Bi-Lipschitz embeddings into low-dimensional Euclidean spaces

JIŘÍ MATOUŠEK
Department of Computer Science,
Charles University
Malostranské nám. 25, 118 00 Praha 1, Czechoslovakia

Abstract

Let (X, d), (Y, ρ) be metric spaces and $f : X \to Y$ an injective mapping. We put

$$\|f\|_{Lip} = \sup\{\frac{\rho(f(x), f(y))}{d(x,y)}; x, y \in X, x \neq y\},$$

and $dist(f) = \|f\|_{Lip}.\|f^{-1}\|_{Lip}$ (the *distortion* of the mapping f). The distortion can be considered as a measure of "faithfulness" of the mapping from a metric point of view. The *Lipschitz distance* of X from subspaces of Y is defined as follows:

$$dist(X, \subseteq Y) = \inf\{dist(f); f : X \to Y \text{ an injective mapping}\}.$$

We investigate the maximum possible Lipschitz distance of a n-point metric space X from subspaces of k-dimensional Euclidean space E^k (where k is a small fixed number); let us denote this quantity by $f(n, k)$. We obtain the following bounds:

$$\begin{array}{lcll} f(n,k) & = & \Theta(n) & (k = 1, 2) \\ f(n,k) & = & O(n^{2/k}(\log n)^{3/2}) & (k \geq 3) \\ f(n,k) & = & \Omega(n^{2/k}) & (k \text{ even}), \\ f(n,k) & = & \Omega(n^{2/(k+1)}) & (k \text{ odd}). \end{array}$$

1 Introduction

We shall consider the following type of problem: Given a finite metric space X, embed it into a low-dimensional Euclidean space as faithfully as possible. Such a problem often arises in practical applications, e.g. in numerical taxonomy. A matrix of distances (dissimilarities) among objects is given (measured, computed), and it is desirable to present this information in a more intutitive way, namely by a two- or three-dimensional picture, where the Euclidean distances among points should correspond to the prescribed distances of objects.

It is easily seen that there are finite metric spaces which cannot be isometrically embedded into an Euclidean space (the isometric embeddability into Euclidean and Hilbert spaces has been nicely characterized by Schoenberg – see e.g. [Scho38]), therefore some approximate solution is desirable. This problem has been considered in the literature for quite a long time. The paper [Kru64] proposes a solution, which has been modified and refined many times: Here the "faithfulness" criterion is some norm of a matrix, arising as the *difference* of the prescribed matrix of distances and the matrix of actual distances of points in the embedding under consideration. For many types of matrix norms, efficient algorithm is known for computing an optimal embedding under this criterion (see [Mat85]).

The above criterion has some drawback, namely it tends to emphasize the faithfulness for large distances in the original matrix, and the small distances may be distorted quite a lot in an optimal embedding (this has been observed e.g. in [Hei88]). The motivation of this paper is to investigate the possibilities of good embeddings under another criterion of faithfulness – so-called distortion, which essentially says how much the embedding spoils the distance ratios. Let us introduce the necessary definitions.

Let (X, d), (Y, ρ) be metric spaces and $f : X \to Y$ an injective mapping. We put

$$\|f\|_{Lip} = \sup\{\frac{\rho(f(x), f(y))}{d(x,y)}; x, y \in X, x \neq y\},$$

and $dist(f) = \|f\|_{Lip}.\|f^{-1}\|_{Lip}$ (the *distortion* of the mapping f). The *Lipschitz distance* of X from subspaces of Y is defined as follows:

$$dist(X, \subseteq Y) = \inf\{dist(f); f : X \to Y \text{ an injective mapping }\}.$$

Let us introduce one more notion needed in the proofs.

We call a mapping $f : (X, d) \to (Y, \rho)$ *non-contracting* if $\rho(f(x), f(y)) \geq d(x, y)$ for every $x, y \in X$. Obviously in f goes into an Euclidean space, we may modify it to a non-contracting mapping without changing its distortion.

In this paper, we shall consider the following problem: what can be the maximal possible Lipschitz distance of a n-point metric space from subspaces of E^k, where k is a (small) prescribed natural number?

The above introduced notions of distortion and Lipschitz distance and also the interest in Lipschitz distances of finite metric spaces from subspaces of Banach spaces are not new. They appear naturally in so-called Local Banach Space Theory, which investigates the structure of finite-dimensional subspaces of various Banach spaces, in particular the L_p-spaces.

