f_2 by approximating (finitely many) elements in $L_p((0,1))$ by step functions in ℓ_p. This gives that $\lambda_1(H(d,k); L_p(\mathcal{M}), 2) \lesssim_p k$ and therefore $\lambda_1(H(d,k); L_p(\mathcal{M}), 2) \asymp_p k$. □

Corollary 8.2. *In the setting of Theorem 8.1, the following hold true.*

(i) (1) *For $p \in [1,2)$, $c_{L_p(\mathcal{M})}(H(d,k)) \asymp_p d^{1-1/p}$.*
(2) *For $p \in [2,\infty)$, $c_{L_p(\mathcal{M})}(H(d,k)) \asymp_p d^{1/2}$.*

(ii) For an infinite sequence $\{H(d_n,k_n)\}_n$ with $\lim_{n\to\infty} d_n = \infty$, the following hold:
(1) *For $p \in [1,2)$, the supremum of the exponents $\alpha \in [0,1]$ such that there exists $C > 0$ such that (t^α, Ct) can be a control pair of $\coprod_n H(d_n,k_n)$ into $L_p(\mathcal{M})$ is $1/p$.*
(2) *For $p \in [2,\infty)$, the supremum of the exponents $\alpha \in [0,1]$ such that there exists $C > 0$ such that (t^α, Ct) can be a control pair of $\coprod_n H(d_n,k_n)$ into $L_p(\mathcal{M})$ is $1/2$.*

Proof. On (i), in both cases, inequalities from below follow from Theorem 1.11 and Theorem 8.1. Inequalities from above can be deduced from the special embeddings of $H(d_n,k_n)$ indicated in the proof of theorem 8.1.

On (ii), inequalities from above follow from the estimations on distoritons in (i) and Lemma 1.13. Ones from below are again from the special embeddings above. □

ACKNOWLEDGMENTS

The author wishes to express his gratitude to the organizers, specially to Florent Baudier and Romain Tessera, for the organization and the invitation to the present fall school "Metric Embeddings : Constructions and Obstructions" in Novermber, 2014. He is also gratefull to Ana Khukhro and Mikhail Ostrovskii for helpful comments.

References

[AT14] G. Arzhantseva and R. Tessera, *Relatively expanding box spaces with no expansion*, preprint, arXiv:1402.1481 (2014).

[Aus11] T. Austin, *Amenable groups with very poor compression into Lebesgue spaces.*, Duke Math. J. **159** (2011), no. 2, 187–222.

[BFGM07] U. Bader, A. Furman, T. Gelander, and N. Monod, *Property (T) and rigidity for actions on Banach spaces.*, Acta Math. **198** (2007), no. 1, 57–105.

[BL76] J. Bergh and J. Löfström, *Interpolation spaces. An introduction.*, Grundlehren der Mathematischen Wissenschaften, vol. 223, Springer-Verlag, Berlin, 1976.

[BL00] Y. Benyamini and J. Lindenstrauss, *Geometric nonlinear functional analysis. vol. 1.*, American Mathematical Society Colloquium Publications, vol. 48, American Mathematical Society, Providence, RI, 2000.

[Chu97] F. R. K. Chung, *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics, vol. 92, American Mathematical Society, 1997.

[GN12] R. I. Grigorchuk and P. W. Nowak, *Diameters, distortion and eigenvalues*, European J. Combin. **33** (2012), no. 7, 1574–1587.

[Gro77] J. L. Gross, *Every connected regular graph of even degree is a Schreier coset graph*, J. Combi. Theory Ser. B **22** (1977), no. 3, 227–232.

[JV14] P.-N. Jolissaint and A. Valette, *L_p-distortion and p-spectral gap of finite graphs*, Bull. London. Math. Soc. **46** (2014), no. 2, 329–341.

[Laf08] V. Lafforgue, *Un renforcement de la propriété* (T), Duke Math. J. **143** (2008), 559–602.

[Mat97] J. Matoušek, *On embedding expanders into ℓ_p spaces*, Israel J. Math. **102** (1997), no. 1, 189–197.

[Mim14] M. Mimura, *Sphere equivalence, Banach expanders, and extrapolation*, Int. Math. Res. Notices, online published (2014).

[MN12] M. Mendel and A. Naor, *Nonlinear spectral calculus and super-expanders*, preprint, arXiv:1207.4705, to appear in Inst. Hautes Études Sci. Publ. Math. (2012).

[Nao14] A. Naor, *Comparison of metric spectral gaps*, Anal. Geom. Metr. Spaces **2** (2014), 1–52.

[OS94] E. Odell and Th. Schlumprecht, *The distortion problem*, Acta Math **173** (1994), 259–281.

[Oza04] N. Ozawa, *A note on non-amenability of $B(\ell_p)$ for $p = 1, 2$*, Internat. J. Math. **15** (2004), 557–565.

[Pis10] G. Pisier, *Complex interpolation between Hilbert, Banach and operator spaces*, Mem. Amer. Math. Soc. **208** (2010), no. 978, vi+78.

[PZ02] I. Pak and A. Żuk, *On Kazhdan constants and random walks on generating subsets*, Int. Math. Res. Not. **36** (2002), 1891–1905.

[Ray02] Y. Raynaud, *On ultrapowers of non commutative L_p spaces.*, J. Operator Theory **48** (2002), 41–68.

[Ric14] E. Ricard, *Hölder estimates for the noncommutative Mazur maps*, preprint, arXiv:1407.8334 (2014).

Masato Mimura, Mathematical Institute, Tohoku University
E-mail address: mimura-mas@m.tohoku.ac.jp

Følner type sets, Property A and coarse embeddings

Thibault Pillon

November 2, 2014

Abstract

Our goal is to expose amenability as a tool to produce good embeddings of metric spaces into Banach spaces. After introducing amenability, focussing on Følner's isoperimetric criterion, we show how Yu's property A generalizes the notion to uniformly discrete metric spaces. We show how to produce proper isometric actions of amenable groups and coarse embeddings of metric spaces with Property A. Finally, by keeping track of the size of Følner sets, we obtain lower bounds on the compression functions of those embeddings.

1 Amenability and proper actions on Hilbert spaces

1.1 The Hausdorff-Banach-Tarski paradox and von Neumann's definition

The Hausdorff-Banach-Tarski paradox states that it is possible to cut a sphere into finitely many pieces and reassemble them with no deformations into two spheres of the same size as the original one. It is called a paradox only because it contradicts our geometrical intuition in a very strong sense. What makes such a cutting possible lies in the use of the axiom of choice and of non-Lebesgue-measurable pieces. In the study of that theorem, the notion of amenability arose as a fundamental group theoretic property forbidding such decompositions.

Theorem 1.1 (Hausdorff, 1914 [Hau] - Banach,Tarski, 1924 [BT]). *Let* $X=\mathbb{S}^2$ *denote the two dimensional unit sphere in* $\mathbb{R}^3$ *and let* $G = \mathrm{SO}_3(\mathbb{R})$ *be its group* *of isometries. There exists a non-measurable partition of* X *into four subsets* $A_1, A_2,$ *and* B_1, B_2 *and rotations* $\alpha_1, \alpha_2, \beta_1, \beta_2 \in G$ *such that*

$$(\alpha_1 \cdot A_1) \sqcup (\alpha_2 \cdot A_2) = G, \text{ and } (\beta_1 \cdot B_1) \sqcup (\beta_2 \cdot B_2) = G.$$

Proof: Consider the subgroup $F = \mathrm{F}(\alpha, \beta)$ of G generated by the two matrices

$$\alpha = \begin{pmatrix} 3/5 & -4/5 & 0 \\ 4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & -4/5 \\ 0 & 4/5 & 3/5 \end{pmatrix}$$

and admit that this subgroup is free. Consider the following partition of F into four subsets :

$$\begin{aligned} \mathcal{A}_+ &= \{\text{reduced words starting with the letter } \alpha\} \\ \mathcal{A}_- &= \{\text{reduced words starting with the letter } \alpha^{-1}\} \\ \mathcal{B}_+ &= \{\text{reduced words starting with the letter } \beta\} \cup \{\beta^{-n},\ n \geq 0\} \\ \mathcal{B}_- &= \{\text{reduced words starting with the letter } \beta^{-1}\} \setminus \{\beta^{-n},\ n \geq 0\} \end{aligned}$$

These sets satisfy the following :

$$\mathcal{A}_+ \sqcup \alpha\mathcal{A}_- = G \text{ and } \mathcal{B}_+ \sqcup \beta\mathcal{B}_- = G.$$

Now fix a set of representatives $\{x_i\}_{i \in \mathcal{I}}$ of the F-orbits in X and define

$$\begin{aligned} A_1 &= \{g \cdot x_i,\ g \in \mathcal{A}_+,\ i \in \mathcal{I}\}, & A_2 &= \{g \cdot x_i,\ g \in \mathcal{A}_-,\ i \in \mathcal{I}\}, \\ B_1 &= \{g \cdot x_i,\ g \in \mathcal{B}_+,\ i \in \mathcal{I}\}, & B_2 &= \{g \cdot x_i,\ g \in \mathcal{B}_-,\ i \in \mathcal{I}\}. \end{aligned}$$

We obtain that $X = A_1 \sqcup (\alpha \cdot A_2) = B_1 \sqcup (\beta \cdot B_2)$. □

Such a decomposition is called a *paradoxical decomposition*. From his study of the Banach-Tarski Paradox, Von Neumann came up with the following definition :

Definition 1.2 (von Neumann, 1929 [vN]). Let G be a discrete group, a *mean* on G is a linear functional $M : \ell^\infty(G) \to \mathbb{R}$ which satisfies

1. $M(f) \geq 0$ whenever $f \geq 0$,
2. $M(\mathbf{1}) = 1$.

A mean is called *left-invariant* if additionally

3. $M(g \cdot f) = M(f)$, for every $g \in G,\ f \in \ell^\infty(G)$.

G is called *amenable* if it admits a left-invariant mean.

Remark 1.3. To get a intuitive understanding of the notion, it is important to note that evaluating a left-invariant mean on indicator functions of subsets of G will give us a left invariant *finitely-additive measure* on G.

The crucial observation of von Neumann is that the existence of paradoxical decompositions of the group is an obstruction to amenability. Tarski later proved that it is actually the only obstruction.

Theorem 1.4 (Tarski, 1938 [Ta])**.** *A discrete group G admits a paradoxical decomposition if and only if it is not amenable.*

In a modern view-point, theorem 1.1 uses non-amenability of a certain isometric action of the free group on the sphere to produce a paradoxical decomposition of that sphere. It is difficult to prove amenability or non-amenability of a group using this definition but let's see some examples.

Example 1.5.

1. Every finite group is amenable. Averaging a function amongst the elements of the group provides a left-invariant mean.
2. Free groups are non-amenable. The case of two generators follows from the proof of Theorem 1.1 and the argument for more generators is completely similar.
3. The group $\mathbb{Z}$ of all integers is an amenable group. Providing an explicit left-invariant mean is impossible since it relies on the axiom of choice. One such mean could be given by taking the limit of bounded functions along a $\mathbb{Z}$-invariant ultrafilter.

1.2 Følner's criterion

The most surprising fact about the concept of amenability is that it admits many equivalent definitions coming from very diverse areas of mathematics : measure theoretic, geometric, dynamical, analytic, spectral, etc. The most important for our exposition is the Følner geometric characterization in terms of sets with small boundaries.

Definition 1.6. Let G be a finitely generated group equipped with the word metric associated to some finite generating set, let A be a subset of G, and let $R > 0$. Define the *R-boundary* of A as

$$\partial_R A = \{g \in G \setminus A \mid d(g, A) \leq R\}.$$

Fix $\varepsilon > 0$, a finite subset A of G is called an (R, ε)-*Følner set* if it satisfies

$$\frac{\#\partial_R A}{\#A} \leq \varepsilon$$

This definition is well-suited to give an intuitive notion of Følner sets as sets with small boundaries, however it is almost always more practical to work with the following :

Definition 1.5 (revisited)**.** A finite subset $A \subset G$ is called an (R, ε)-*Følner set* if it satisfies

$$\frac{\#(g \cdot A \triangle A)}{\#A} \leq \varepsilon$$

for every $g \in G$ such that $|g| \leq R$.

The equivalence between the two definitions relies on the fact that the size of the symmetric difference between A and one of its close translates is roughly equal to the size of its boundary. Note that in order to pass from one definition to the other we may have to multiply ε or R by some fixed constant.

Theorem 1.6 (Følner, 1955 [Føl])**.** *A finitely generated group G is amenable if and only if, for every $\varepsilon > 0$ and for every $R > 0$, G contains an (R, ε)-Følner set.*

Remark 1.7. Fixing $R = 1$ in the theorem would give the exact same class of groups. This is due to the fact that R-boundaries for large R can be controlled in terms of 1-boundaries. So to obtain an (R, ε)-Følner set, one can choose a $(1, \delta)$-Følner set for a sufficiently small δ. In this setting, a sequence of $(1, \varepsilon_n)$-Følner sets (F_n) is called a *Følner sequence* if $\varepsilon_n \to 0$. It will always satisfy

$$\lim_{n \to \infty} \frac{\#g \cdot F_n \triangle F_n}{\#F_n} = 0$$

However, it is very convenient to keep the flexibility of fixing R

Proof: We only give a sketch.

