ISRAEL JOURNAL OF MATHEMATICS, Vol. 52, Nos. 1-2, 1985

ON LIPSCHITZ EMBEDDING OF FINITE METRIC SPACES IN HILBERT SPACE

BY

J. BOURGAIN

Department of Mathematics, Vrije Universiteit Brussel, Pleinlaan 2, 10F7, 1050, Brussels, Belgium

ABSTRACT

It is shown that any n point metric space is up to $\log n$ lipeomorphic to a subset of Hilbert space. We also exhibit an example of an n point metric space which cannot be embedded in Hilbert space with distortion less than $(\log n)/(\log\log n)$, showing that the positive result is essentially best possible. The methods used are of probabilistic nature. For instance, to construct our example, we make use of random graphs.

1. Introduction and statement of results

This note fits in the program of investigating the geometry of finite metric spaces developed more intensively over the last years. Among the motivations for this research, let us mention a relation with work of M. Gromov on Riemannian manifolds [2] and the development of the non-linear theory of Banach spaces (cf. [1]). We refer the reader to the survey of J. Lindenstrauss [4] for a detailed exposition of this theme.

Several notions appearing in the theory of finite-dimensional normed spaces can be reformulated for finite metric spaces. The linear isomorphisms are replaced by bi-Lipschitz maps and the so-called Banach–Mazur distance by the Lipschitz distance or distortion. Thus, denoting "cardinal" by $|\ |$

DEFINITION. Let X, d and Y, δ be (finite) metric spaces, $|X| = |Y|$. Let

$$\operatorname{dist}(X, Y) = \inf\{\|F\|_{\mathrm{Lip}}\|F^{-1}\|_{\mathrm{Lip}}\,;\ F \text{ one-one map from } X \text{ onto } Y\}$$

where

$$\|F\|_{\mathrm{Lip}} = \sup_{s \neq t \text{ in } X} \frac{\delta(F(s), F(t))}{d(s,t)}$$

Received August 27, 1984

The by now classical theorem of F. John asserts that any n-dimensional normed space is at distance at most $\sqrt{n}$ from Hilbert space. Since the logarithm of the cardinal of the (finite) metric space is the analogue of the dimension of the linear space (for reason of entropy), the natural question arising is whether or not given a finite metric space X, there exists a subset Y of Hilbert space satisfying $|X| = |Y|$ and $\operatorname{dist}(X, Y) \leqq (\log |X|)^{1/2}$. A result due to W. B. Johnson and J. Lindenstrauss (see [3]) asserts that this set Y can then be chosen in a Hilbert space of dimension at most $C. \log |X|$, a fact which will be exploited in this paper.

The previous embedding problem is by now almost completely solved in following two statements:

PROPOSITION 1. *Given a finite metric space X, d there exists a one-to-one map F from X into Hilbert space such that*

$$\|F\|_{\text{Lip}} \|F^{-1}\|_{\text{Lip}} \leqq C \log |X|.$$

PROPOSITION 2. *There exists a metric space X, d of cardinal $|X| = n$ (for any positive integer n) so that the distance from X to an n-point subset of Hilbert space is at least $c(\log n)/(\log\log n)$.*

In Propositions 1 and 2, the letters $C < \infty$, $c > 0$ stand for numerical constants.

The proof of Proposition 1 depends on the following inequality.

PROPOSITION 3. *Let X, d be a finite metric space and denote for each positive number s by $\mathcal{P}_s$ the set of all subsets of X with cardinal $[s]$. There is a numerical constant C such that for each pair x, y of points in X*

$$(1) \qquad d(x, y) \leqq C \int_1^{|X|} \frac{1}{s} \left\{ |\mathcal{P}_s|^{-1} \sum_{A \in \mathcal{P}_s} |d(x, A) - d(y, A)| \right\} ds.$$

As usual, $d(x, A) = \inf_{z \in A} d(x, z)$.