For earlier results, we mention the classical and elegant papers of Enflo [Enf69a, Enf69b]. Both these papers give lower bounds on the Lipschitz distances of certain finite metric spaces from subspaces of Hilbert space, and this is used to establish non-existence of a uniform embedding of some Banach space into Hilbert space. There are many more papers with such a flavor, and recently the embedding of finite metric spaces into Banach spaces has been investigated as a subject of independent interest (see e.g. [AM83, Bou85, BFM86, BMW86, JLS87]).

As for the embeddings of finite metric spaces into Euclidean spaces, J.Bourgain [Bou85] proved that any n-point metric space can be embedded into E^n with distortion $O(\log n)$. An interesting result ("flattening lemma") is proved in [JL84]; we shall quote this result and use some ingredients of the proof in Section 4.

We know of no previous quantitative results concerning our problem, i.e. the minimum necessary distortion needed to embed an n-point metric space into E^k (for small fixed k). We shall obtain some asymptotic upper and lower bounds; for $k = 1$ and all even k the upper bound essentially matches the lower bound, while for odd $k \geq 3$ there remains a gap. The most surprising result we get is that as for the worst-case distortion, the spaces E^1 and E^2 differ only insignificantly (by a constant multiple), while E^2 and E^3 differ much more substantially.

2 Statement of results

For $k = 1$ we shall get the most exact upper bound by a specific method, which does not seem to generalize to higher dimensions.

Theorem 2.1 *Every n-point metric space can be embedded into E^1 with distortion at most $12n$.*

The above upper bound is also the best one we can get for $k = 2$. For $k \geq 3$ we have the following result:

Theorem 2.2 *Let X be a n-point subset of E^n and let $3 \leq k \leq \ln n$. Then*

$$dist(X, \subseteq E^k) = O(n^{2/k}\sqrt{\log n/k}).$$

Let us remark that the condition $k \leq \ln n$ could be replaced by $k \leq C.\log n$ (then the constant of proportionality in the bound on the Lipschitz distance will depend on C); this would made the calculations slightly less transparent.

It is known that every n-point metric space can be embedded into E^n with distortion $O(\log n)$ [Bou85], while the distortions needed for embeddings into Euclidean spaces of small (fixed) dimension are of much larger order. For this reason we have restricted ourselves to the case when X is a subspace of some Euclidean space in the above theorem.

Let us now turn to the lower bound question. An easy way to get a lower bound is a "volume argument": Let D_n denote a n-point metric space, in which every two distinct points have distance 1. A simple consideration shows that any embedding of D_n into E^k must have distortion at least $\Omega(n^{1/k})$:

Suppose that $f : D_n \to E^k$ is a non-contracting map with Lipschitz constant L. Choose $x_0 \in D_n$. Each $f(y)$ lies in the ball $B(f(x_0), L)$, and so each ball $B(f(y), 1/2)$ is contained in $B(f(x_0), L + 1/2)$. At the same time the balls $B(f(y), 1/2)$ are disjoint for distinct y's, and so their total volume does not exceed the volume of $B(f(x_0), L + 1/2)$. This gives $(2L + 1)^k \geq n$, or $L \geq (n^{1/k} - 1)/2$.

The above lower bound is asymptotically tight for $k = 1$. However, for larger k there are more serious obstructions to a "nice" embedding of a finite metric space into E^k, given by topological properties of E^k (namely by the non-realizability of some m-dimensional simplicial complexes in E^{2m}). This method also yields the "bad" space for E^k lying in E^{k+1} or E^{k+2}.

Theorem 2.3 *For every fixed $m \geq 1$ and every n, there exists a n-point subspace X of E^{2m+1}, with*

$$dist(X, \subseteq E^{2m}) = \Omega(n^{1/m})$$

.

It seems that the metric spaces from the previous theorem, requiring distortion of order $n^{1/m}$ for embedding into E^{2m}, are pathological in some sense and that for "typical" n-point metric spaces the necessary distortion is much smaller. This cannot be fully formalized, but one can show e.g. the following:

Let $1 \leq k < m$ be fixed integers; put $d = (\lceil m/k \rceil - 1)/m$. Let a metric space X arise by an independent random choice of n points from the uniform distribution in the unit cube of E^m. Then for every fixed $\varepsilon > 0$, the probability $Prob(dist(X, \subseteq E^k) > n^{d+\varepsilon})$ tends to 0 for $n \to \infty$.