Suppose that G satisfies Følner's criterion and let $F_n \subset G$ be $(n, \frac{1}{n})$-Følner sets. Define functionals M_n on $\ell^\infty(G)$ by

$$M_n(\varphi) = \frac{1}{\#F_n} \sum_{g \in F_n} \varphi(g).$$

The M_n are unit functionals on $\ell^\infty(G)$, and by compactness of the unit sphere in $\ell^\infty(G)^*$ we can assume that the sequence (M_n) converges to a weak-* limit

M. It is easy to check that M is a mean, and left-invariance is a consequence of the asymptotic invariance of the F_n's.

For the converse, we use that $\ell^1(G)$ is dense in its bidual $\ell^\infty(G)^*$. Given M a left-invariant mean, choose a sequence $\phi_n \in \ell^1(G)$ of finite support functions converging to M. Moreover, choose each ϕ_n so that there exist $N > 0$ such that ϕ_n takes value in $\{0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N}, 1\}$. By left-invariance of M, $g \cdot \phi_n - \phi_n$ must become small as n goes to infinity. Considering the sets $F_n^k = \{x \in G \mid \phi_n(x) \leq \frac{k}{N}\}$, we see that by a pigeon-hole principle, at least one of them must be close to its translate by g. □

Let us now revisit our previous examples from Følner's point of view.

Example 1.8.

1. Every finite group is amenable. Indeed, the group itself is an (R, ε)-Følner set for any R and ε.

2. Free groups are non-amenable. Indeed, the Cayley graph of a free group of rank k is a $2k$-regular tree. We can easily check that any connected sub-tree containing n points has a 1-boundary of size $n(2k-2)+2$ forbidding the existence of $(1, \varepsilon)$-Følner sets for small values of ε.

3. The group $\mathbb{Z}$ of all integers is an amenable group. Intervals of the form $[0, n]$ are (R, ε)-Følner at least when $n > \varepsilon/R$.

4. One goes easily from $\mathbb{Z}$ to $\mathbb{Z}^d$ and to any abelian finitely generated group.

Følner's criterion naturally raises the following question : when does an infinite sequence of balls form a Følner sequence? The following gives a complete answer to this question.

Corollary 1.9. *All groups with subexponential growth are amenable.*

Proof: We'll prove the converse statement, i.e. that non-amenabe groups have exponential growth.

Let G be a finitely generated group. Denote by $B(n)$ the ball of radius n and by $S(n)$ the sphere of radius n in G. We have

$$\begin{aligned}
\#B(n) &= \#B(n-1) + \#S(n) \\
&= \#B(n-1)\left(1 + \frac{\#S(n)}{\#B(n-1)}\right) \\
&= \#B(n-2)\left(1 + \frac{\#S(n-1)}{\#B(n-2)}\right)\left(1 + \frac{\#S(n)}{\#B(n-1)}\right) \\
&= \prod_{i=1}^{n}\left(1 + \frac{\#S(i)}{\#B(i-1)}\right).
\end{aligned}$$

It is immediate that $\partial B(n) = S(n+1)$, so by non-amenability of G, the general term of the product must be uniformly bounded away from 1. This implies exponential growth. □

Note that the proof also tells us that in a non-amenable group of sub-exponential growth, at least a subsequence of the balls forms a Følner sequence.

1.3 Gromov's a-T-menability

Let us recall a few facts about groups actions on Hilbert spaces.

Definition 1.10. An *affine isometric action* α of G on a Banach space E is a homomorphism of G into the group of affine isometric transformations $\mathrm{Aff}(E)$.

Such an action is called *proper* if moreover for some (equivalently for all) $\xi \in E$

$$\|\alpha(g)\xi\| \to \infty \text{ whenever } |g| \to \infty$$

Definition 1.11 (Gromov, 1988 [?]). A group G is called *a-T-menable* if it admits a proper affine isometric action on a Hilbert space.

A-T-menability was introduced by Gromov as a strong negation of Kazhdan's property (T) which requires that every affine isometric action of the group on a Hilbert space has bounded orbits. The terminology follows from the fact that a-T-menability is a weak form of amenability, although this is not clear from the definition.

Example 1.12.

1. $\mathbb{Z}^d$ is a-T-menable. Indeed, the action

$$\alpha(m_1, \ldots, m_d)(x_1, \ldots, x_d) = (x_1 + m_1, \ldots, x_d + m_d)$$

   is proper.

2. The free group on two generators $\mathbb{F}_2 = F(a, b)$ acts properly on a Hilbert space.

   Proof: Consider the action of $\mathbb{F}_2$ on its Cayley graph $\Gamma = (V, E)$ for the standard generating set. Equip Γ with the natural orientation where edges have positive orientation from g to ag or bg and negative orientation otherwise. Consider now the Hilbert space $\mathcal{H} = \ell^2(E)$ of square summable functions on the edges of Γ. The left-action of G on Γ lifts to a unitary representation of $\mathcal{H}$. Define now $b : G \to \mathcal{H}$ by

$$b(g)(e) = \begin{cases} 1 & \text{if } e \notin [e, g] \\ -1 & \text{if } e \notin [g, e] \\ 0 & \text{otherwise} \end{cases}$$

where $[x,y]$ denotes the oriented geodesic from x to y. It is easily checked that the formula

$$\alpha(g)\xi = g \cdot \xi + b(g)$$

defines a proper affine isometric action of G. □

The following theorem shows that amenable groups are a-T-menable, it is essential to us since it gives an explicit construction of a proper action given Følner sets on the group. The same approach will be applied in the non-equivariant setting and in both cases we will be able to obtain quantitative information about the actions (resp. coarse maps) obtained this way.

Theorem 1.13 (Bekka-Cherix-Valette, 1993 [BCV])**.** *Any amenable group admits a proper affine isometric action on a Hilbert space.*

Proof: Let G be an amenable group, and let F_n be $(n, 1/n^2)$-Følner sets in G. Consider the Hilbert sum $\mathcal{H} = \bigoplus_{i=1}^{\infty} \ell^2(G)$ equipped with the natural diagonal action of G. Now define $\xi_n \in \ell^2(G)$ by

$$\xi_n = \frac{1}{\sqrt{(\#F_n)}}\chi_{F_n},$$

where χ_{F_n} denotes the indicator function of F_n, and define $b(g) \in \mathcal{H}$ by $b(g) = \bigoplus_n g \cdot (\xi_n - \xi_n)$. Note that $b(g)$ belongs to $\mathcal{H}$ since

$$\begin{aligned} \|b(g)\|^2 &= \sum_{n=1}^{\infty} \|g \cdot \xi_n - \xi_n\|^2 \\ &= \sum_{n=1}^{\infty} \frac{\#(g \cdot F_n \bigtriangleup F_n)}{\#F_n} \end{aligned}$$

and by Følner's condition when n becomes large enough, the summand is dominated by $1/n^2$ which insures that the series converges. Define $\alpha : G \to \mathrm{Aff}(\mathcal{H})$ by

$$\alpha(g)v = g \cdot v + b(g).$$

This is a well-defined affine isometric action of G. To see that it is proper, notice that as $|g|$ grows larger and larger, so does the amount of indices n sucht that $g \cdot F_n$ and F_n are disjoint. Hence

$$\begin{aligned} \|b(g)\| &\geq 2 \cdot \#\{n \mid F_n \cap g \cdot F_n = \emptyset\} \\ &\to \infty \quad \text{as } |g| \to \infty. \end{aligned}$$

□

2 Property A and coarse embeddings

2.1 Property A

Definition 2.1 (Yu, 2000 [Yu]). Let X be a uniformly discrete metric space. We say that X has *Property A* if for every $\varepsilon > 0$ and $R > 0$, there exists a collection $(A_x)_{x \in X}$ of finite subsets of $X \times \mathbb{N}$ and $S > 0$ such that

(a) $\dfrac{\# A_x \triangle A_y}{\# A_x \cap A_y} \leq \varepsilon$ whenever $d(x, y) \leq R$, and

(b) $A_x \subset B(x, S) \times \mathbb{N}$.

Such subsets are called (R, ε)-Følner type sets.

Observe that condition (a) is similar to Følner's condition; sets associated to close points are close. Condition (b), however, replaces equivariance. Indeed, in group it is always the case that finite subsets are disjoint from their far translates. Here, we make it a requirement.

The use of the extra dimension $\mathbb{N}$ allows us to count points with multiplicity and is necessary for technical reasons.

Example 2.2. Amenable groups, seen as uniformly discrete spaces have property A. Indeed, fix $R, \varepsilon > 0$ and let F be a (R, δ)-Følner set for a suitable δ. Then the family of sets $A_g = gF \times \{1\}$ satisfies property A for R and ε.

The question whether Property A for groups is equivalent to amenability is natural and the following example shows that it isn't. Indeed, free groups have trees as Cayley graphs.

Example 2.3. Infinite trees have property A.

Proof: Let T be such a tree and choose x_0 a root in T. From any $x \in T$ there exists a unique minimal path from x to x_0. Fix $n > 0$ and build a set $A_x \subset T \times \mathbb{N}$ in the following way : assign weight 1 to x (meaning put the point $x \times \{0\}$ in the set A_x) then follow the path to x_0 to the next vertex. Assign weight 1 to this vertex and keep going until either $\# A_x = n$ or you reach x_0. If x_0 is reached, assign the correct weight to x_0 so that $\# A_x = n$.

Computations show that $\# A_x \triangle A_y \leq 2d(x, y)$ and $\# A_y \cap A_y \geq n - 2d(x, y)$. Hence

$$\lim_{n \to \infty} \frac{\# A_x \triangle A_y}{\# A_y \cap A_y} = 0$$

which is enough to insure the existence of (R, ε)-A sets fo any R and ε. $\square$

2.2 Asymptotic dimension

Since proving Property A is not easy in general, we give one important criterion which insures it.

Definition 2.4 (Gromov, 2000 [Gro2])**.** Let $\mathcal{U} = \{U_i\}_{i \in I}$ be a cover of the metric space X. Given $R > 0$, the R-multiplicity of $\mathcal{U}$ is the smallest integer n such that every ball of radius R in X intersects at most n elements of $\mathcal{U}$.

The *asymptotic dimension* of X, $\mathrm{AsDim}(X)$ is the smallest integer n such that for any $R > 0$ there exists a uniformly bounded cover of X with R-multiplicity $n + 1$.

Asymptotic dimension is suited to the large scale point of view. Intuitively, we want to associate a dimension to a metric space which corresponds to the topological dimension of the space seen from afar. It shares many features with more classical notions of dimension and gives intuitive results on familiar objects (see items 1. 2. and 3. below)

Example 2.5.

1. Compact metric spaces have asymptotic dimension 0.
2. Real trees have asymptotic dimension 1.
3. $\mathrm{AsDim}(\mathbb{Z}^n) = n$.
4. Hyperbolic metric spaces have finite asymptotic dimension, but there exist hyperbolic spaces with arbitrarily large asymptotic dimension.
5. $\mathbb{Z}^{(\infty)}$ and the wreath product $\mathbb{Z} \wr \mathbb{Z}$ bith have infinite asymptotic dimension.

The following result gives a practical criterion for having property A, we state if without proof.

Theorem 2.6 (Higson-Roe, 2000 [HR])**.** *Let X be a uniformly discrete metric space. If X has finite asymptotic dimension, then X has property A.* □

2.3 Coarse embeddings

Recall the following definitions :

Definition 2.7. A map $F : X \to Y$ is callled *coarse* if there exist control functions $\rho_+, \rho_- : \mathbb{R}_+ \to \mathbb{R}_+$, with $\lim_{t \to \infty} \rho_- = +\infty$, such that

$$\rho_-(d(x,y)) \leq d(F(x), F(y)) \leq \rho_+(d(x,y)), \text{ for all } x, y \in X.$$

Furthermore, the maximal map ρ_- for that condition (namely $\rho_-(t) = \inf\{d(F(x), F(y)) \mid d(x,y) \leq t\}$) is called the *compression function* of F.

The study of spaces, especially groups, which admit embeddings into Hilbert spaces (or more general Banach spaces) has been very important in connection with conjectures coming from index theory and geometry. Property A was designed by Yu as a tool to produce such embeddings.