From (1), the deduction of Proposition 1 is straightforward. Denote by 2^X the set of all subsets X and consider the map $u : X \to l^\infty_{2^X}$ given by $u(x) = \{d(x, A)\}_{A \subset X}$. The inequality

$$|d(x, A) - d(y, A)| \leqq d(x, y)$$

implies that $\|u\|_{\text{Lip}} \leqq 1$. Let next Λ be the diagonal map from $l^\infty_{2^X}$ to $l^1_{2^X}$ defined by $\Lambda_{(A)} = s^{-1} |\mathcal{P}_s|^{-1}$ where $|A| = s$. Thus Λ is a linear operator and

$$\|\Lambda\|_{l^\infty \to l^1} = \|\Lambda\|_{N(l^\infty, l^1)} = \int_1^{|X|} \frac{ds}{s} = \log |X|$$

where $N(.,.)$ refers to the corresponding space of nuclear opeartors (see [5] for definitions and basic theory). Clearly

$$\|\Lambda \circ u\|_{\mathrm{Lip}(X, l^1_{2^X})} \leqq \log |X|$$

and inequality (1) asserts that

$$\|(\Lambda \circ u)^{-1}\|_{\mathrm{Lip}} \leqq C. \tag{2}$$

Notice at this point that we just obtained an analogue of the fact that given an n-dimensional normed space E, there is a linear map $j : E \to l^1$ satisfying $\|j\|_{N(E,l^1)} \leqq n$ and $\|j^{-1}\| \leqq 1$. This result is best possible.

To complete the proof of Proposition 1, factor Λ through $l^2_{2^X}$ as $\Lambda = \Lambda''\Lambda'$ where $\Lambda' : l^\infty_{2^X} \to l^2_{2^X}$, $\Lambda'' : l^2_{2^X} \to l^1_{2^X}$ are diagonal operators defined by $\Lambda'_{(A)} = \Lambda''_{(A)} = \Lambda^{1/2}_{(A)}$. Then

$$\|\Lambda'\| \leqq (\log |X|)^{1/2} \quad \text{and} \quad \|\Lambda''\| \leqq (\log |X|)^{1/2}$$

and hence, in view of (2),

$$\|\Lambda' \circ u\|_{\mathrm{Lip}(X, l^2_{2^X})} \leqq (\log |X|)^{1/2}; \quad \|(\Lambda' \circ u)^{-1}\|_{\mathrm{Lip}} \leqq C(\log |X|)^{1/2}.$$

This completes the proof.

Proposition 2 now implies that there is no metric analogue of the F. John ellipsoid for normed spaces, or equivalently the $(\dim E)^{1/2}$ estimates for the 2-summing norm of the identity operator (see again [5] for details). By the distance estimate (in the linear and hence Lipschitz sense)

$$\mathrm{dist}(E, l^2_{\dim E}) \leqq (\dim E)^{1/2}$$

we get also as a consequence of Proposition 2

COROLLARY 4. *There is no uniform estimation for the distortion of n-point metric spaces to a subset of some* $(\log n)$*-dimensional normed space.*

This gives a negative solution to a question raised in [3]. Notice that the map $F(x) = \{d(x,y)\}_{y \in X}$ gives a Lipschitz embedding of X, d into $l^\infty_{|X|}$ with $\|F\|_{\mathrm{Lip}} \|F^{-1}\|_{\mathrm{Lip}} = 1$. In particular, if we denote by $\psi(n)$ the smallest integer such that any n-point metric space is at Lipschitz distance at most 2 from a subset of some $\psi(n)$-dimensional normed space, we get

$$\left\{\frac{\log n}{\log \log n}\right\}^2 \leqq \psi(n) \leqq n.$$

At the time of writing, no more information on ψ seems to be known.