So, in particular, although there exist n-point subsets of E^3, whose embedding into the plane requires the distortion of linear order, a randomly chosen n-point subset of the unit cube in E^3 typically admits embedding with distortion of order at most $n^{1/3+\varepsilon}$.

We shall not give a proof of the above statement here; it can be found in [Ma89].

3 Embedding into real line

In this section we shall prove Theorem 2.1.

For the following considerations, it will be useful to imagine the metric space (X, ρ) as a complete graph on $|X|$ vertices, whose edges are labeled by the corresponding distances. This weighted graph will be denoted also by X, and the weight of an edge will be called its *length*. A graph $G(X, \leq d)$ (where d is a positive number) arises from X by deleting all edges with length greater than d.

Further let us denote

$$\sigma(X) = \inf\{d; G(X, \leq d) \text{ is connected}\}.$$

Obviously $\sigma(X)$ can also be expressed as

$$\sigma(X) = \inf\{\rho(A, B);\ A \neq \emptyset \neq B, (A, B) \text{ is a partition of } X\},$$

and from this one sees that $\sigma(X)$ is also the length of the longest edge of any minimum spanning tree for X.

First we prove the following

Lemma 3.1 *Let (X, ρ) be a n-point metric space, $v \in X$ a point of X. Let us put $A = 2n\sigma(X)$. Then there exists a non-contracting mapping $\phi : X \to [0, A]$, such that $\phi(v) = 0$ and for every $x \in X$ the following holds:*

$$\phi(x) \geq \rho(v, x), \qquad A - \phi(x) \geq \rho(v, x).$$

Proof: Let T be a minimum spanning tree for X. T has $n-1$ edges and the length of every edge is at most $\sigma(X)$, hence the total length of edges of X is at most $(n-1)\sigma(X) < A/2$. It is well known that any weighted graph, whose weights satisfy the triangle inequality, contains a Hamiltonian circuit of length at most $2D$, where D denotes the total edge length of a minimum spanning tree for this graph. So let $H = (v = x_1, x_2, \ldots, x_n, x_1 = v)$ be a Hamiltonian circuit of length at most A, i.e.

$$\rho(x_1, x_2) + \rho(x_2, x_3) + \ldots + \rho(x_n, x_1) \leq A.$$

Let us define

$$\phi(x_i) = \rho(x_1, x_2) + \rho(x_2, x_3) + \ldots + \rho(x_{i-1}, x_i), \qquad i = 1, \ldots, n.$$

The triangle inequality gives for every pair x_i, x_j $(i < j)$

$$\rho(x_i, x_j) \leq \rho(x_i, x_{i+1}) + \rho(x_{i+1}, x_{i+2}) + \cdots \rho(x_{j-1}, x_j) = |\phi(x_i) - \phi(x_j)|,$$

so that the mapping ϕ is non-contracting. The condition involving the vertex v can be verified similarly. $\Box$

Proof of Theorem 2.1 : In the proof, we shall work with *rooted trees.* A rooted tree arises from an usual graph-theoretic tree by choosing one vertex as the *root*. Then we say that a vertex u is a *successor* of v if v lies on the (unique) path from u to the root; u is a *son* of v if it is its a successor and u, v are connected by an edge. The inverse relation to successor (resp. son) is called *predecessor* (resp. *father*).

We shall prove the following stronger statement by induction on n:

Statement 3.2 *Let (X, ρ) be a n-point metric space and $v \in X$ some its point. Let us put $B = 6n\sigma(X)$. Then there exists a non-contracting mapping $f : X \to [0, B]$ such that $\|f\|_{Lip} \leq 12n$ and for every $x \in X$ it is*

$$f(x) \geq \rho(x, v), \quad B - f(x) \geq \rho(x, v).$$

This statement is obviously true for $n = 1$ (in this case we put $\sigma(X) = 0$).

Let $|X| = n > 1$, $\sigma = \sigma(X)$ and let T be a fixed minimum spanning tree of X. Let us consider the graph $G(X, \leq \sigma/2)$. This graph is disconnected; let $\mathcal{K}$ denote the set of all its components.