Proposition 2.8 (Yu, 2000)**.** *Let X be a uniformly discrete metric space. If X has property A then X embeds coarsely into a Hilbert space.*

Proof: The construction is very similar to the proof of theorem 1.13. We'll define an embedding in $\bigoplus \ell^2(X \times \mathbb{N})$. First, for each $n > 0$ fix a family $\left(A_x^{(n)}\right)$ of $(n, \frac{1}{n^2})$-Følner type sets. Then define $\xi_x^{(n)} \in \ell^2(X \times \mathbb{N})$ by

$$\xi_x^{(n)} = \frac{\chi_{A_x^{(n)}}}{\sqrt{(\# A_x^{(n)})}}.$$

Now fix a base point $z \in X$ and define $F : X \to \bigoplus_n \ell^2(X \times \mathbb{N})$ by

$$F(x) = \bigoplus_{n=1}^{\infty} \left(\xi_z^{(n)} - \xi_x^{(n)}\right).$$

We need to check that this map is well-defined and is indeed a coarse embedding. Fix $x, y \in X$ and choose k minimal so that $d(x, y) \leq k + 1$, we have

$$\begin{aligned}
\|F(x) - F(y)\|^2 &= \sum_{n=1}^{\infty} \left\|\xi_z^{(n)} - \xi_x^{(n)}\right\| \\
&= \sum_{n=1}^{\infty} \frac{\# \left(A_x^{(n)} \triangle A_y^{(n)}\right)}{\# A_x^{(n)}} \\
&\leq \sum_{n=1}^{k} \frac{\# \left(A_x^{(n)} \triangle A_y^{(n)}\right)}{\# A_x^{(n)}} + \sum_{n=k+1}^{\infty} \frac{1}{n^2} \\
&\leq 2k + 8 \leq 2d(x, y) + 10.
\end{aligned}$$

In the case $y = z$ this gives us that $F(x)$ is well-defined. The general statement gives an upper control function for the map F. For the lower control function, note that by condition (b) in definition 2.1, there exists a sequence S_n such that

$$\operatorname{supp}(A_x^{(n)}) \subset B(x, n)$$

It is straightforward that in order to satisfy condition (a), the sequence (S_n) must tend to infinity. Without loss of generality suppose (S_n) is increasing and define $\phi(k) = \max\{n \mid 2S_n < k \leq d(x,y)\}$, this ensures that $A_x^{(n)}$ and $A_y^{(n)}$ are disjoint whenever $n \leq \phi(k)$ We obtain

$$\begin{aligned} \|F(x) - F(y)\| &= \sum_{n=1}^{\phi(k)} \frac{\#\left(A_x^{(n)} \triangle A_y^{(n)}\right)}{\#A_x^{(n)}} + \sum_{n=\phi(k)+1}^{\infty} \frac{\#\left(A_x^{(n)} \triangle A_y^{(n)}\right)}{\#A_x^{(n)}} \\ &\geq 2\phi(k). \end{aligned}$$

□

This proposition gives us the first obstruction to property A. A space which doesn't embed coarsely into a Hilbert space can not satisfy Property A, hence families of expander graphs don't have A. Giving more examples of spaces without this property is difficult and whether the last proposition admits a converse is even harder. See A. Khukhro's notes and talk for more about the subject.

3 Quantitative properties and compression functions

The purpose of this section is to sharpen the notions of Følner and Følner type sets to obtain lower control on the compression functions of the embeddings we constructed. All following material is due to Tessera [Te1, Te2].

Definition 3.1. Let G be an amenable group, a Følner sequence $(F_n)_{n\geq 1}$ of G is called *controlled* if there exists $C > 0$ such that

$$\operatorname{diam} F_i \leq \frac{C}{\varepsilon}$$

whenever F_n is $(1, \varepsilon)$-Følner.

So, in addition to the existence of sets with small boundaries, we require that such sets can be chosen small enough. For combinatorial reasons, the condition above is the sharpest one can ask for. In other words, groups with controlled Følner sequences are as good as it gets. The following proposition shows that these groups embed in L^p spaces with very good compression functions. We provide it without proof.

Theorem 3.2 ([Te2]). *Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing function satisfying*

$$\int_1^\infty \left(\frac{f(t)}{t}\right)^p \frac{dt}{t} < \infty$$

and let G be a amenable group with controlled Følner sets. Then there exists an affine isometric action of G on a Hilbert space whose compresion function ρ satisfies

$$\rho(t) \succcurlyeq \alpha(t).$$

□

Example 3.3. Groups with polynomial growth have controlled Følner sequences. Indeed if $\#B(n) \approx n^\alpha$ it is easily checked that $\frac{\#S(n+1)}{\#B(n)} \approx 1/n$. So the family of all balls form a controlled Følner sequence.

Proposition 3.4. *The following classes of groups have controlled Følner sequences:*

1. *Polycyclic groups.*
2. *Amenable connected Lie groups.*
3. *Some algebraic semi-direct products, in particular amenable Baumslag-Solitar groups.*
4. *Wreath products of the form $F \wr \mathbb{Z}$ with F finite.*

□

The same idea applied to Følner type sets gives the following definition:

Definition 3.5. Let X be a uniformly discrete metric space, let $J : \mathbb{R}_+ \to \mathbb{R}_+$ be some increasing function and fix $1 \leq p < \infty$. We say that X has *quantitative property A(J,p)* if for each $n > 0$ there exists a family $\left(A_x^{(n)}\right)_{x\in X}$ such that

1. $\#A_x^{(n)} \geq J(n)^p$,
2. $\#\left(A_x^{(n)} \bigtriangleup A_y^{(n)}\right) \leq d(x,y)^p$,
3. $\operatorname{supp} A_x^{(n)} \subseteq B(x,n)$.

Theorem 3.6. *Let X be a metric space with property A(J,p) as above and let f be an increasing function satisfying*

$$\int_1^\infty \left(\frac{f(t)}{J(t)}\right)^p \frac{dt}{t} < \infty.$$

Then there exists a large scale Lipschitz coarse embedding of X into an L^p space with compression function ρ satisfying

$$\rho \succcurlyeq f.$$

Proof: Fix a base point $z \in X$ and fix families $\left(A_x^{(n)}\right)$ as in the definition. Define $F_n : X \to \ell^p(X)$ by

$$F_n(x) = \left(\frac{f(2^n)}{J(2^n)}\right)\left(\chi_{A_x^{(2^n)}} - \chi_{A_z^{(2^n)}}\right)$$

and set $F : X \to \left(\bigoplus \ell^p(X)\right)_p$, $F(x) = \bigoplus F_n(x)$ We need to prove that F is well-defined and that it satisfies the requirement of the theorem. We have

$$\begin{aligned}
\|F(x) - F(y)\|_p^p &= \sum_{n=1}^{\infty} \|F_n(x) - F_n(y)\|_p^p \\
&= \sum_{n=1}^{\infty} \left(\frac{f(2^n)}{J(2^n)}\right)^p \# \left(A_x^{(2^n)} \triangle A_y^{(2^n)}\right) \\
&\leq d(x,y)^p \int_1^{\infty} \left(\frac{f(2^u)}{J(2^u)}\right)^p \mathrm{d}u \\
&= d(x,y)^p \int_1^{\infty} \left(\frac{f(t)}{J(t)}\right)^p \frac{\mathrm{d}t}{t}.
\end{aligned}$$

This both shows that F is well-defined (set $y = z$) and that it is Lipschitz. On the other hand, fix $x, y \in X$ and choose N maximal such that $d(x,y) > 2^{(N+1)}$. This condition ensures that $A_x^{(2^N)}$ and $A_y^{(2^N)}$ are disjoint. We obtain

$$\begin{aligned}
\|F(x) - F(y)\|_p^p &\geq \|F_N(x) - F_N(y)\|_p^p \\
&= \left(\frac{f(2^N)}{J(2^N)}\right)^p \# \left(A_x^{(2^N)} \triangle A_y^{(2^N)}\right) \\
&\geq \left(\frac{f(2^N)}{J(2^N)}\right)^p 2J(2^N)^p \\
&= 2f(2^N) \geq 2f\left(d(x,y)\right)
\end{aligned}$$

which shows that $\rho_F \succcurlyeq f$. □

We expose some classes of metric spaces for which this approach is fruitful. As in the equivariant case, looking at balls as potential controlled Følner type sets gives us results linking growth and compression functions.

Theorem 3.7. *1. Let X be a quasi-geodesic metric space with subexponential growth ν i.e.*

$$\#B(x,r) \leq \nu(r),\ \forall x \in X, r > 0.$$

Then X has $A(J_p,p)$ for every $1 \leq p < \infty$, where $J_p(t) \approx (t/\log \nu(t))^{1/p}$.

2. Moreover, if we assume homogeneity on the size of balls, namely that

$$\#B(x,n) < C\nu(n) \text{ for some } C > 0,$$

one can choose $J_p(t) \approx t/\log v(t)$ *for all* $1 \leq p < \infty$.

3. Moreover, is X *is a uniformly doubling metric space, i.e such that* ν *satisfies*

$$\nu(2r) \leq C'\nu(r),$$

then one can choose $J(t) \approx t$.

Theorem 3.8. *Let* X *be an homogeneous Riemannian manifold. Then* X *has property* $A(J,p)$ *for all* $p \geq 1$ *and* $J \approx t$.

References

[BT] Banach, St. and Tarski, A. (1929). Sur la décomposition des ensembles de points en parties respectivement congruents, Fund. Math., 14, 127-131.

[BCV] Bekka, M. E. B., Chérix, P.-A. and Valette, A. (1995). Proper affine isometric actions of amenable groups. Novikov Conjectures, Index Theorems, and Rigidity, 226, 1.

[Føl] Flner, E. (1955), On groups with full Banach mean value, Math. Scand.3, 243254.

[Gro1] Gromov, M. (1988). Rigid transformations groups. Géométrie différentielle (Paris, 1986), 33, 65-139.

[Gro2] Gromov, M. (1996). Geometric group theory, Vol. 2: Asymptotic invariants of infinite groups. Bull. Amer. Math. Soc, 33, 0273-0979.

[Hau] Hausdorff, F. (1914). Bemerkung ber den Inhalt von Punktmengen. Math. Ann. 75(3), 428-433

[HR] Higson, N., Roe J. (2000). Amenable group actions and the Novikov conjecture. J. Reine Angew. Math. (519), 143-153.

[Ta] Tarski, A. (1938), Algebraische Fassung de Massproblems, Fund. Math. 31, 47-66

[Te1] Tessera, R. (2008). Quantitative property A, Poincar inequalities, L p-compression and L p-distortion for metric measure spaces. Geometriae Dedicata, 136(1), 203-220.

[Te2] Tessera, R. (2011). Asymptotic isoperimetry on groups and uniform embeddings into Banach spaces. Commentarii Mathematici Helvetici, 86(3), 499-535.

[vN] von Neumann, J. (1929). Zur allgemeinen Theorie des Masses. Fund. Math. 13, 73116.

[Yu] Yu, G. (2000). The coarse BaumConnes conjecture for spaces which admit a uniform embedding into Hilbert space. Inventiones Mathematicae, 139(1), 201-240.

ON RAMSEY TECHNIQUES IN QUANTITATIVE METRIC GEOMETRY: *THE MINIMUM DISTORTION NEEDED TO EMBED A BINARY TREE INTO ℓ_p*

DANIEL GALICER

ABSTRACT. It is commonly said that a result is typical of the Ramsey theory, if in any finite coloring of some mathematical object one can extract a sub-object (usually having some kind of desired structure), which is monochromatic. In this essay we discuss in detail a clever Ramsey-type argument due to Jiří Matoušek utilized in the context of embedding theory. Namely, to study the smallest constant $C = C(n)$ for which a complete binary tree of height n can be C-embedded into a given uniformly convex Banach space. As a consequence, the quantitative lower bound of $const \cdot (\log n)^{\min(1/2,1/p)}$ in the distortion needed to embed this space into ℓ_p (for $1 < p < \infty$) is explained.

1. A GLIMPSE TO RAMSEY-TYPE RESULTS AND BOURGAIN'S WORK ON BINARY TREES

Let us begin with a seemingly banal but enlightening question: How many people should be on a party to ensure that three of them are either mutual acquaintances (each one knows the other two) or mutual strangers (each one does not know either of the other two)? This query is usually known as the problem of friends and strangers. For our purposes, it is convenient to phrase this question in a graph-theoretic language. Denote by n the number of people at the party and suppose that each person is represented by a vertex of a complete graph (a simple undirected graph in which every pair of distinct vertices is connected by a unique edge) K_n of order n. Given two partygoers (or vertices), we paint in red the edge that links them if they know each other and in blue otherwise. Therefore, our problem translates into the following: How big n must be to assert the existence of a complete subgraph of order 3 in K_n painted entirely in red or blue?

Ramsey's classical theorem [Ram30] points in the same direction as this question. Colloquially speaking, it states that in any coloring of the edges (using a palette with a finite number of colors) of a sufficiently large complete graph, one will find monochromatic (i.e., of the same color) complete subgraphs. This foundational tool in combinatorics initiated a new perspective that is now framed as part of the Ramsey theory. But what exactly do people mean

Key words and phrases. Graph trees, bi-Lipschitz structure, Ramsey techniques, Uniformly convex Banach spaces.

when they refer to a statement as of Ramsey-type? Perhaps the most popular result of this type (although quite naive) is the well-known pigeonhole principle: if $f : \{1,\dots n\} \to \{1,\dots m\}$ and $n > m$ then f can not be injective (if you have fewer pigeon holes than pigeons and you put every pigeon in a pigeon hole, then there must result at least one pigeon hole with more than one pigeon). The typical scenario of the Ramsey theory starts with some mathematical object which is divided into several pieces. The question that arises in this context is how big should be the original object in order to affirm that at least one of the pieces has a given interesting property. Going back to our friends and strangers' example, we wanted to know how big had to be our study set (n = the number of partygoers) to ensure the existence of a certain structure (three "friends" or three "complete strangers"). By the way... the answer is $n \geq 6$ and it is an interesting challenge to prove this, but this is another matter.