If only embeddings in l_k^∞-spaces are allowed, some n-point metric spaces will require k as large as $\log k \sim \log n$. This happens for instance if we take for X an n-point set in the unit sphere of l_m^2, $m = c(\log n)$, forming a ρ-net for the Euclidean distance. If $F: X \to l_k^\infty$, then by extending the coordinate maps to Lipschitz maps on the l_m^2-sphere, one gets indeed

$$\log k \geqq \rho \log n \qquad \text{whenever } \|F\|_{\text{lip}} \|F^{-1}\|_{\text{lip}} \leqq c/\rho$$

as a consequence of measure-concentration phenomenon on Euclidean spheres (see [3] for details on this subject).

2. Proof of Proposition 3

Inequality (1) may clearly be rewritten in the form

$$\sum_{p=1}^{\log|X|} \left\{ |\mathcal{P}_{2^p}|^{-1} \sum_{A \in \mathcal{P}_{2^p}} |d(x,A) - d(y,A)| \right\} \geqq cd(x,y). \tag{3}$$

Let us introduce some notation. Denote for $x \in X$, $\rho > 0$ by $B(x,\rho)$ the ball in X, d with midpoint x and radius ρ, thus

$$B(x,\rho) = \{y \in X;\ d(x,y) < \rho\}$$

and consider $\varphi_x(\rho) = |B(x,\rho)|$, which is a decreasing function of ρ.

Fix now distinct points x, y in X, $d(x,y) = \varepsilon > 0$. Let $\{\rho_t\}_{0 \leqq t \leqq t^*}$ be the increasing sequence of positive numbers defined by

$$\begin{aligned} &\rho_0 = 0, \\ &\rho_t = \inf\{\rho > 0;\ \varphi_x(\rho) \geqq 2^t \text{ and } \varphi_y(\rho) \geqq 2^t\} \qquad (t < t^*), \\ &\rho_{t^*} = \varepsilon/2. \end{aligned}$$

Denoting φ^+ (resp. φ^-) the right (resp. left) limit, it follows that

$$\varphi_x^+(\rho_t) \geqq 2^t \quad \text{and } \varphi_y^+(\rho_t) \geqq 2^t \qquad \text{for } 0 \leqq t < t^*,$$

$$\varphi_x^-(\rho_t) \leqq 2^t \quad \text{or} \quad \varphi_y^-(\rho_t) \leqq 2^t \qquad \text{for } 0 \leqq t \leqq t^*.$$

Fix $0 < t \leqq t^*$. Take $s = 2^p$ such that $s.2^t \sim n/10$. Let $0 < \rho < \rho_t$ and $0 < \delta < \rho_t$. Suppose $\varphi_x^-(\rho_t) \leqq \varphi_y^-(\rho_t)$, hence $|B(x,\rho)| < 2^t$. Since $B(x,\rho_t) \cap B(y,\rho_t) = \varnothing$, we may write by elementary probabilistic considerations

$$\begin{aligned} &|\mathcal{P}_s|^{-1} \sum_{A \in \mathcal{P}_s} |d(x,A) - d(y,A)| \\ &\quad \geqq (\rho - \rho_t + \delta) |\mathcal{P}_s|^{-1} |\{A \in \mathcal{P}_s;\ A \cap B(x,\rho) = \varnothing \text{ and } A \cap B(y, \rho_t - \delta) \neq \varnothing\}| \\ &\quad \geqq c(\rho - \rho_t + \delta) |\mathcal{P}_s|^{-1} |\{A \in \mathcal{P}_s;\ A \cap B(y,\rho_t - \delta) \neq \varnothing\}|. \end{aligned}$$

Now, if $\delta < \rho_t - \rho_{t-1}$, $|B(y, \rho_t - \delta)| \geqq \varphi_y^+(\rho_{t-1}) \geqq 2^{t-1}$ by construction. Consequently, letting $\rho \to \rho_t$

$$|\mathcal{P}_s|^{-1} \sum_{A \in \mathcal{P}_s} |d(x, A) - d(y, A)| \geqq c(\rho_t - \rho_{t-1}), \qquad s = 2^p. \tag{4}$$

Notice that the values of p corresponding to $t = 1, \dots, t^*$ may be chosen distinct.

