ISRAEL JOURNAL OF MATHEMATICS 79 (1992), 103–111

FINITE METRIC SPACES NEEDING HIGH DIMENSION FOR LIPSCHITZ EMBEDDINGS IN BANACH SPACES

BY

JUAN ARIAS-DE-REYNA AND LUIS RODRÍGUEZ-PIAZZA

Departamento de Análisis Matemático
Universidad de Sevilla, Apdo. 1160. 41080 Sevilla, Spain

ABSTRACT

We construct a sequence of metric spaces (M_n) with card $M_n = n$ satisfying that for every $c < 2$, there exists a real number $a(c) > 0$ such that, if the Lipschitz distance from M_n to a subset of a Banach space E is less than c, then $\dim(E) \geq a(c)n$. We also prove several results about embeddings of metric spaces whose non-zero distance values are in the interval [1,2].

1. Introduction

We recall the notion of Lipschitz distance or distortion between two metric spaces M and N; it is defined as:

$$\mathrm{dist}(\mathrm{M},\mathrm{N}) = \inf\{\|F\|_{\mathrm{Lip}}\|F^{-1}\|_{\mathrm{Lip}}; F \text{ bijective map from M onto N}\}.$$

Following Bourgain we will say that two metric spaces M and N are c-lipeomorphic if $\mathrm{dist}(\mathrm{M},\mathrm{N}) \leq c$.

Given a real number $c \geq 1$ let us define $\psi_c(n)$ as the least natural number such that every metric space M of cardinality n is c-lipeomorphic to a subset of a Banach space E whose dimension is $\dim(\mathrm{E}) \leq \psi_c(n)$. We are interested in the asymptotic behavior of $\psi_c(n)$ as n tends to infinity.

Since every metric space M is isometric to a subset of $\mathcal{C}(\mathrm{M})$ it is clear that $\psi_c(n) \leq n-1$. Using volume arguments it is easy to prove that $\psi_c(n)$ increases

Received June 3, 1991

at least as $c \log n$; this fact led Johnson and Lindenstrauss [JL] to ask whether $\psi_c(n) \leq k(c) \log n$. Using random graphs Bourgain [B] disproved this conjecture, establishing the inequality

$$\frac{K}{c^2}\left(\frac{\log n}{\log\log n}\right)^2 \leq \psi_c(n).$$

Recently, Johnson, Lindenstrauss and Schechtman [JLS] proved that there exists a constant K such that, for $0 < \beta < 1$,

$$\psi_{K/\beta}(n) \leq c(\beta) n^{\beta}.$$

Here we are going to prove that for small distortions c we have $\psi_c(n) \sim n$. We will construct a sequence of metric spaces $(\mathbf{M}_n)$, with $\operatorname{card} \mathbf{M}_n = n$ satisfying the following

THEOREM: *For every positive constant $c < 2$, there exists $a(c) > 0$, such that if $\mathbf{M}_n$ is c-lipeomorphic to a subset of a Banach space $\mathbf{E}$, then* $\dim(\mathbf{E}) \geq a(c)n$.

We also prove that the initial conjecture of Johnson and Lindenstrauss holds if we deal with metric spaces whose non-zero distance values are in the interval $[1, 2]$. Indeed we have

THEOREM: *Let $1 \leq d < 2$. Then there exists a constant $c(d)$ such that every metric space with n points and whose non-zero distance values are in the interval $[1, d]$ is isometric to a subset of a Banach space of dimension $\leq c(d) \log n$.*

On the other hand, there exists a sequence of metric spaces $(\mathbf{X}_n)$ such that:

(1) $\mathbf{X}_n$ *has* $2n$ *points.*
(2) *The non-zero distance values of* $\mathbf{X}_n$ *are in the interval* $[1, 2]$.
(3) $\mathbf{X}_n$ *is not isometric to any subset of a Banach space of dimension* $< n - 1$.

2. Bounds on the dimension for small distortion

Let $\Omega = \{-1, 1\}^m$ endowed with the canonical probability. The space $L^\infty(\Omega)$ is isometric to ℓ_∞^n, with $n = 2^m$. The Rademacher functions $x_i \in L^\infty(\Omega)$ are the coordinate functions defined on Ω. The Walsh functions are the products $x_A = \prod_{i \in A} x_i$ where $A \subset \{1, 2, \ldots, m\}$. We can consider these n functions as elements of $L^2(\Omega)$ where this family $(x_A)_A$ is a complete orthonormal system. In fact, when we consider Ω as a group the Walsh functions are its characters.