Summarizing, a statement has essentially a Ramsey-type flavor if it ensures the existence of some kind of rigid substructure in a given set having enough members. Being a bit extreme, Ramsey-type results give certain regularity amid disorder. These techniques have proven to be extremely useful in various contexts alien to it (allowing to solve, for example, long-standing problems in analysis; see [AT06] for a proper treatment on several important applications). Of course, Ramsey theory may be labeled undoubtedly as a part of combinatorics or discrete mathematics, and in general these branches seem to be quite distant, at least at first glance, from embedding theory or metric geometry. The aim of this note is to show how to apply this kind of discrete techniques to study the smallest distortion needed in a particular embedding problem. Before going into details, let us start with a couple of definitions in order to clarify all the notions we deal with.

Given two metric spaces (M, d_M), (N, d_N), and a mapping $f : M \to N$, we denote the Lipschitz constant of f by $\|f\|_{\text{Lip}} := \sup\{\frac{d_N(f(x),f(y))}{d_M(x,y)} : x \neq y\}$. If f is injective then the (bi-Lipschitz) distortion of f is defined as $\mathsf{dist}(f) = \|f\|_{\text{Lip}} \cdot \|f^{-1}\|_{\text{Lip}}$. Informally, the distortion is a measure of the amount by which a function warps distances. Note that a function with distortion 1 does not necessarily preserve mutual distances but it may re-scale them in the same ratio. We write $M \overset{C}{\hookrightarrow} N$ if there exist an embedding $f : M \to N$ with $\mathsf{dist}(f) \leq C$ (such an embedding is called a C-embedding or a C-isomorphism). The smallest distortion with which M embeds into N is denoted $c_N(M)$, namely,

$$c_N(M) = \inf\{C : \ M \overset{C}{\hookrightarrow} N\}.$$

We say that $f : M \hookrightarrow N$ is non-contracting if $d_M(x, y) \leq d_N(f(x), f(y))$ for every $x, y \in M$ (i.e., $\|f^{-1}\|_{\text{Lip}} \leq 1$). In this working we focus on the case where the target space N is a Banach space $(X, \|\cdot\|)$, so we can compute $c_X(M)$ (by re-scaling if necessary) as $\inf\{\|f\|_{\mathrm{Lip}} : f : M \hookrightarrow X \text{ non-contracting}\}$. If $X = \ell_p$ for some $p \geq 1$ we use the shorter notation $c_p(M) = c_{\ell_p}(M)$. The parameter $c_2(M)$ is usually known as the Euclidean distortion of X.

Lipschitz (or uniform and coarse) embeddings of metric spaces into Banach spaces with "good geometrical properties" have found many significant applications, specially in computer science and topology. The advantages of low distortion embeddings are based on the fact that for spaces with "good properties" one can apply several geometric tools which are generally not available for typical metric spaces. The most significant accomplishments throughout these lines were obtained in the design of algorithms (the information obtained from concrete geometric representations of finite spaces is used to obtain efficient approximation algorithms and data structures). In this context, the spaces with "good geometrical features" are mostly separable Hilbert spaces (or certain classical Banach spaces such as L_p spaces).

The bi-Lipschitz structure of arbitrary trees and its applications to different context have been studied extensively during the last years. We refer to [Dre84, Mat90, Bar98, JLPS02, LS03, Dra03, FRT03, BS05, NPS+06] and the references therein for a detailed treatment. Recall that a (graph-theoretical) tree is an undirected graph $T = (V, E)$ in which any two vertices are connected by exactly one path. In other words, any connected graph without simple cycles is a tree. The present essay is devoted to the study of the Euclidean (and L_p) distortion of complete binary trees.

Just to be in tune, we denote by B_n the complete rooted binary tree of height (or depth) n. This is a graph defined as follows: B_0 is a single vertex (the root), and B_{n+1} arises by taking one vertex (the root) and connecting it to the roots of two disjoint copies of B_n. We also consider k-ary trees of height h (each non-leaf vertex has k successors), which we denote by $T_{k,h}$. These spaces are metric space endowed with the path-metric: the distance between two vertices is the number of edges in the path connecting them (i.e., we consider the graph-theoretic distance on the vertex set, with edges of unit length).

A famous result in embedding theory due to Bourgain [Bou86] states the following.

Theorem 1.1. *Let $1 < p < \infty$, for any embedding $f : B_n \hookrightarrow \ell_p$ we have $\mathsf{dist}(f) \geq c \log(n)^{\min(1/2, 1/p)}$, where c is a constant depending only on p.*

In other words, he showed that $c_p(B_n) = \Omega_p\left(\log(n)^{\min(1/2,1/p)}\right)$. Among Bourgain's contributions we find a noteworthy characterization (in terms of their metric structure) of a linear property of Banach spaces. Namely, he showed that a Banach space X is superreflexive (see definition below) if and only if $\lim_{m\to\infty} c_X(B_n) = \infty$. He also established the following interesting dichotomy: For a Banach space X either $c_X(B_n) = 1$ for all n, or there exists $\alpha > 0$ such that $c_X(B_n) = \Omega\left((\log n)^{\alpha}\right)$. Bourgain used this result to solve a question posed by Gromov, showing that the hyperbolic plane does not admit a bi-Lipschitz Euclidean embedding. The arguments involved in his work are based on the use of some technical probabilistic tools (diadic Walsh-Paley martingales). We highlight that Bourgain derived Theorem 1.1 as a particular case of a much more general result involving structural properties of Banach spaces. We recall some classic definitions from Banach space theory in order to state all this.

The modulus of (uniform) convexity $\delta_X(\varepsilon)$ of a Banach space X endowed with norm $\|\cdot\|$ is defined as

$$\delta_X(\varepsilon) := \inf\left\{1 - \left\|\frac{x+y}{2}\right\| \,\Big|\, \|x\| = \|y\| = 1 \text{ and } \|x-y\| \geqslant \varepsilon\right\},$$

for $\varepsilon \in (0,2]$. The space X is said to be uniformly convex of type $q \geqslant 2$ if $\delta_X(\varepsilon) \geqslant c\varepsilon^q$ for some $c > 0$. Put simply, the modulus of convexity measures how deep inside (in the unit ball of X) must lie the midpoint of a line segment with extremes in the sphere of X in terms of the length of the segment. Intuitively, if a space has a "big" modulus of convexity then the center of a line segment included in the unit ball must lie very deep inside the ball (i.e., has small norm) unless the segment is short. If the function $\delta_X(\cdot)$ is never zero, we say that X is uniformly convex (or uniformly rotund). Spaces with this property are common examples of reflexive Banach spaces (this is a consequence of the classical Milman-Pettis theorem [Mil38, Pet39]). Since the converse does not hold, this justifies the name given to those spaces that are isomorphic to uniformly convex spaces; that is, superreflexive Banach space.

It is well-known that, for $1 < p < \infty$, the ℓ_p space (or any L_p-space) is uniformly convex. The asymptotic behavior of its moduli (as computed by Hanner [Han56]) is given by

$$\delta_p(\varepsilon) = \begin{cases} \frac{(p-1)\varepsilon^2}{8} + o(\varepsilon^2) & \text{if } 1 < p \leq 2, \\ \frac{\varepsilon^p}{p2^p} + o(\varepsilon^p) & \text{if } 2 \leq p < \infty. \end{cases} \tag{1}$$

In particular, $\delta_p(\varepsilon) \geq c\varepsilon^{\max(2,p)}$ where $c = c(p)$.

Now that we have the definition of uniform convexity in mind, we are able to state Bourgain's embedding theorem on binary trees.

Theorem 1.2. *Let X be a uniformly convex Banach space whose modulus of uniform convexity satisfies $\delta_X(\varepsilon) \geqslant c\varepsilon^q$ for some $q \geq 2$ and $c > 0$ (i.e., X uniformly convex of type $q \geqslant 2$). Then* *the minimum distortion necessary for embedding B_n into X is at least c_1 $(\log n)^{1/q}$ for some* $c_1 = c_1(c,q) > 0$.

Observe that Theorem 1.1 becomes a direct consequence of this theorem since, by Equation (1), any L^p-space ($1 < p < \infty$) is uniformly convex of type $q = \max(2,p)$. It should be noted that Bourgain's bound in Theorem 1.1 is optimal, as proven by Bourgain himself in his seminal work for the Euclidean case ($p = 2$) and by Matoušek [Mat99] for every $1 < p < \infty$. Thus, $c_p(B_n) = \Theta\left(\log(n)^{\min(1/2,1/p)}\right)$.

Several proofs of Theorem 1.2 have been published over the years (e.g., [Bou86, Mat99, LS03, LNP09, MN13, Klo14]). This note aims to present an elementary proof (due to Matoušek in [Mat99]), where a shrewd use of a Ramsey-type result is displayed.

2. Matoušek's proof or Theorem 1.2

Matoušek's argument has a geometric ingredient and a combinatorial one. The former is the simplest and relates uniform convexity to embeddings of some special trees. Consider the four-vertices tree with one root v_0 which has one son v_1 and two grandchildren v_2, v_2'. We denote by S this tree (with edges of unit length). We say that a subset $F = \{x_0, x_1, x_2, x_2'\}$ of a metric space (M, d_M) is an δ*-fork* if there exist a function $f : S \to F$ mapping v_i to x_i (for $i = 0,1,2$) and v_2' to x_2', such that the restricted functions $f|_{\{v_0,v_1,v_2\}} : \{v_0, v_1, v_2\} \to \{x_0, x_1, x_2\}$ and $f|_{\{v_0,v_1,v_2'\}} : \{v_0, v_1, v_2'\} \to \{x_0, x_1, x_2'\}$ are $(1+\delta)$-isomorphisms. Qualitatively, for small δ, the mutual distances between elements of the sets $\{x_0, x_1, x_2\}$ and $\{x_0, x_1, x_2'\}$ are similar to those of $\{0,1,2\} \subset \mathbb{R}$. It should be noted that in this definition, no information about the distance between the vertices v_2 and v_2' is inherited by F. We call the points x_2 and x_2' the *tips* of F. The name "fork" (which is obviously given by mnemonic purposes) comes by understanding how this object should look like in the Euclidean space $\mathbb{R}^3$ for a small δ. The following lemma states that if a fork F in a uniformly convex Banach space has a rigid structure (i.e. δ is small) then its tips are very close to each other.

Lemma 2.1. *(Fork Lemma) Let X be a uniformly convex Banach space whose modulus of uniform convexity satisfies $\delta_X(\varepsilon) \geqslant c\varepsilon^q$ for some $q \geq 2$ and $c > 0$, and let $F = \{x_0, x_1, x_2, x_2'\} \subset X$ be an δ-fork. Then $\|x_2 - x_2'\| = O(\delta^{1/q})\|x_0 - x_1\|$.*

Proof. (of Lemma 2.1) By translating and re-scaling if necessary, we may presume that $x_0 = 0$ and $\|x_1\| = 1$.

Set

$$z := x_1 + \frac{x_2 - x}{\|x_2 - x_1\|}.$$

Obvious computations show $\|z - x_1\| = 1$, $\|x_2 - x_1\| \leq 1 + 2\delta$, $\|z - x_2\| \leq 2\delta$ and $\|z\| \geq 2 - 2\delta$. Put $u = z - 2x_1$. Observe that $x = x_1$ and $y := x_1 + u$ lie on the unit sphere of X, and for the midpoint

$$\frac{x+y}{2} = x + \frac{u}{2} = \frac{z}{2}$$

we have

$$\left\| \frac{x+y}{2} \right\| \geq 1 - \delta.$$

Using the uniform convexity condition we obtain $\|y - x\| = \|u\| = O(\delta^{1/q})$, thus

$$\|x_2 - 2x_1\| \leq \|x_2 - z\| + \|\underbrace{z - 2x_1}_{u}\| \leq 2\delta + O(\delta^{1/q}) = O(\delta^{1/q}),$$

for δ small. Note that the constant of proportionality in the last $O(\cdot)$ notation depends only on c and q. Analogously (by symmetry), we also get $\|x_2' - 2x_1\| = O(\delta^{1/q})$, hence $\|x_2 - x_2'\| = O(\delta^{1/q})$, concluding the proof. □

For our purposes it will be useful to compute the smallest distortion needed to embed complete k-ary trees into uniformly convex spaces instead of dealing with complete binary trees. Any complete k-ary tree can be 2-embedded into a complete binary tree of height large enough. This is stated in the following lemma. Recall that the level of a vertex of $T_{k,h}$ is just its distance from the root.

Lemma 2.2. *Let $T_{k,h}$ be a complete k-ary tree of height h. Then $T_{k,h}$ can be 2-embedded into the complete binary tree B_n for height $n = 2h\lceil \log_2 k \rceil$.*

Proof. Note that it is sufficient to demonstrate this for powers of 2 (i.e., $k = 2^s$). We obviously map the root of $T_{2^s,h}$ into the root of B_{2hs}. We now follow an inductive procedure. If a vertex v of $T_{2^s,h}$ has already being mapped to a vertex u at some level l of B_{2hs}, we map the 2^s successors of v to 2^s vertices above u at level $l + 2s$ whose mutual distances are all between $2s$ and $4s$. Indeed, without loss of generality we can assume that $l = 0$, now note that B_{2s} is constructed by gluing to each leaf of B_s another disjoint copy of B_s. For each of these copies we select a leaf and map the successors of v to them. □

Given a rooted tree T, we denote by $SP(T)$ the set all pairs of vertices $\{x, y\}$ of T such that x lies in between the way from y to the root. The following Ramsey-type result, whose proof is simple and short, can be regarded as the cornerstone towards the proof of Theorem 1.2.