Clearly, summutation of (4) over t yields (3). This completes the proof.

3. Metric spaces associated to graphs; proof of Proposition 2

A (non-oriented) graph G on a set X (which elements are called the vertices of G) is a collection of pairs (the edges of G) of distinct elements of X. A path Q in G is a chain of consecutive edges; their number is the path length $|Q|$ of Q. If G is a connected graph, i.e. any two points in X are endpoints of some chain in G, G defines a metric on X,

$$d_G(x, y) = \inf\{|Q|;\ Q \text{ is a path in } G \text{ connecting } x \text{ and } y\}.$$

Our approach to Proposition 2 consists in analysing metric spaces associated to random graphs.

Let G be a random graph on n vertices with edge probability $\delta = M/n$, where M is to be fixed later. Denote $\mathcal{G}$ the sample space, which identifies with $\{0,1\}^{\binom{n}{2}}$ equipped with product measure $\mathbf{P}$.

LEMMA 1. *If*

$$k < \lambda \equiv c \frac{\log n}{\log M} \quad \textit{and} \quad x \neq y \textit{ in } X,$$

then

$$\mathbf{P}_{\mathcal{G}}[d_G(x, y) \leqq k] < \frac{1}{200}.$$

PROOF. Estimate $\mathbf{P}_{\mathcal{G}}[d_G(x, y) \leqq k]$ by $\sum_{j=0}^{k-1} \binom{n-2}{j} \delta^{j+1}$.

Next, let

$$\mathcal{H}_1 = \{G \in \mathcal{G};\ d_G(x, y) \geqq \tfrac{1}{2}\lambda \text{ for at least } \tfrac{1}{4}n^2 \text{ pairs } \{x, y\}\}.$$

LEMMA 2. $\mathbf{P}_{\mathcal{G}}[\mathcal{H}_1] > 1 - \frac{1}{50}$.

PROOF. Denoting Δ_X the diagonal of X, write

$$\int_{\mathcal{G}} \sum_{\mathcal{P}_2(X) \setminus \Delta_X} \min\{d_G(x, y), \lambda\} \geqq \binom{n}{2} \lambda \left\{1 - \sup_{x \neq y} \mathbf{P}[d_G(x, y) < \lambda]\right\}$$

from where, by Lemma 1,

$$\binom{n}{2}\lambda\,\mathbf{P}[\mathcal{H}_1]+\left\{\binom{n}{2}\frac{\lambda}{2}+\frac{n^2}{8}\lambda\right\}\mathbf{P}[\mathcal{G}\setminus\mathcal{H}_1]\geqq\frac{199}{200}\binom{n}{2}\lambda.$$

The minoration on $\mathbf{P}[\mathcal{H}_1]$ follows.

We now examine the probability for connectivity of G. Define $\mathcal{H}_2=\{G\in\mathcal{G};\ G \text{ connected}\}$.

LEMMA 3. *$\mathbf{P}_{\mathcal{G}}[\mathcal{H}_2]>1-\frac{1}{50}$ provided $M>C\log n$.*

PROOF. The fact that G is not connected is equivalent to the existence of a nontrivial subset A of X such that no point in A is connected to a point in $X\setminus A$ by an edge in G. Hence

$$\mathbf{P}[\mathcal{G}\setminus\mathcal{H}_2]\leqq\sum_{j=1}^{(n-1)}\binom{n}{j}(1-\delta)^{j(n-j)}$$

implying the lemma.