We prove first the following property of the Walsh functions:

PROPOSITION 1: *Let $0 < \alpha < 1$ and $\mathbf{F} \subset L^2(\Omega)$ a subspace of dimension $q < \alpha 2^m$. Then there exists a Walsh vector x_A such that for every $x \in \mathbf{F}$*

$$\|x - x_A\|_\infty > \sqrt{1-\alpha}.$$

Proof: We compute the mean of the distances of the points x_A to $\mathbf{F}$. For this we choose an orthonormal system $e_1, e_2, \ldots, e_q$ in $\mathbf{F}$. We have

$$d_2(x_A, \mathbf{F})^2 = 1 - \sum_{i=1}^{q} |\langle x_A, e_i\rangle|^2.$$

Then

$$\sum_A d_2(x_A, \mathbf{F})^2 = 2^m - \sum_A \sum_{i=1}^{q} |\langle x_A, e_i\rangle|^2$$
$$= 2^m - \sum_{i=1}^{q} \|e_i\|_2^2 = 2^m - q.$$

So

$$\frac{1}{2^m} \sum_A d_2(x_A, \mathbf{F})^2 = 1 - \frac{q}{2^m}.$$

As we suppose $q < \alpha 2^m$, we obtain that the mean is $> 1 - \alpha$. Hence there exists a certain x_A such that $d_2(x_A, \mathbf{F}) > \sqrt{1-\alpha}$.

It follows that for every space $\mathbf{F}$ of dimension $q < \alpha 2^n$ there exists a point x_A in the set of Walsh functions such that

$$d_\infty(x_A, \mathbf{F}) \geq d_2(x_A, \mathbf{F}) > \sqrt{1-\alpha}. \qquad \blacksquare$$

We are now able to prove the main result. We begin by defining the metric spaces that will play a special role in the proof.

Let $n = 2^m$ and let $\mathbf{M}_{3n}$ be the subset of ℓ_∞^n consisting of the $3n$ points

$$y_i = 2e_i, \qquad z_i = -2e_i, \qquad x_j = \sum_{i=1}^{n} w_{i,j} e_i,$$

where (e_i) is the canonical basis of ℓ_∞^n and $(w_{i,j})$ is the Walsh matrix, associated with $n = 2^m$, that is the matrix whose entries are all equal to 1 or -1 and whose rows are the Walsh system.

For every n, there is only one natural number m such that $3 \cdot 2^m \leq n < 3 \cdot 2^{m+1}$ and we can define the metric space $\mathbf{M}_n$ by adding points to $\mathbf{M}_{3 \cdot 2^m}$ so that all non-zero distance values of $\mathbf{M}_n$ are between 1 and 4. These metric spaces satisfy

THEOREM 2: *For every positive constant $c < 2$, there exists a real number $a(c) > 0$, such that if $\mathbf{M}_n$ is c-lipeomorphic to a subset of a Banach space $\mathbf{E}$, then* $\dim(\mathbf{E}) \geq a(c)n$.

Proof: Obviously we only need to prove the theorem for $\mathbf{M}_{3n}$ being $n = 2^m$. Let $T : \mathbf{M}_{3n} \to \mathbf{E}$ be a mapping such that

$$\frac{1}{c}d(x,y) \leq \|Tx - Ty\| \leq d(x,y),$$

for every pair of points $x, y \in \mathbf{M}$, with $c < 2$ and $\dim \mathbf{E} = k$.

Take $x_i^* \in \mathbf{E}^*$, such that $\|x_i^*\| = 1$ and

$$x_i^*(Ty_i - Tz_i) = \|Ty_i - Tz_i\| \geq \frac{4}{c} = 2 + 2\left(\frac{2-c}{c}\right).$$

The dimension k of $\mathbf{E}$ is greater than or equal to the rank of the matrix $\big(x_i^*(Tx_j)\big)$. This rank changes at most one unit if we subtract $\big(x_i^*(Ty_i) + x_i^*(Tz_i)\big)/2$ from row i. Each point x_j is one unit distant from y_i or z_i according to the sign of $w_{i,j}$. If $d(x_j, y_i) = 1$, we have

$$|x_i^*(Tx_j - Ty_i)| \leq d(x_j, y_i) = 1.$$

So

$$\begin{aligned} &x_i^*(Tx_j) - \big(x_i^*(Ty_i) + x_i^*(Tz_i)\big)/2 \\ \geq\ &\frac{x_i^*(Ty_i) - x_i^*(Tz_i)}{2} - |x_i^*(Tx_j) - x_i^*(Ty_i)| \geq 1 + \frac{2-c}{c} - 1 = \frac{2-c}{c}. \end{aligned}$$