The following is a simple but important observation: The set of vertices of every component $K \in \mathcal{K}$ induces a connected subgraph in T (in other words, if there is a path between two vertices u, v in the graph G, in which every edge has length at most $\sigma/2$, then also the unique path between u and v in the spanning tree T has this property). It is easily seen that if T did not have this property, then we would be able to find a spanning tree of smaller total edge length.

We shall now define an auxiliary rooted tree S. The set of vertices will be $V(S) = \mathcal{K}$. The root of S is chosen arbitrarily (let it be a component R). Components $K, L \in \mathcal{K}$ will be joined by an edge in S iff there is an edge joining them in the tree T. For every component $K \in \mathcal{K}$ we shall fix one its vertex v_K: For R, v_R is chosen arbitrarily, and for the remaining components K, v_K will be the vertex of K which is joined by an edge in T to a vertex of the father of K in S.

For every $K \in \mathcal{K}$ we have $\sigma(K) \leq \sigma/2$ and $|K| < n$. By the inductive hypothesis (Statement 3.2) we can choose for every K a non-contracting mapping $f_K : K \to [0, 3|K|\sigma]$ with Lipschitz constant at most $12|K|$, such that for every point $x \in K$

$$f_K(x) \geq \rho(x, v_K), \quad 3\sigma|K| - f_K(x) \geq \rho(x, v_K).$$

The required embedding of the metric space X into E^1 will now be built inductively, proceeding from the leaves of the tree S to its root, using the embeddings of the components we already have at our disposal.

Let a component K be a vertex of S and let S_K be the subtree of S determined by K (i.e. containing just the successors of K in S, including K). Further, let us denote by U_K the set of all vertices of the components contained in the subtree S_K.

Let $K \in \mathcal{K}$. We shall prove the following statement:

Statement 3.3 *There exists a mapping $g_K : U_K \to [0, B_K]$, where $B_K \leq 6\sigma|U_K|$, satisfying*

(i) g_K is non-contracting

(ii) $\|g\|_{Lip} \leq 12|U_K|$

(iii) For every $x \in U_K$ the following holds:

$$g_K(x) \geq \rho(x, v_K) + \sigma, \quad B_K - g_K(x) \geq \rho(x, v_K) + \sigma.$$

The validity of this statement for $K = R$ (the root of S) already implies Statement 3.2.

The proof of 3.3 is easy when K is a leaf of the tree S. Then it suffices to put $g_K(x) = f_K(x) + \sigma$ for every $x \in U_K$ (in (iii), we use the inequality $3\sigma|U_K| \leq 6\sigma|U_K| - 2\sigma$).

So let us assume that K is an inner vertex of S and let Statement 3.3 hold for each of its sons in S. Further let $\phi : K \to [0, |K|\sigma]$ be the non-contracting mapping guaranteed by Lemma 3.1 for $X = K$ and $v = v_K$ (here one again uses $\sigma(K) \leq \sigma/2$). Let us denote the vertices of K by $u_1, \ldots, u_q$ in such a way that $\phi(u_1) < \phi(u_2) < \ldots < \phi(u_q)$ (so $v_K = u_1$). For $j = 1, 2, \ldots, q$, let $M_j = \{K_{j,1}, K_{j,2}, \ldots, K_{j,m(j)}\}$ be the set of those components of S_K, which are directly connected by an edge in T to the vertex u_j. Let $M = M_1 \cup M_2 \cup \ldots \cup M_q$ (i.e. M contains just all sons of K in S).

Now we shall define auxiliary numbers y_{ij} and the mapping g_K as follows:

$$\begin{array}{rcll}
g_K(x) & = & f_K(x)+\sigma & (x\in K)\\
y_{1,1} & = & \sigma+3\sigma|K|+\phi(u_1)=\sigma+3\sigma|K| & \\
g_K(x) & = & g_{K_{1,1}}(x)+y_{1,1} & (x\in K_{1,1})\\
y_{1,2} & = & y_{1,1}+6\sigma|U_{K_{1,1}}| & \\
g_K(x) & = & g_{K_{1,2}}(x)+y_{1,2} & (x\in K_{1,2})\\
y_{1,3} & = & y_{1,2}+6\sigma|U_{K_{1,2}}| & \\
 & & \vdots & \\
g_K(x) & = & g_{K_{1,m(1)}}(x)+y_{1,m(1)} & (x\in K_{1,m(1)})\\
y_{2,1} & = & y_{1,m(1)}+6\sigma|U_{K_{1,m(1)}}|+\phi(u_2)-\phi(u_1) & \\
 & & \vdots & \\
g_K(x) & = & g_{K_{2,1}}(x)+y_{2,1} & (x\in K_{2,1})\\
y_{2,2} & = & y_{2,1}+6\sigma|U_{K_{2,1}}| & \\
 & & \vdots & \\
g_K(x) & = & g_{K_{q-1,m(q-1)}}(x)+y_{q-1,m(q-1)} & (x\in K_{q-1,m(q-1)})\\
y_{q,1} & = & y_{q-1,m(q-1)}+6\sigma|U_{K_{q-1,m(q-1)}}|+\phi(u_q)-\phi(u_{q-1}) & \\
 & & \vdots & \\
g_K(x) & = & g_{K_{q,m(q)}}(x)+y_{q,m(q)} & (x\in K_{q,m(q)})\\
B_K & = & y_{q,m(q)}+6\sigma|U_{K_{q,m(q)}}|+|K|\sigma-\phi(u_q)+\sigma &
\end{array}$$