As a consequence of these observations

LEMMA 4. *Fix $C\log n<M\ll n$. There is a collection $\mathcal{C}$ of connected graphs on X ($|X|=n$) such that*

(.) $|\mathcal{C}|>\exp cMn$,

(..) $\operatorname{dist}(X,d_G;X,d_{G'})>c\left(\frac{\log n}{\log M}\right)^2$ *if $G\neq G'$ in $\mathcal{C}$.*

PROOF. Defining $\mathcal{H}=\mathcal{H}_1\cap\mathcal{H}_2$, it follows $\mathbf{P}_{\mathcal{G}}[\mathcal{H}]>\frac{24}{25}$ from Lemmas 2 and 3. Fix an integer $N=\exp cMn$ and consider the product space $\Omega=\bigotimes_{i=1,\dots,N}\mathcal{H}^{(i)}$ where each factor $\mathcal{H}^{(i)}$ is a copy of $\mathcal{H}\subset\mathcal{G}$ equipped with normalized measure. Thus the elements $\vec{G}$ of Ω are sequences of graphs on X. Take

$$\Omega'=\bigcap_{1\leqq i\neq j\leqq N}\ \bigcap_{\substack{f:X\to X\\ f \text{ one-to-one}}}\left\{\vec{G}\in\Omega;\ \|f\|_{\mathrm{Lip}(G_i\to G_j)}>\frac{\lambda}{2}\right\}$$

where G_i (resp. G_j) refers to X, d_{G_i}.

Then, by definition of $\mathcal{H}_1$

$$\mathbf{P}(\Omega\setminus\Omega')\leqq N^2n!\sup_{G'\in\mathcal{H}_1;f}\mathbf{P}_{\mathcal{H}}\left[G;\{x,y\}\notin G \text{ whenever } d_{G'}(f(x),f(y))\geqq\frac{\lambda}{2}\right]$$

$$\leqq 2N^2n!(1-\delta)^{n^2/4}.$$

For appropriate choice of N, it follows that $\Omega \backslash \Omega'$ has small measure and any element $\vec{G}$ in Ω' provides a collection $\mathscr{C}$ satisfying the lemma.

PROOF OF PROPOSITION 2. Let Y be an n-point metric space corresponding to a subset of Hilbert space. It follows from the Johnson–Lindenstrauss result [3] that Y is 2 Lipeomorphic to a subset of l_m^2 where the dimension $m < C \log n$. Let $\mathscr{C}$ be the family of graphs on n vertices given by Lemma 4 for some choice of M. In computing the distortion of X, d_G $(G \in \mathscr{C})$ to a subspace Y or l_m^2, we may suppose

$$\frac{1}{n} \leqq \|s - t\|_2 \leqq n^2 \qquad \text{if } s \neq t \text{ in } Y.$$

Therefore Y may be choosen in a $1/n$-net in the n^2-radius ball of l_m^2, leading to at most $\exp[Cn(\log n)^2]$ samples.

If $M > C(\log n)^2$, $|\mathscr{C}|$ will be larger (by (.) of Lemma 4). By construction and Lemma 4 (..), it follows that for some G in $\mathscr{C}$, X, d_G is only up to $c(\log n)/(\log \log n)$ Lipschitz isomorphic to a subset of Hilbert space.

ACKNOWLEDGMENT

The present paper originates from a visit of the author at the Hebrew University which he thanks for the hospitality. The idea to use random graphs in order to obtain lower bounds (Proposition 2) originates from J. Lindenstrauss.

REFERENCES

1. J. Bourgain, V. Milman and H. Wolfson, *Theorie of type in finite metric spaces*, Trans. Am. Math. Soc., to appear.

2. M. Gromov, *Filling Riemannian manifolds*, J. Diff. Geom. **18** (1983), 1–147.

3. W. Johnson and J. Lindenstrauss, *Extensions of Lipschitz mappings into a Hilbert space*, Contemp. Math. **26** (1984), 189–206.

4. J. Lindenstrauss, Proceedings Missouri Conf., Missouri — Columbia (1984), to appear.

5. A. Pietsch, *Operator Ideals*, Springer-Verlag, Berlin, 1978.