Analogously, if $d(x_j, z_i) = 1$ we obtain

$$x_i^*(Tx_j) - \big(x_i^*(Ty_i) + x_i^*(Tz_i)\big)/2 \leq -\frac{2-c}{c}.$$

In the first case, when $d(x_j, y_i) = 1$ we have

$$\begin{aligned} 0 &\leq x_i^*(Tx_j) - \big(x_i^*(Ty_i) + x_i^*(Tz_i)\big)/2 \\ &= x_i^*(Tx_j) - x_i^*(Tz_i) - \frac{x_i^*(Ty_i) - x_i^*(Tz_i)}{2} \leq 3 - \frac{2}{c} \end{aligned}$$

and when $d(x_j, z_i) = 1$ we obtain

$$\begin{aligned} 0 &\geq x_i^*(Tx_j) - \big(x_i^*(Ty_i) + x_i^*(Tz_i)\big)/2 \\ &= x_i^*(Tx_j) - x_i^*(Ty_i) + \frac{x_i^*(Ty_i) - x_i^*(Tz_i)}{2} \geq -3 + \frac{2}{c}. \end{aligned}$$

It follows that the entries of the matrix

$$A = \left(x_i^*(Tx_j) - \frac{x_i^*(Ty_i) + x_i^*(Tz_i)}{2} \right)$$

have absolute values between $(2-c)/c$ and $3-2/c$, and the same sign as those of $(w_{i,j})$; that is the $\|\cdot\|_\infty$-distance from $(w_{i,j})$ to A is less than or equal to $2-2/c$. We know, from Proposition 1, that if $\mathbf{F}$ is a subspace of dimension q and $q < \alpha 2^m$ there exists a vector x_A in the Walsh system such that $d_\infty(x_A, \mathbf{F}) > \sqrt{1-\alpha}$. It follows that

$$k+1 \geq \left(1-(2-2/c)^2\right)n = \frac{(3c-2)(2-c)}{c^2}n. \quad \blacksquare$$

3. Metric spaces with restrictions on the distance

We give first a result that allows one to give low-dimensional embeddings of certain subsets of ℓ_∞^m.

THEOREM 3: *Given $\varepsilon > 0$, there exists a constant $K > 0$ such that if $\mathbf{M}$ is a subset of ℓ_∞^m with n points, where $n \leq m$, $\alpha > 0$, and*

(a) *the points $\mathbf{a}_i$ of $\mathbf{M}$ have coordinates $(a_{ij})_{j=1}^m$ satisfying $\|(a_{ij})_{i=1}^n\|_2 \leq \alpha$ for every j, $1 \leq j \leq m$,*

(b) *$1 \leq d(x,y)$ for every pair of different points of $\mathbf{M}$,*

then there exists a Banach space $\mathbf{E}$ with

$$\dim \mathbf{E} \leq K(\varepsilon)\alpha^2 \log(1+n)$$

that contains a subset $(1+\varepsilon)$-lipeomorphic to $\mathbf{M}$.

Proof: Recall the definition of the Kolmogorov numbers of an operator $u : \mathbf{X} \to \mathbf{Y}$ between two Banach spaces

$$d_k(u) = \inf\left\{ \|Q_{\mathbf{S}} u\| \,\middle|\, \mathbf{S} \subset \mathbf{Y}, \dim \mathbf{S} < k \right\},$$

where $Q_{\mathbf{S}}$ is the quotient operator from $\mathbf{Y}$ to $\mathbf{Y}/\mathbf{S}$. For the identity operator $i_n : \ell_2^n \to \ell_\infty^n$ Garnaev and Gluskin [GG] obtained the bound

$$d_k(i_n) \leq \min\left\{ 1, \left(c\frac{\log(1+n/k)}{k} \right)^{1/2} \right\}.$$

Let $\mathbf{b}_j$ be the points of ℓ_2^n whose coordinates are $(a_{i,j})_{i=1}^n$. By condition (a) we know that $\|\mathbf{b}_j\|_2 \leq \alpha$.