Lemma 2.3. *Suppose that each of the pairs of the set $SP(T_{k,h})$ is painted with a color from a palette of r colors. If $k \geq r^{(h+1)^2}$, then there exist a subtree $T' \subset T_{k,h}$ which is a copy of the complete binary tree B_h, such that the color of any pair $\{x,y\} \in SP(T')$ depends exclusively on the levels of x and y.*

Proof. We start proving the following simple claim: Suppose that all the leaves of $T_{k,h}$ (i.e., vertices at level h) are colored by r' colors and $k > r'$ then there exist a copy of B_h in $T_{k,h}$ such that all its leaves have the same color. The case $h = 0$ is trivial. For $h \geq 1$, consider all the k subtrees isomorphic to $T_{k,h-1}$ connected to the root of $T_{k,h}$. By inductive hypothesis we can pick a copy of B_{h-1} with monochromatic leaves. Since $k > r'$ by the pigeonhole principle, two of this copies have the same color of leaves. If we connect this copies to the root we get the copy of B_h with the desired property.

Going back to our problem... Label each leaf $z \in T_{k,h}$ by a vector having the colors of the pairs $\{x,y\} \in SP(T_{k,h})$ lying on the path form z to the root (we write the coordinates of the vectors using a predetermined order common for all leaves). We want show the existence of subtree $T' \subset T_{k,h}$, which is a copy of B_h, such that the color of the pair $\{x,y\} \in SP(T')$ depends only on the levels of x and y. This can be rephrased into finding a copy of B_h in $T_{k,h}$ whose all leaves are labeled with the same vector. Note that each of this vector have $\binom{h+1}{2} < (h+1)^2$ coordinates; hence, the leaves of $T_{k,h}$ are colored with $r' < r^{(h+1)^2}$ possible colors. The result now follows from our preliminary claim. □

The following lemma states that if a copy of the metric space $P_h = \{0,1,\ldots,h\} \subset \mathbb{R}$ is embedded with a constant-bounded distortion into a given metric space, and h is large enough, then we can find a 3-term arithmetic progression such that the restriction of our embedding to this set has distortion near 1.

Lemma 2.4. *(Path embedding Lemma) For any given constants $\alpha > 0$ and $\beta \in (0,1)$ the exists a constant $C = C(\alpha,\beta)$ with the following property: for every non-contracting mapping f defined in the metric space $P_h = \{0,1,\ldots,h\} \subset \mathbb{R}$ into some metric space (M,d_M) with $h \geq 2^{CK^\alpha}$, for $K = \|f\|_{\text{Lip}}$, there exists an arithmetic progression $Z = \{x, x+a, x+2a\} \subset P_h$ such that the restriction of f onto Z is a $(1+\varepsilon)$-isomorphism with*

$$\varepsilon = \beta\left(\frac{d_M(f(x), f(x+a))}{a}\right)^{-\alpha}.$$

The proof of the lemma is a bit cumbersome. Maybe should put it aside on a first read.

Proof. (of Lemma 2.4) We define, for $a \in \{1,\dots,h\}$ the number

$$K(a) := \max\left\{\frac{d_M(f(x), f(y))}{|x-y|} : x, y \in P_h, |x-y| = a\right\}.$$

By the triangle inequality $K(a) \geq K(2a)$ for every a.

We also define an decreasing sequence of numbers $x_0 > x_1 > x_2 > \dots$ by setting $x_0 = K$ and $x_{j+1} = \frac{x_j}{\left(1+\frac{\beta}{4x_j^\alpha}\right)}$. We denote by t the first index with $x_t \leq 1$. It can be seen that $t = O(K^\alpha)$, and therefore we can assume that $2^t \leq h$ (by picking C large enough). Observe that in the sequence $K(2^0) \geq K(2^1) \geq K(2^2) \geq \cdots \geq K(2^t)$, there must be two consecutive values, say $K(2^i)$ and $K(2^{i+1})$, belonging to the same interval $[x_{j+1}, x_j)$. Thus,

$$1 \leq \frac{K(2^i)}{K(2^{i+1})} \leq 1+\eta,$$

where $\eta = \frac{\beta}{4K(2^i)^\alpha}$. We consider the number $a := 2^i$ and we fix the points $x, x+2a \in P_h$ such that $K(2a) = K(2^i)$ is attained. This means that, $d_M(f(x), f(x+2a)) = 2aK(2a)$. We therefore have

$$d_M(f(x), f(x+a)) \leq aK(a) \leq a(1+\eta)K(2a),$$

and also

$$d_M(f(x+a), f(x+2a)) \leq a(1+\eta)K(2a).$$

In addition, we have

$$\begin{aligned} d_M(f(x), f(x+a)) &\geq d_M(f(x), f(x+2a)) - d_M(f(x+a), f(x+2a)) \\ &\geq 2aK(2a) - a(1+\eta)K(2a) \\ &= a(1-\eta)K(2a). \end{aligned}$$

From the equations above, the result easily follows. □

We are now able to display Matoušek's proof of Theorem 1.2. First we sketch the main steps of his argument. We pick k, h adequate natural numbers such that $T_{k,h} \overset{2}{\hookrightarrow} B_n$ (according to Lemma 2.2) and consider a non-contracting mapping $f : T_{k,h} \to X$ such that $\|f\|_{\text{Lip}}$ is is smaller than our expected bound (i.e., $\|f\|_{\text{Lip}} = c_1(\log n)^{1/q}$, for c_1 small enough), we will get to an absurdity. Using cunningly Lemma 2.3 we are able to find a complete binary tree inside $T_{k,h}$ for which f embeds "identically" every path between a root to a leaf (this is the key point and is based heavily on mixing combinatorics with distortion). This fact, together with Lemma 2.4, allow us to find a 0-fork in $T_{k,h}$ (for some $a \in \mathbb{N}$) mapped by f to an δ-fork in X for δ small. But, according to Lemma 2.1, this can not happen: the tips of the 0-fork in $T_{k,h}$ are far apart.

A rigorous and detailed proof is the following.

Proof. (of Theorem 1.2) First we declare the parameters involved in the proof, their values will be fixed later. Let $\beta > 0$ be small enough (depending on q and c), suppose n is large and let k, h be natural numbers (depending on n). Fix $f : T_{k,h} \to X$ a non-contracting mapping with $\|f\|_{\mathrm{Lip}} = K = c_1(\log n)^{1/q}$ for c_1 small, we will get a contraction.

Let $r = \left\lceil \frac{2K^{q+1}}{\beta} \right\rceil$, and suppose $k \geq r^{(h+1)^2}$. We label the pairs in $SP(T_{k,h})$ according to the distortion of their distance by f; that is to say, each pair $\{x, y\} \in SP(T_{k,h})$ is colored with the number

$$\left\lfloor \frac{K^q}{\beta} \frac{\|f(x) - f(y)\|}{d_{T_{k,h}}(x, y)} \right\rfloor \in \{0, 1, \ldots, r-1\},$$

where $d_{T_{k,h}}$ stands for the path-metric in $T_{k,h}$. By our Ramsey-type result, Lemma 2.3, we can find a subtree T', which is a copy of B_h, inside $T_{k,h}$ such that the color of each pair $\{x, y\} \in SP(T')$ depends exclusively on the levels of x and y. This is the core of Matoušek's argument: we manage to find a binary tree on which the mutual distances induced by f only depend on the position of the vertices.

Fix P a path from a root to a leaf in T' (note that this path, is isometric to $P_h = \{0, \ldots, h\} \subset \mathbb{R}$). If h is big enough, say $h = 2^{CK^q}$ where $C = C(q, \beta)$ is as in Lemma 2.4, we can pick three vertices y_0, y_1, y_2 of P whose levels form an arithmetic progression with common difference a (i.e., this vertices are at levels $l, l+a, l+2a$, respectively), such that the restriction of f to this triple becomes a $(1+\delta)$ isomorphism for

$$\delta = \beta \left(\frac{\|f(y_0) - f(y_1)\|}{a} \right)^{-q}.$$

Let y_2' be a vertex at T' at the same level as y_2 (i.e, $l+2a$) and at distant $2a$ from y_2 (note that this also implies that y_2' is at distant a and $2a$ from y_1 and y_0, respectively). By the level dependence of the colors we have that the pairs $\{y_i, y_2\}$ and $\{y_i, y_2'\}$ in $SP(T')$ are equally labeled ($i = 0, 1$). Precisely, for $i = 0, 1$ we have

$$\left\lfloor \frac{K^q}{\beta} \frac{\|f(y_i) - f(y_2)\|}{d_{T_{k,h}}(y_i, y_2)} \right\rfloor = \left\lfloor \frac{K^q}{\beta} \frac{\|f(y_i) - f(y_2')\|}{d_{T_{k,h}}(y_i, y_2')} \right\rfloor. \tag{2}$$

This implies that the restriction of f to the triple $\{y_0, y_1, y_2'\}$ is a $(1+2\delta)$-isomorphism (a priori we can not ensure to be a $(1+\delta)$ isomorphism since the equality in Equation (2) is given only for the integer parts). Therefore, the set $\{f(y_0), f(y_1), f(y_2), f(y_2')\}$ is a 3δ-fork in X. By Lemma 2.1, we obtain

$$2a \leq \|f(y_2) - f(y_2')\| = O(\delta^{1/q} a) = O(\beta^{1/q}) \|f(y_0) - f(y_1)\| = O(\beta^{1/q} a).$$

Recall that the constant of proportionality in the last $O(\cdot)$ notation depends only on c and q (and not on β). Thus, by choosing β small enough we have a contradiction.

We have made several assumptions... It is time to see how to choose properly the parameters involved. We had $h = 2^{CK^q}$, hence if c_1 in the expression $K = c_1(\log n)^{1/q}$ is small enough, we can ensure that $h < n^{1/4}$. On the other hand, we had $k = r^{(h+1)^2}$, thus $\log_2 k = (h+1)^2 \log_2 r = O\left(\sqrt{n}(\log\log n)\right)$, therefore

$$h \log_2 k = O(n^{5/6}) < n, \tag{3}$$

for n large enough. Equation (3) and Lemma 2.2 ensures that the tree $T_{k,h}$ with which we have dealt can be embedded with distortion at most 2 into the complete binary B_n. This completes the proof. □

REFERENCES

[AT06] S. A. Argyros and S. Todorcevic. *Ramsey Methods in Analysis*. Springer, 2006.

[Bar98] Yair Bartal. On approximating arbitrary metrices by tree metrics. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 161–168. ACM, 1998.

[Bou86] J. Bourgain. The metrical interpretation of superreflexivity in banach spaces. *Israel Journal of Mathematics*, 56(2):222–230, 1986.

[BS05] Sergei Buyalo and Viktor Schroeder. Embedding of hyperbolic spaces in the product of trees. *Geometriae Dedicata*, 113(1):75–93, 2005.

[Dra03] A Dranishnikov. On hypersphericity of manifolds with finite asymptotic dimension. *Transactions of the American Mathematical Society*, 355(1):155–167, 2003.

[Dre84] A. W. Dress. Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: a note on combinatorial properties of metric spaces. *Advances in Mathematics*, 53(3):321–402, 1984.

[FRT03] Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 448–455. ACM, 2003.

[Han56] O. Hanner. On the uniform convexity of l_p and l_p. *Arkiv för Matematik*, 3(3):239–244, 1956.

[JLPS02] William B Johnson, Joram Lindenstrauss, David Preiss, and Gideon Schechtman. Lipschitz quotients from metric trees and from banach spaces containing l< sub> 1</sub>. *Journal of Functional Analysis*, 194(2):332–346, 2002.

[Klo14] Benoît R Kloeckner. Yet another short proof of bourgainŠs distortion estimate for embedding of trees into uniformly convex banach spaces. *Israel Journal of Mathematics*, pages 1–4, 2014.

[LNP09] James R Lee, Assaf Naor, and Yuval Peres. Trees and markov convexity. *Geometric and Functional Analysis*, 18(5):1609–1659, 2009.

[LS03] Nathan Linial and Michael Saks. The euclidean distortion of complete binary trees. *Discrete and Computational Geometry*, 29(1):19–22, 2003.

[Mat90] Jiří Matoušek. Extension of lipschitz mappings on metric trees. *Commentationes Mathematicae Universitatis Carolinae*, 31(1):99–104, 1990.

[Mat99] J. Matoušek. On embedding trees into uniformly convex banach spaces. *Israel Journal of Mathematics*, 114(1):221–237, 1999.

[Mil38] D. Milman. On some criteria for the regularity of spaces of the type (b). *Doklady Akad. Nauk SSSR*, 20:243–246, 1938.

[MN13] Manor Mendel and Assaf Naor. Markov convexity and local rigidity of distorted metrics. *Journal of the European Mathematical Society*, 15(1):287–337, 2013.

[NPS+06] Assaf Naor, Yuval Peres, Oded Schramm, Scott Sheffield, et al. Markov chains in smooth banach spaces and gromov-hyperbolic metric spaces. *Duke Mathematical Journal*, 134(1):165–197, 2006.