The total length of the interval containing the image of g_K according to the above scheme is

$$B_K\leq 2\sigma+3\sigma|K|+\sigma|K|+\sum_{L\in M}(6\sigma|L|)\leq 6\sigma|U_K|-3\sigma|K|+\sigma|K|+2\sigma\leq 6\sigma|U_K|.$$

It remains to verify that the above defined mapping g_K satisfies all the conditions of Statement 3.3.

Let x,y be two distinct points of U_K. If both x,y lie in the same set U_L ($L\in M$), resp. they both lie in K, then (i),(ii) follows from the assumptions on g_L (resp. on f_K). Suppose that none of these cases occurs. It is $\rho(x,y)\geq\sigma/2$, while $|g_K(x)-g_K(y)|\leq 6\sigma|U_K|$ (as it was shown above), hence $\|g_K\|_{Lip}\leq 12|U_K|$, which gives (ii) for these x,y.

Let us now verify (i) for x,y. Let us assume that e.g. $x\in U_P$, where P is a component from M_a, and $y\in U_Q, Q\in M_b$. Then we have

$$\rho(x,y)\leq\rho(x,v_P)+\rho(v_P,u_a)+\rho(u_a,u_b)+\rho(u_b,v_Q)+\rho(v_Q,y).$$

The first addend is (by the assumptions on g_P) at most $\min(g_P(x),B_P-g_P(x))-\sigma$, the second one is at most σ and the third one is bounded from above by $|\phi(u_b)-\phi(u_a)|$. The fourth and fifth addend are estimated similarly as the second and first one, obtaining the bound

$$\rho(x,y)\leq\min(g_P(x),B_P-g_P(x))+|\phi(u_b)-\phi(u_a)|+\min(g_Q(y),B_Q-g_Q(y))$$

in total. At the same time, from the construction one easily gets exactly the above expression as the lower bound for the distance of $g_K(x)$ and $g_K(y)$. A similar reasoning proves (i) also in case when $x\in K$, $y\in L\in M_j$.

The reader who followed the proof upto this point will easily verify also the condition (iii) along similar lines.□

The constant of proportionality in Theorem 2.1 can be improved by a more precise reasoning in the above proof, but it cannot be lowered to 1 by the above used method. Nevertheless, it seems that for every n-point metric space X there might exist an embedding into real line with distortion and most $n-1$.

4 Upper bound on the distortion

In this part we prove an upper bound on $dist(X, \subseteq E_k)$, where X is a n-point metric space. The proof uses the method of random orthogonal projection. Such methods working with a randomly chosen k-dimensional subspace are used in Banach space theory already for quite a long time (see e.g. [FLM77, MiS86]) Johnson a Lindenstrauss used such a method to prove the following result ("flattening lemma" for finite subspaces of Euclidean spaces):

Theorem 4.1 [JL84] *For every $\varepsilon > 0$ there exists a constant $C = C(\varepsilon)$, such that if X is a n-point subset of E^n for some $n \geq 2$, then $dist(X, \subseteq E^{C \log n}) \leq 1+\varepsilon$.* □

Let us remark that Theorem 2.2 does not give us such a precise result for k of order $\log n$; to get this result, we would have to use fines estimates in the proof.