Fix $\delta > 0$ and let r be

$$r = \frac{\alpha^2 \log(1+n)}{c^2\delta^2},$$

where c is the constant in the inequality of Garnaev and Gluskin. If k is the least natural number greater than r, we obtain

$$d_k(i_n) \leq \frac{\delta}{\alpha}.$$

This means that there exists a space $\mathbf{S}$ of dimension $< k$ (thus $\leq r$), and such that for every unit vector $\mathbf{b}$ of ℓ_2^n there exists a vector $\mathbf{v}$ in $\mathbf{S}$ satisfying $\|\mathbf{b} - \mathbf{v}\|_\infty \leq \delta/\alpha$. So, for every $\mathbf{b}_j$, there exists a vector $\mathbf{v}_j$ such that

$$\left\| \frac{\mathbf{b}_j}{\alpha} - \mathbf{v}_j \right\|_\infty \leq \frac{\delta}{\alpha}.$$

That is, for every $\mathbf{b}_j$, there exists a vector $\mathbf{u}_j \in \mathbf{S}$ such that

$$\|\mathbf{b}_j - \mathbf{u}_j\|_\infty \leq \delta.$$

If $\mathbf{u}_j = (u_{i,j})_{i=1}^n$, we have that the points $\mathbf{x}_i$ of ℓ_∞^m with coordinates $(u_{i,j})_{j=1}^m$ are contained in a space of dimension $\leq r$ and satisfy

$$|d(\mathbf{x}_i, \mathbf{x}_h) - d(\mathbf{a}_i, \mathbf{a}_h)| \leq 2\delta.$$

Since $d(\mathbf{x}_i, \mathbf{x}_h) \geq 1$ for every $i \neq h$, the distortion of the mapping that sends $\mathbf{a}_i$ to $\mathbf{x}_i$ is less than $(1+2\delta)/(1-2\delta)$. Finally, if δ is small enough, this distortion is less than $1+\varepsilon$. ∎

From here we can prove the following

COROLLARY 4: *Given $\varepsilon > 0$, there exists a real number $c(\varepsilon) > 0$ such that for every metric space* $\mathbf{M}$ *with n points, whose non-zero distance values are in the interval $[1,2]$, there exists a Banach space E of dimension* $\dim(\mathbf{E}) \leq c(\varepsilon)\log n$ *that contains a subset $(1+\varepsilon)$-lipeomorphic to* $\mathbf{M}$.

Proof: Let $x_1, x_2, \ldots x_n$ be the n points of $\mathbf{M}$ and put $m = \binom{n}{2}$. There exists an embedding of $\mathbf{M}$ into ℓ_∞^m defined as follows: For every pair (i,j) of natural numbers $i < j \leq n$ put $f_{(i,j)}(x_r) = 0$ if $r \notin \{i,j\}$, $f_{(i,j)}(x_i) = d(x_i, x_j)/2$

and $f_{(i,j)}(x_j) = -d(x_i, x_j)/2$. Finally let $f : \mathbf{M} \to \ell_\infty^m$ be the mapping with coordinates $f_{(i,j)}$. Note that f is an isometry since the non-zero distances of $\mathbf{M}$ are between 1 and 2.

Applying now the above theorem with $\alpha = \sqrt{2}$ we obtain the required result. ∎

Remark: The non-zero distances of the spaces $\mathbf{M}_n$ of Theorem 2 are between 1 and 4. So there exists a space $\mathbf{M}'_n$ 2-lipeomorphic with $\mathbf{M}_n$ and whose non-zero distances are between 1 and 2. By the Corollary it follows that $\mathbf{M}_n$ is $(2+\varepsilon)$-lipeomorphic with a subset of a Banach space of dimension $\le c(\varepsilon)\log n$. ∎

There is a problem that arises naturally observing the properties of these spaces:

PROBLEM: *Does there exist a constant $\alpha > 0$ with the property that for every real number $c > 2$ there exists $b(c) > 0$ such that*

$$\psi_c(n) \le b(c)(\log n)^\alpha?$$

By Bourgain's inequality, this number α must be ≥ 2.

Now we can improve the above Corollary in the following way.