[Pet39] B. J. Pettis. A proof that every uniformly convex space is reflexive. *Duke Mathematical Journal*, 5(2):249–253, 1939.

[Ram30] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, 2(1):264–286, 1930.

DEPARTAMENTO DE MATEMÁTICA - PAB I, FACULTAD DE CS. EXACTAS Y NATURALES, UNIVERSIDAD DE BUENOS AIRES, (1428) BUENOS AIRES, ARGENTINA AND CONICET

E-mail address: dgalicer@dm.uba.ar

NON-EMBEDDABILITY OF THE URYSOHN SPACE INTO SUPERREFLEXIVE BANACH SPACES

ADRIANE KAÏCHOUH

Abstract. We present Pestov's proof that the Urysohn space does not embed uniformly into a superreflexive Banach space ([P]). Its interest lies mainly in the fact that the argument is essentially combinatorial. Pestov uses the extension property for the class of finite metric spaces ([S2]) to build affine representations of the isometry group of the Urysohn space.

Contents

1. Uniform embeddings

We recall the notion of uniform embedding of metric spaces.

Let X and Y be two metric spaces. A **uniform embedding** of X into Y is an embedding of X into Y as uniform spaces. Equivalently, a map $f : X \to Y$ is a uniform embedding if there exist two non-decreasing functions ρ_1 and ρ_2 from $\mathbb{R}_+$ to $\mathbb{R}_+$, with $0 < \rho_1 \leqslant \rho_2$ and $\lim_{r\to 0} \rho_2(r) = 0$, such that for all x, x' in X, one has

$$\rho_1(d_X(x, x')) \leqslant d_Y(f(x), f(x')) \leqslant \rho_2(d_X(x, x')).$$

In particular, a uniform embedding is uniformly continuous.

Uniform embeddings transpose the local structure of metric spaces: what matter are small neighborhoods of points. We are interested in the existence of uniform embeddings into nice Banach spaces, where niceness begins at reflexivity.

2. The Urysohn space

The Urysohn space $\mathbb{U}$ is a **universal** Polish space: it is a complete separable metric space that contains an isometric copy of every (complete) separable metric space. Moreover, the Urysohn space is remarkable for its strong homogeneity properties: up to isometry, it is the unique Polish space that is both universal and ultrahomogeneous.

Definition 2.1. A metric space X is **ultrahomogeneous** if every isometry between finite subsets of X extends to a global isometry of X.

The space $\mathbb{U}$ was built by Urysohn in the early twenties ([U1]), but was long forgotten after that. Indeed, another universal Polish space, $\mathcal{C}([0,1],\mathbb{R})$ (Banach-Mazur, see [B] and [S1]), put the Urysohn space in the shade for sixty years. It regained interest in the eighties when Katětov ([K2]) provided a new construction of the Urysohn space. From this construction, Uspenskij ([U2]) proved that not only is $\mathbb{U}$ universal but also its isometry group[1] is a universal Polish group (every Polish group embeds in $\mathrm{Iso}(\mathbb{U})$ as a topological subgroup).

We will see that in fact, the Urysohn space enjoys a much stronger homogeneity property than ultrahomogeneity. In the next section, we will present this strengthening of ultrahomogeneity.

First, let us present Katětov's construction of the Urysohn space and explain how it yields the universality of its isometry group.

2.1. **Katětov spaces.** Let X be a metric space.

Definition 2.2. A **Katětov map** on X is a map $f : X \to \mathbb{R}^+$ such that for all x and x' in X, one has

$$|f(x) - f(x')| \leqslant d(x,x') \leqslant f(x) + f(x').$$

A Katětov map corresponds to a metric one-point extension of X: if f is a Katětov map on X, then we can define a metric on $X \cup \{f\}$ that extends the metric on X by putting, for all x in X,

$$d(f,x) = f(x).$$

This will indeed be a metric because Katětov maps are exactly those which satisfy the triangle inequality.

Example 2.3. If x is a point in X, then the map $\delta_x : X \to \mathbb{R}^+$ defined by $\delta_x(x') = d(x,x')$ is a Katětov map on X. It correspond to a trivial extension of X: we are adding the point x to X.

We denote by $E(X)$ the space of all Katětov maps on X. We equip the space $E(X)$ with the supremum metric, which geometrically represents the smallest possible distance between the two extension points.

The maps δ_x of example 2.3 define an isometric embedding of the space X into $E(X)$. We therefore identify X with its image in $E(X)$ via this embedding. This observation will allow us to build towers of extensions in the next section. The essential property of those towers is the following.

Proposition 2.4. Every isometry of X extends uniquely to an isometry of $E(X)$.

In particular, the uniqueness implies that the extension defines a group homomorphism from $\mathrm{Iso}(X)$ to $\mathrm{Iso}(E(X))$.

[1] Isometry groups are endowed with the topology of pointwise convergence. Basic open sets are the sets of all isometries that extend a given partial isometry between finite subsets. When X is a complete separable metric space, its isometry group $\mathrm{Iso}(X)$ is a Polish group.

Proof. Let φ be an isometry of X. If ψ extends φ, we must have $d(\psi(f), \delta_x) = d(f, \delta_{\varphi^{-1}(x)}) = f(\varphi^{-1}(x))$ for all x in X and f in $E(X)$, hence the uniqueness.

Thus, we extend φ to the space $E(X)$ by putting $\psi(f) = f \circ \psi^{-1}$ for all f in $E(X)$. It is easy to check that the map ψ is an isometry of $E(X)$ that extends φ. □

In general, the space $E(X)$ is unfortunately not separable. Since we are interested only in Polish spaces, we circumvent this problem by considering only Katětov maps with finite support.

Definition 2.5. Let S be a subset of X and let f be a Katětov map on X. We say that S is a **support** for f if for all x in X, we have

$$f(x) = \inf_{y \in S} f(y) + d(x, y).$$

In other words, S is a support for f if the map f is the biggest 1-Lipschitz map on X that coincides with f on S.

We denote by $E(X, \omega)$ the space of all Katětov maps that admit a finite support[2]. If the metric space X is separable, then $E(X, \omega)$ is separable, it still embeds X isometrically, and isometries of X still extend uniquely to isometries of $E(X, \omega)$. Moreover, the extension homomorphism from $\mathrm{Iso}(X)$ to $\mathrm{Iso}(E(X, \omega))$ is continuous (see [M2, proposition 2.5]).

2.2. **Tower construction of the Urysohn space.** The construction of the Urysohn space we present highlights its universality: we start with an arbitrary Polish space and we build a copy of the Urysohn space around it. Besides, the construction keeps track of the isometries of the original Polish space, which points to the universality of its isometry group as well.

Let X be our starting Polish space. We build an increasing sequence (X_n) of metric spaces recursively, by setting

- $X_0 = X$;
- $X_{n+1} = E(X_n, \omega)$.

The discussion above guarantees that isometries extend continuously at each step: every isometry of X_n extends to an isometry of X_{n+1} and the extension homomorphism from $\mathrm{Iso}(X_n)$ to $\mathrm{Iso}(X_{n+1})$. Thus, if we write $X_\infty = \bigcup_{n \in \mathbb{N}} X_n$, we obtain a continuous extension homomorphism from $\mathrm{Iso}(X)$ to $\mathrm{Iso}(X_\infty)$.

Now, consider the completion $\widehat{X_\infty}$ of X_∞. Since all the X_n are separable, the space $\widehat{X_\infty}$ is Polish. Moreover, isometries of X_∞ extend to isometries of $\widehat{X_\infty}$ by uniform continuity, so we get a continuous extension homomorphism from $\mathrm{Iso}(X)$ to $\mathrm{Iso}(\widehat{X_\infty})$.

It remains to explain why the space $\widehat{X_\infty}$ is the promised ultrahomogeneous and unique Urysohn space. The key defining property of $\widehat{X_\infty}$ is that every one-point metric extension of a finite subset of $\widehat{X_\infty}$ is realized in $\widehat{X_\infty}$ over this finite set.

Definition 2.6. A metric space X is said to have the **Urysohn property** if for every finite subset A of X and every Katětov map $f \in E(A)$, there exists x in X such that for all a in A, we have $d(x, a) = f(a)$.

Theorem 2.7. (Urysohn) Let X be a complete separable metric space. If X has the Urysohn property, then X is ultrahomogeneous.

Proof. We carry a back-and-forth argument. Let $i : A \to B$ an isometry between two finite subsets of X. Enumerate a dense subset $\{x_n : n \geqslant 1\}$ of X. Recursively, we build finite subsets A_n and B_n of X and isometries $i_n : A_n \to B_n$ such that

[2]The letter ω is the set-theoretic name for $\mathbb{N}$.

- $A_0 = A$ and $B_0 = B$;
- $i_0 = i$;
- $A_n \subseteq A_{n+1}$ and $B_n \subseteq B_{n+1}$;
- $x_n \in A_n \cap B_n$;
- i_{n+1} extends i_n.

To this aim, assume A_n and B_n have been built. Consider the metric extension of A_n by x_{n+1}: the corresponding Katětov map is $\delta_{x_{n+1}}$. We push it forward to a Katětov map on B_n via the isometry i_n. Now, since the space X satisfies the Urysohn property, we can find an element y_{n+1} that realizes it; we add it to B_n and extend i_n by setting $i'_{n+1}(x_{n+1}) = y_{n+1}$. This constitutes the *forth* step.

For the *back* step, we apply the same argument to the inverse of the isometry i'_{n+1} to find a preimage to x_{n+1}.

In the end, the union of all the isometries i_n defines an isometry of a dense subset of X, so it extends to an isometry of the whole space X (because X is complete). This is the desired extension of i. □

Another back-and-forth argument shows that any two complete separable metric spaces with the Urysohn property are isomorphic (see [G, theorem 1.2.5]). Thus, we may for instance define the Urysohn space $\mathbb{U}$ to be the space obtained from $X = \{0\}$ by applying the tower construction above. This uniqueness result guarantees that $\mathbb{U}$ indeed embeds every Polish space isometrically. Moreover, the construction also yields that its isometry group $\mathrm{Iso}(\mathbb{U})$ embeds all isometry groups of Polish spaces. A beautiful result of Gao and Kechris ([GK]) states that these actually encompass all Polish groups, so we conclude that $\mathrm{Iso}(\mathbb{U})$ is a universal Polish group.

In particular, $\mathrm{Iso}(\mathbb{U})$ contains the group $\mathrm{Homeo}_+[0,1]$ of orientation-preserving homeomorphisms of the unit interval. In the proof of theorem 7.1, we will use this fact, together with the following result of Megrelishvili ([M1]), to show that the Urysohn space does not admit any uniform embedding into a superreflexive Banach space.

Theorem 2.8. (Megrelishvili) The only continuous representation of $\mathrm{Homeo}_+[0,1]$ by linear isometries on a reflexive Banach space is the trivial representation.

3. The extension property

In 1992, Hrushovski ([H2]) proved that for every finite graph, there exists a bigger finite graph such that every partial graph isomorphism of the smaller graph extends to a global graph automorphism of the bigger graph. It turns out that this phenomenon occurs in several other structures, and in particular for metric spaces.

Definition 3.1. A metric space has the **extension property** if for every finite subset A of X, there exists a finite subset B of X that contains A such that every partial isometry of A extends to a global isometry of B.

The extension property is indeed a strengthening of ultrahomogeneity.

Proposition 3.2. Let X be a complete separable metric space. If X has the extension property, then X is ultrahomogeneous.

Proof. Let $i : A \to B$ be an isometry between two finite subsets A and B of X. We wish to extend i to a global isometry of X. First, the extension property gives a finite subset Y_0 of X containing A and B such that the partial isometry i extends to a global isometry j_0 of Y_0.

Enumerate a dense subset $\{x_n : n \geqslant 1\}$ of X. Recursively, we build an increasing chain of finite subsets Y_n of X, with $Y_{n+1} \supseteq Y_n \cup \{x_n\}$, and an increasing chain of global isometries j_n of Y_n by applying the extension property.

Let now Y be the union of all the Y_n's. The map j defined by $j(x) = j_n(x)$ if $x \in Y_n$ is a global isometry of Y. Since Y contains all the points x_n, it is dense in X, so j extends to an isometry of the whole space X (because X is complete). □

Independently, Vershik ([V]) announced and Solecki ([S2]) proved that the Urysohn space satisfies the extension property. Consequently, the extension property is sometimes also called the *Hrushovski-Solecki-Vershik property*. Note that this is really a result about the class of *all* metric spaces. It means that for every finite metric space, there exists a bigger finite metric space such that every partial isometry of the smaller metric space extends to a global isometry of the bigger metric space.

In fact, the Urysohn space satisfies an even stronger form of extension property ([S3]): we can choose the extension of those partial isometries to be compatible with the group structure. Thus, the extension will provide a group homomorphism from the isometry group of the smaller metric space to the isometry group of the bigger one. This **coherent extension property** has a very powerful consequence on the isometry group, which is the heart of the argument for theorems 6.1 and 5.5.

Proposition 3.3. Let X be a complete separable metric space. If X satisfies the coherent extension property, then its isometry group $\mathrm{Iso}(X)$ contains a dense locally finite subgroup.

A group is said to be **locally finite** if every finitely generated subgroup is finite.

Proof. We carry the same construction as in the proof of proposition 3.2: we recursively build finite subsets Y_n of X such that

- $Y_n \subseteq Y_{n+1}$;
- every partial isometry of Y_n extends to a global isometry of Y_{n+1};
- (coherence) moreover, the extension defines a group embedding from $\mathrm{Iso}(Y_n)$ to $\mathrm{Iso}(Y_{n+1})$;
- the union $Y = \bigcup_{n \in \mathbb{N}} Y_n$ of all the Y_n's is dense in X.

Since the extension is coherent, the union $G = \bigcup_{n \in \mathbb{N}} \mathrm{Iso}(Y_n)$ is an increasing union of subgroups of $\mathrm{Iso}(Y)$. Thus, as the increasing union of finite groups, it is a locally finite group. We show that the group G is dense in $\mathrm{Iso}(Y)$. By density of Y in X, the group $\mathrm{Iso}(Y)$ is dense in $\mathrm{Iso}(X)$, so this will complete the proof.

Consider a basic open set in $\mathrm{Iso}(Y)$. It is given by a partial isometry $i : A \to B$ between finite subsets of Y. Since A and B are finite, there exists an integer n such that both A and B are contained in Y_n. But then the partial isometry i of Y_n extends to a global isometry of Y_{n+1}, which is in G. Thus, the basic open set contains an element of G, and G is indeed dense in $\mathrm{Iso}(Y)$. □

Remark 3.4. In [P], Pestov states the above result for metric spaces which satisfy only the extension property, without any coherence assumption. It is not clear, then, how to build a dense locally finite subgroup recursively, as the groups $\mathrm{Iso}(Y_n)$ need not even be subgroups of $\mathrm{Iso}(Y)$, nor be included in one another.

4. Ultrapowers of Banach spaces

4.1. **Ultrafilters.** Dually to ideals giving a notion of smallness, ultrafilters give a way to declare some sets as *large*. More precisely, a **filter** on a set I is a collection $\mathcal{F}$ of subsets of I such that

- (non-triviality) the whole set I is in $\mathcal{F}$ but the empty set is not in $\mathcal{F}$;
- if A is in $\mathcal{F}$, then any subset B of I containing A also is in $\mathcal{F}$;
- the intersection of two elements of $\mathcal{F}$ is again in $\mathcal{F}$.

An **ultrafilter** is a maximal filter (with respect to inclusion). Equivalently, a filter $\mathcal{U}$ on I is an ultrafilter if and only if for each subset A of I, either A is in $\mathcal{U}$ or $I \setminus A$ is in $\mathcal{U}$.

The point of ultrafilters, aside from brewing ultracoffee, is to make arbitrary sequences converge.

Definition 4.1. Let X be a topological space. Let I be a set (of indices) and let $\mathcal{F}$ be a filter on I. Let $(x_i)_{i\in I}$ be a family of elements of X and let x be a point in X. We say that x is the **limit** of $(x_i)_{i\in I}$ **along** $\mathcal{F}$, and we write $x = \lim_{i\to\mathcal{F}} x_i$, if for every neighborhood V of x in X, the set $\{i \in I : x_i \in V\}$ is in $\mathcal{F}$.

The usual notion of convergence for sequences indexed by the integers thus corresponds to the convergence along the filter of cofinite subsets of $\mathbb{N}$: this filter contains all the intervals $[n;\infty[$.

Proposition 4.2. Let $(x_i)_{i\in I}$ be a family of elements of reals and let $\mathcal{U}$ be an ultrafilter on I. If $(x_i)_{i\in I}$ is bounded, then the family $(x_i)_{i\in I}$ has a limit along $\mathcal{U}$.

Proof. We use the classical Bolzano-Weierstrass cutting-in-half argument. Assume that the family takes its values in the bounded interval $[a, b]$. Cut the interval in two and look at which elements of the sequence fall in which half: consider the two sets

$$L = \left\{ i \in I : x_i \in \left[a, \frac{a+b}{2} \right] \right\} \text{ and } R = \left\{ i \in I : x_i \in \left[\frac{a+b}{2}, b \right] \right\}.$$

Since $\mathcal{U}$ is an ultrafilter, exactly one of the sets L and R belongs to $\mathcal{U}$, say L.

Then we do that again in L: we consider the sets

$$L' = \left\{ i \in I : x_i \in \left[a, \frac{3a+b}{4} \right] \right\} \text{ and } R' = \left\{ i \in I : x_i \in \left[\frac{3a+b}{4}, \frac{a+b}{2} \right] \right\}.$$

This time, either L' is in $\mathcal{U}$, or its complement, which is $R' \cup R$ is. But we know that L is in the ultrafilter $\mathcal{U}$ too, so the intersection $L \cap (R' \cup R) = R'$ belongs to $\mathcal{U}$; and so on.

Thus, inductively, we find a decreasing sequence of intervals $[a_n, b_n]$ of length $\frac{b-a}{n}$ such that for all n, the set $\{i \in I : x_i \in [a_n, b_n]\}$ is in the ultrafilter $\mathcal{U}$. It follows that the intersection point of all those intervals $[a_n, b_n]$ is the limit of the family $(x_i)_{i\in I}$ along the ultrafilter $\mathcal{U}$. □

The same argument readily adapts to families in any compact space (see e.g. [E2, theorem 3.1.24]).

4.2. **Ultraproducts of metric spaces.** Let $(X_i)_{i\in I}$ be a family of metric spaces. We choose a distinguished point x_i in each X_i. We consider the following subset of the product of the X_i's:

$$\ell^\infty(X_i, x_i, I) = \{y \in \prod_{i\in I} X_i : \sup_{i\in I} d_{X_i}(x_i, y_i) < \infty\}.$$

Let $\mathcal{U}$ be an ultrafilter on I. The boundedness assumption above allows us to equip $\ell^\infty(X_i, x_i, I)$ with the following pseudometric:

$$d(y, z) = \lim_{i\to\mathcal{U}} d_{X_i}(y_i, z_i).$$

The **metric space ultraproduct** along $\mathcal{U}$ of the family $(X_i)_{i\in I}$ centered at $(x_i)_{i\in I}$ is the metric quotient of the pseudometric space $(\ell^\infty(X_i, x_i, I), d)$. We denote it $\left(\prod_{i\in I}(X_i, x_i)\right)_{\mathcal{U}}$.

Remark 4.3. Any ultraproduct of complete metric spaces is easily seen to be complete.

In a normed space, the origin is a canonical choice for a distinguished point. The ultraproduct of a family of normed spaces, centered at the family of origins, comes with a natural structure of normed space. If all the normed spaces are Banach spaces, then by the above remark, their ultraproduct also is a Banach space. This normed space then induces a structure of affine normed space on the ultraproduct of normed spaces centered in an arbitrary family of points. Moreover, the choice of distinguished points does not matter too much.

Proposition 4.4. Let $(E_i)_{i\in I}$ be a family of normed spaces. Let $(x_i)_{i\in I}$ and $(x'_i)_{i\in I}$ be two families of distinguished points. Let $\mathcal{U}$ be an ultrafilter on I. Then the ultraproducts of $\left(\prod_{i\in I}(E_i, x_i)_{i\in I}\right)_{\mathcal{U}}$ and $\left(\prod_{i\in I}(E_i, x'_i)_{i\in I}\right)_{\mathcal{U}}$ are affinely isomorphic and isometric.

Proof. Consider the linear translation $(y_i)_{i\in I} \mapsto (y_i - x_i + x'_i)_{i\in I}$ in the product $\prod_{i\in I} E_i$. It sends $\ell^\infty(X_i, x_i, I)$ to $\ell^\infty(X_i, x'_i, I)$ and preserves the pseudometric. Hence, it defines an isometry between the two ultraproducts.

Moreover, since the isometry comes from a translation, the two ultraproducts are affinely isomorphic. □

When all the normed spaces E_i's are equal, say to a Banach space E, an ultraproduct of the family $(E_i)_{i\in I}$ centered at the family of origins is a Banach space, called a **Banach space ultrapower** of E.

5. Superreflexive Banach spaces

A Banach space E is said to be **superreflexive** if every Banach space ultrapower of E is reflexive. Enflo exhibited a characterization of superreflexivity in terms of convexity properties ([E1, corollary 3]): a Banach space is superreflexive if and only if it admits an equivalent norm that is uniformly convex.

Remark 5.1. In Enflo's theorem, superreflexivity is defined a bit differently; see [HM, theorem 2.3] and [S4, proposition 1.1] for the equivalence of the two definitions.

Definition 5.2. A Banach space $(E, \|\cdot\|)$ is **uniformly convex** if for every $\epsilon > 0$, there exists $\delta > 0$ such that for every x, y in E with $\|x\| = 1$, $\|y\| = 1$, one has

$$\|x - y\| \geqslant \epsilon \Rightarrow \left\| \frac{x+y}{2} \right\| \leqslant 1 - \delta.$$

In other words, a Banach space is uniformly convex if and only if its unit ball is strictly convex, this in a uniform way.

Examples 5.3. The following Banach spaces are superreflexive.

- Hilbert spaces.
- L^p spaces, for $1 < p < \infty$. This is a consequence of the Clarkson inequalities ([C, theorem 2]).

Superreflexivity is preserved under taking ℓ^2-type sums (the key argument is the Minkowski inequality).

Proposition 5.4. (Day, [D, theorem 2]) Let E be a superreflexive Banach space and X an arbitrary set. Then the Banach space $\ell^2(X, E)$ is superreflexive too.

Though uniform convexity is more workable a notion, it is intrinsically metric and it is not stable under Banach space isomorphisms, whereas superreflexivity is. Hence, since both uniform and coarse structures are invariant under isomorphisms, we state the embeddings results with superreflexivity rather than with uniform convexity.

The result we will present the proof of in the next two sections is the following.

Theorem 5.5. (Pestov) The Urysohn space does not admit any uniform embedding into a superreflexive Banach space.

Remark 5.6. Just around the same time Pestov's paper was published, a stronger result was proven by Kalton in [K1]: that the space c_0 does not admit any uniform embedding into a *reflexive* Banach space. Since c_0 is a Polish space, it embeds isometrically into the Urysohn space, so it follows that $\mathbb{U}$ does not admit any uniform embedding into a reflexive Banach space either. Still, Pestov's proof is based on very different techniques and is worth presenting.

Superreflexivity is a strengthening of reflexivity that invites ultraproducts constructions. The next section contains the main argument of Pestov's proof, an ultraproduct construction designed to *smoothen* actions on Banach spaces.

6. Averaging distances

Theorem 6.1. Let G be a locally finite group acting by isometries on a metric space X. Suppose that X admits a mapping φ into a normed space E such that for some functions $\rho_1, \rho_2 : \mathbb{R}_+ \to \mathbb{R}_+$:

$$\rho_1(d_X(x,x')) \leqslant \|\varphi(x) - \varphi(x')\| \leqslant \rho_2(d_X(x,x')).$$

Then there is a map ψ of X into a Banach space ultrapower of some $\ell^2(\mathcal{U}, E)$, satisfying the same inequalities

$$\rho_1(d_X(x,x')) \leqslant \|\psi(x) - \psi(x')\| \leqslant \rho_2(d_X(x,x')), \tag{1}$$

and such that the action of G on $\psi(X)$ extends to an action of G by affine isometries on the affine span of $\psi(X)$.

Proof. Let Ξ be the set of all finite subgroups of G. For every finite subgroup F in Ξ, we define a map $\psi_F : X \to \ell^2(F, E)$ by

$$\psi_F(x)(f) = \frac{1}{\sqrt{\operatorname{Card} F}} \varphi(f^{-1} \cdot x),$$

for every x in X and f in F.

Since G acts on X by isometries, the maps ψ_F satisfy the inequalities (1):

$$\rho_1(d_X(x,x')) \leqslant \|\psi_F(x) - \psi_F(x')\| \leqslant \rho_2(d_X(x,x')).$$

Indeed, let x and x' be two elements of X. Then we have:

$$\begin{aligned}
\|\psi_F(x) - \psi_F(x')\|_2 &= \left(\sum_{f\in F} \|\psi_F(x)(f) - \psi_F(x')(f)\|_E^2 \right)^{1/2} \\
&= \left(\frac{1}{\operatorname{Card} F} \sum_{f\in F} \|\varphi(f^{-1}\cdot x) - \varphi(f^{-1}\cdot x')\|_E^2 \right)^{1/2} \\
&\leqslant \left(\frac{1}{\operatorname{Card} F} \sum_{f\in F} \rho_2^2(d_X(f^{-1}\cdot x, f^{-1}\cdot x')) \right)^{1/2} \\
&= \left(\frac{1}{\operatorname{Card} F} \sum_{f\in F} \rho_2^2(d_X(x,x')) \right)^{1/2} \\
&= \rho_2(d_X(x,x'))
\end{aligned}$$

and similarly

$$\begin{aligned}\|\psi_F(x) - \psi_F(x')\|_2 &= \left(\frac{1}{\operatorname{Card} F} \sum_{f \in F} \|\varphi(f^{-1} \cdot x) - \varphi(f^{-1} \cdot x')\|_E^2 \right)^{1/2} \\ &\geqslant \left(\frac{1}{\operatorname{Card} F} \sum_{f \in F} \rho_1^2(d_X(f^{-1} \cdot x, f^{-1} \cdot x')) \right)^{1/2} \\ &= \left(\frac{1}{\operatorname{Card} F} \sum_{f \in F} \rho_1^2(d_X(x, x')) \right)^{1/2} \\ &= \rho_1(d_X(x, x')).\end{aligned}$$

We would like to find a map that is compatible with the action of G. The group F acts on $\ell^2(F, E)$ by isometries, via the left regular representation: for $r \in \ell^2(F, E)$ and f, g in F, we define

$$^g r(f) = r(g^{-1} f).$$

Then the map ψ_F becomes F-equivariant:

$$\begin{aligned}^g(\psi_F(x))(f) &= \psi_F(x)(g^{-1} f) \\ &= \frac{1}{\sqrt{\operatorname{Card} F}} \varphi(f^{-1} g \cdot x) \\ &= \psi_F(g \cdot x)(f).\end{aligned}$$

Now we are average out all the maps ψ_F's. Choose an ultrafilter $\mathcal{U}$ on Ξ with the property that for each F in Ξ, the set $\{H \in \Xi : F \subseteq H\}$ is in $\mathcal{U}$. The local finiteness of the group G guarantees that such an ultrafilter exists.