THEOREM 5: *Let $1 \le d < 2$, there exists a constant $c(d)$ such that every metric space with n points whose non-zero distances are in $[1, d]$ embeds isometrically into a Banach space of dimension $\le c(d)\log n$.*

Proof: Let $\mathbf{M}$ be a metric space with n points and consider the embedding into ℓ_∞^m defined in the proof of Corollary 4. We saw there that there exist functions $h_{(i,j)}$ such that $\|h_{(i,j)}\|_\infty \le \varepsilon$ and the vector space generated by the functions $g_{(i,j)} = f_{(i,j)} + h_{(i,j)}$ is of dimension $\le c(\varepsilon)\log n$.

Consider a pair (i,j) and put $g = g_{(i,j)}$ in order to simplify the notation. Choose $\alpha > 0$ so that

$$|\alpha g(x_i) - \alpha g(x_j)| = d(x_i, x_j).$$

We claim that αg is a Lipschitz mapping with Lipschitz constant 1, if we choose $\varepsilon < (2-d)/8$.

In fact, observe that

$$\Big||g(x_i) - g(x_j)| - d(x_i, x_j)\Big| \le 2\varepsilon.$$

So

$$\left|\frac{d(x_i, x_j)}{|g(x_i) - g(x_j)|} - 1\right| \le \frac{2\varepsilon}{|g(x_i) - g(x_j)|} \le \frac{2\varepsilon}{1 - 2\varepsilon}.$$

That is $|\alpha - 1| < 2\varepsilon(1 - 2\varepsilon)^{-1}$. It follows that

$$\alpha < (1 - 2\varepsilon)^{-1}. \tag{1}$$

To prove that $\|\alpha g\|_{\mathrm{Lip}} = 1$ it suffices to prove that $|\alpha g(x_i) - \alpha g(z)| \le d(x_i, z)$. As $|g(z)| \le \varepsilon$ and $g(x_i) \le (d/2) + \varepsilon$, we only need

$$\alpha\left(\frac{d}{2} + 2\varepsilon\right) < 1.$$

(1) then leads to the condition

$$(1 - 2\varepsilon)^{-1}\left(\frac{d}{2} + 2\varepsilon\right) < 1,$$

that is equivalent to $\varepsilon < (2 - d)/8$. ∎

Finally we will prove that the above Theorem cannot be extended to the case $d = 2$. We define the metric space $\mathbf{X}_n$ as the subset of ℓ_∞^n consisting of the basic vectors e_i and the vectors $x_i = b - 2e_i$, where $b = (1, 1, \ldots, 1)$. It is easy to see that all the distance values are 0, 1 or 2.

THEOREM 6: *The metric space $\mathbf{X}_n$ cannot be embedded isometrically in a Banach space of dimension $< n - 1$.*

Proof: Let $J : \mathbf{X}_n \to \mathbf{E}$ be an isometric mapping into a Banach space. Then there exists a linear form $x_i^* \in \mathbf{E}^*$, with $\|x_i^*\| = 1$ and such that $x_i^*(J(e_i) - J(x_i)) = d(e_i, x_i) = 2$. The composition $f_i = x_i^* \circ J$ is a Lipschitz mapping with Lipschitz constant 1. We know that $f_i(e_i) - f_i(x_i) = 2$ and for every $j \ne i$

$$|f_i(x_i) - f_i(e_j)| \le 1, \qquad |f_i(e_i) - f_i(e_j)| \le 1.$$

So the matrix

$$\left(f_i(e_j) - \frac{f_i(x_i) + f_i(e_i)}{2}\right)_{i,j=1}^n$$

is the identity matrix and the rank of the matrix $(f_i(e_j))$ is greater than or equal to $n - 1$; but it is clear that this rank is less than the dimension of $\mathbf{E}$. ∎

References

[B] J. Bourgain, *On Lipschitz embedding of finite metric spaces in Hilbert space*, Isr. J. Math. **52** (1985), 46–52.

[GG] A. Yu. Garnaev and E. D. Gluskin, *On width of the euclidean ball*, Soviet Math. Dokl. **30** (1984), 200–204.

[JL] W. B. Johnson and J. Lindenstrauss, *Extension of Lipschitz mappings into a Hilbert space*, Contemp. Math. **26** (1984), 189–206.

[JLS] W. B. Johnson, J. Lindenstrauss and G. Schechtman, *On Lipschitz embedding of finite metric spaces in low dimensional normed spaces*, in *Geometric Aspects of Functional Analysis*, Israel Seminar 1985–86, Lecture Notes in Math. **1267**, Springer, New York, 1987, pp. 177–184.