Choose a point x^* in X. This yields distinguished points $\psi_F(x^*)$ in the $\ell^2(F, E)$'s. More precisely, let

$$V = \left(\prod_{F \in \Xi} (\ell^2(F, E), \psi_F(x^*)) \right)_{\mathcal{U}}$$

be the ultraproduct of the spaces $\ell^2(F, E)$ along $\mathcal{U}$ centered at the family $(\psi_F(x^*))_{F \in \Xi}$.

We now prove that for every x in X, the family $(\psi_F(x))_{F \in \Xi}$ is at finite distance from the distinguished family $(\psi_F(x^*))_{F \in \Xi}$, hence its class defines an element of V. Let x be an element of X.

$$\begin{aligned}\sup_{F \in \Xi} \|\psi_F(x) - \psi_F(x^*)\| &\leqslant \sup_{F \in \Xi} \rho_2(d_X(x, x^*)) \\ &= \rho_2(d_X(x, x^*)).\end{aligned}$$

This implies we can define a map $\psi : X \to V$ by

$$\psi(x) = [(\psi_F)_{F \in \Xi}]_{\mathcal{U}} .$$

Moreover, the action of G on the space V is well-defined: let g be an element of G. Since G is locally finite, the subgroup of G generated by g is finite, hence in Ξ. We chose the ultrafilter $\mathcal{U}$ in such a way that the set of all F in Ξ that contain $\langle g \rangle$ is in $\mathcal{U}$. From this, it follows that g acts on $\ell^2(F, E)$ for $\mathcal{U}$-every F in Ξ.

Since the action of F on each $\ell^2(F, E)$ is an action by isometries, so is the action of G on V. For this action, the map ψ is G-equivariant as desired.

It remains to identify the ultraproduct V with a Banach space ultrapower of $\ell^2(\mathcal{U}, E)$. First, note that $\ell^2(\mathcal{U}, E)$ contains every $\ell^2(F, E)$ as a normed space (this embedding is not canonical;

this is just because F is finite and $\mathcal{U}$ is bigger). Thus, V is contained in a suitable ultraproduct of $\ell^2(\mathcal{U}, E)$, which is isometrically and affinely isomorphic to the corresponding Banach space ultrapower of $\ell^2(\mathcal{U}, E)$ by proposition 4.4.

□

7. Obstruction to a uniform embedding

Theorem 7.1. The Urysohn space $\mathbb{U}$ cannot be uniformly embedded into a superreflexive Banach space.

Proof. Suppose it can and let $\varphi : \mathbb{U} \to E$ be a uniform embedding of $\mathbb{U}$ into a superreflexive Banach space E. Let also ρ_1 and ρ_2 be two decreasing functions from $\mathbb{R}_+$ to $\mathbb{R}_+$, with $0 < \rho_1 \leqslant \rho_2$ and $\lim_{r\to 0} \rho_2(r) = 0$, witnessing that φ is a uniform embedding: such that for all x, x' in $\mathbb{U}$, one has

$$\rho_1(d_{\mathbb{U}}(x, x')) \leqslant \|\varphi(x), \varphi(x')\|_E \leqslant \rho_2(d_{\mathbb{U}}(x, x')).$$

Let G be a dense locally finite subgroup of $\mathrm{Iso}(\mathbb{U})$ (such a subgroup exists by proposition 3.3). By proposition 6.1, there exists a mapping ψ of $\mathbb{U}$ into a Banach space ultrapower V of $\ell^2(\mathcal{U}, E)$ such that for all x, x' in $\mathbb{U}$, one has

$$\rho_1(d_{\mathbb{U}}(x, x')) \leqslant \|\psi(x), \psi(x')\|_V \leqslant \rho_2(d_{\mathbb{U}}(x, x')), \tag{2}$$

and such that the action of G extends to an action by affine isometries on the affine span S of $\psi(\mathbb{U})$ in V, making ψ G-equivariant.

Note that V is reflexive as an ultrapower of the superreflexive space $\ell^2(\mathcal{U}, E)$, as to proposition 5.4.

The inequalities (2) guarantee that ψ is a uniform isomorphism on its image. In particular, ψ is a homeomorphism. So the topology on G of pointwise convergence on $\mathbb{U}$ coincides with the topology of pointwise convergence on $\psi(\mathbb{U})$, and consequently, on S as G acts by affine isometries.

Moreover, since ψ is a uniformly continuous, so is the representation of G on S. Thus, by density of G in $\mathrm{Iso}(\mathbb{U})$, the action of G extends to a uniformly continuous action of $\mathrm{Iso}(\mathbb{U})$ on S for which the map ψ remains equivariant. It follows that the representation of $\mathrm{Iso}(\mathbb{U})$ on S is faithful: if g and h are isometries such that for all x in $\mathbb{U}$, one has $g \cdot \psi(x) = h \cdot \psi(x)$, then by equivariance, one has $\psi(g \cdot x) = \psi(h \cdot x)$ for all x in $\mathbb{U}$. But since ψ is an isomorphism, this implies that for all x in $\mathbb{U}$, one has $g \cdot x = h \cdot x$, hence $g = h$.

Write this affine representation of $\mathrm{Iso}(\mathbb{U})$ on S is a continuous homomorphism from $\mathrm{Iso}(\mathbb{U})$ to the group $\mathrm{Iso}(S) = \mathrm{LIso}(S) \ltimes S_+$, where S_+ is the additive group of S (group of translations) and $\mathrm{LIso}(S)$ the group of linear isometries of S. Let also π denote the standard (continuous) projection from $\mathrm{LIso}(S) \ltimes S_+$ onto $\mathrm{LIso}(S)$.

Now recall that the group $\mathrm{Iso}(\mathbb{U})$ is a universal Polish group (Uspenskij [U2], see section 2). In particular, it contains $\mathrm{Homeo}_+[0, 1]$ as a topological subgroup. Therefore, we have a faithful continuous affine representation of the group $\mathrm{Homeo}_+[0, 1]$ in the reflexive Banach space V.

But Megrelishvili proved in [M1] that the only continuous representation of $\mathrm{Homeo}_+[0, 1]$ by linear isometries on a reflexive Banach space is the trivial representation (see theorem 2.8). Therefore, the linear part of the restriction of π to $\mathrm{Homeo}_+[0, 1]$ is trivial. $\mathrm{Homeo}_+[0, 1]$ then has to act by translations, but by faithfulness of the representation, this implies that $\mathrm{Homeo}_+[0, 1]$ is abelian, a contradiction. □

8. Concluding remarks

Let us mention which (non-)embeddability properties of the Urysohn space remain when we relax or sharpen our notion of embedding.

8.1. **Coarse embeddability.** Whereas the uniform structure gives the local behavior of metric spaces, the coarse structure, or *large-scale structure*, of a metric space describes its geometry *at infinity*.

A map $f : X \to Y$ is a **coarse embedding** of X into Y if there exist two non-decreasing unbounded functions ρ_1 and ρ_2 from $\mathbb{R}_+$ to $\mathbb{R}_+$, with $0 < \rho_1 \leqslant \rho_2$, such that for all x, x' in X, one has

$$\rho_1(d_X(x, x')) \leqslant d_Y(f(x), f(x')) \leqslant \rho_2(d_X(x, x')).$$

In particular, for a fixed x' in X, the distance $d_X(x, x')$ tends to infinity if and only if $d_X(f(x), f(x'))$ does. Note that a coarse embedding is not necessarily continuous.

Pestov also applies the techniques of theorem 6.1 to coarse embeddings to prove that the Urysohn space does not admit any coarse embedding into a superreflexive Banach space either. The proof is way more technical though[3]. Moreover, it is based on a strengthening of theorem 6.1 ([P, corollary 4.4]), the proof of which I did not understand. It states that if the locally finite group G acts *almost* transitively on the space X, then the image $\psi(X)$ we build is a *metric transform* of X, meaning that the distance $\|\psi(x) - \psi(x')\|$ depends only on $d(x, x')$.

In [K1], Kalton proved a stronger result: the Urysohn space does not even admit any coarse embedding into a reflexive Banach space. It follows from the same result for the space c_0 (see also remark 5.6).

8.2. **Isometric embeddability.** We could also simply consider isometric embeddings of the Urysohn space, which are a very special case of uniform embeddings. However, this proves to be too restrictive: there is only one way to embed the Urysohn space isometrically into a Banach space. Whenever $\mathbb{U}$ embeds isometrically into a Banach space, then the span of its image is the *Holmes space* ([H1, theorem 6]).

In conclusion, it is quite hard to embed $\mathbb{U}$ nicely into Banach spaces!

References

[B] S. Banach, *Théorie des opérations linéaires*, Chelsea Publishing Co., New York, 1955.

[C] J. A. Clarkson, *Uniformly convex spaces*, Trans. Amer. Math. Soc. **40** (1936), no. 3, 396–414.

[D] M. M. Day, *Some more uniformly convex spaces*, Bull. Amer. Math. Soc. **47** (1941), 504–507.

[E1] P. Enflo, *Banach spaces which can be given an equivalent uniformly convex norm*, Proceedings of the International Symposium on Partial Differential Equations and the Geometry of Normed Linear Spaces (Jerusalem, 1972), 1972, pp. 281–288 (1973).

[E2] R. Engelking, *General topology*, Second, Sigma Series in Pure Mathematics, vol. 6, Heldermann Verlag, Berlin, 1989. Translated from the Polish by the author.

[G] S. Gao, *Invariant descriptive set theory*, Pure and Applied Mathematics (Boca Raton), vol. 293, CRC Press, Boca Raton, FL, 2009.

[GK] S. Gao and A. S. Kechris, *On the classification of Polish metric spaces up to isometry*, Memoirs of Amer. Math. Soc. **766** (2003).

[H1] M. R. Holmes, *The Urysohn space embeds in Banach spaces in just one way*, Topology Appl. **155** (2008), no. 14, 1479–1482.

[H2] E. Hrushovski, *Extending partial isomorphisms of graphs*, Combinatorica **12** (1992), no. 4, 411–416.

[HM] C. W. Henson and L. C. Moore Jr., *Subspaces of the nonstandard hull of a normed space*, Trans. Amer. Math. Soc. **197** (1974), 131–143.

[K1] N. J. Kalton, *Coarse and uniform embeddings into reflexive spaces*, Q. J. Math. **58** (2007), no. 3, 393–414.

[K2] M. Katětov, *On universal metric spaces*, General topology and its relations to modern analysis and algebra, VI (Prague, 1986), 1988, pp. 323–330.

[M1] M. G. Megrelishvili, *Every semitopological semigroup compactification of the group $H_+[0,1]$ is trivial*, Semigroup Forum **63** (2001), no. 3, 357–370.

[3] "Now, one can verify, by considering 17 separate cases, that..."!

[M2] J. Melleray, *Some geometric and dynamical properties of the Urysohn space*, Topology Appl. **155** (2008), no. 14, 1531–1560.

[P] V. G. Pestov, *A theorem of Hrushovski-Solecki-Vershik applied to uniform and coarse embeddings of the Urysohn metric space*, Topology Appl. **155** (2008), no. 14, 1561–1575.

[S1] W. Sierpiński, *Sur un espace métrique séparable universel*, Fund. Math. **33** (1945), 115–122.

[S2] S. Solecki, *Extending partial isometries*, Israel J. Math. **150** (2005), 315–331.

[S3] ______, *Notes on a strengthening of the herwig-lascar extension theorem*, 2009.

[S4] J. Stern, *Ultrapowers and local properties of Banach spaces*, Trans. Amer. Math. Soc. **240** (1978), 231–252.

[U1] P. S. Urysohn, *Sur un espace métrique universel*, Bull. Sci. Math. **51** (1927), 43–64 and 74–96.

[U2] V. V. Uspenskij, *On the group of isometries of the Urysohn universal metric space*, Comment. Math. Univ. Carolin. **31** (1990), no. 1, 181–182.

[V] A. M. Vershik, *Globalization of the partial isometries of metric spaces and local approximation of the group of isometries of Urysohn space*, Topology Appl. **155** (2008), no. 14, 1618–1626.