The proof method of [JL84] is the following: one proves, that the set of orthogonal projections $f : E^n \to E^k$, where $k = C \log n$, for which $dist(f|_X) > 1+\varepsilon$, has a smaller measure than the set of all orthogonal projections. The way how this is computed in [JL84] is very elegant, but it is not clear how to use it for distortions of large order. We present here a different way of calculation, which gives (as we shall see later) almost exact result for a fixed k. In the definition of a random projection, we follow [JL84].

Proof of Theorem 2.2: Let $Q_k : E_n \to E_n$ denote the projection on the first k coordinates ($Q_k(x_1, \ldots, x_n) = (x_1, \ldots, x_k, 0, \ldots, 0)$). Let us consider the normalized (probabilistic) Haar measure σ on the orthogonal group $O(n)$ of all rotations of the space E^n around the origin. A *random projection of rank k* will be a random variable F on the probability space $(O(n), \sigma)$, whose values are mappings $E_n \to E_n$, defined by $F(U) = U^* Q_k U, U \in O(n)$. In the sequel we shall write just f for $F(U)$.

For every unit vector $z \in E^n$, we shall bound the probability that a random orthogonal projection deforms it in such a way that its length falls outside an interval $[a, b]$ Due to the rotational symmetry we have for any unit vector z and number $t \in [0, 1]$

$$Prob(\|f(z)\| \leq t) = \mu(x \in S_{n-1}; \|Q_k(x)\| \leq t),$$

where $\|.\|$ stands for the Euclidean norm, S_{n-1} is the unit sphere in E^n and μ is the rotationally invariant probabilistic measure on it.

We shall introduce the notation

$$P_{n,k}(t) = \mu(x \in S_{n-1}; \|Q_k(x)\| \le t).$$

The expression for $P_{n,k}(t)$ is doubtlessly known (e.g. in statistics), but so far the author did not encounter an explicit reference. It is also not too difficult to calculate it, yielding a result

$$P_{n,k}(t) = C_{n,k}\int_0^t x^{k-1}(1-x^2)^{(n-k-2)/2}dx\ , \qquad \text{where} \tag{1}$$

$$C_{n,k} = \left(\int_0^1 x^{k-1}(1-x^2)^{(n-k-2)/2}dx\right)^{-1}. \tag{2}$$

The above integral can also be expressed using the incomplete Beta-function (using a substitution $u = x^2$). This allows numerical estimates of the upper bounds on distortions, obtained by this method. We shall need to estimate $C_{n,k}$, so we express it using Beta-function:

$$C_{n,k}^{-1} = \int_0^1 u^{k/2}(1-u)^{(n-k)/2}du = B(k/2,(n-k)/2),$$

where $B(p,q)$ denotes the Beta-function. Using Stirling's formula and by elementary calculations, one obtains the estimate

$$C_{n,k} \le k(2en/k)^{k/2}. \tag{3}$$

If X is a n-point subset of E^n, the points of X determine at most $\binom{n}{2}$ direction vectors. If a, b will be such numbers that $0 < a < b < 1$ and

$$\binom{n}{2}(P_{n,k}(a) + (1 - P_{n,k}(b)) < 1 \tag{4}$$

then there exists such an orthogonal projection f of rank k, which contracts every distance in X at most a times and at least b times, so $dist(f) \le b/a$.

First we show that for $a = (3n^{2/k}\sqrt{n/k})^{-1}$ it is $\binom{n}{2}.P_{n,k}(a) < 1/2$. Here it suffices to use the simple estimate

$$P_{n,k}(a) \le C_{n,k}\int_0^a x^{k-1}dx \le C_{n,k}a^k/k,$$

and (3).

The estimate for b is a bit more complicated. For every $u \in (0,1)$ the inequality $(1-u) \le e^{-u}$ holds. Using this for $u = x^2$, we get that the expression under the integration sign in (1) is at most $x^{k-1}e^{-cx^2} = e^{(k-1)\ln x - cx}$, where we denoted $c = (n-k-2)/2$. By an elementary calculation one verifies that the function $g(x) = e^{(k-1)\ln x - cx^2}$ has its maximum in the point $m = \sqrt{(k-1)/(n-k-2)} \le \sqrt{k/n}$ and it is decreasing on the interval $(m,1)$, so for $b \ge m$ it is

$$g(b) \ge \int_b^1 g(x)dx \ge (1 - P_{n,k}(b))/C_{n,k}.$$

Let us put $b = 5\sqrt{\ln n/n}$. In the following calculation, we use the bound (3) for $C_{n,k}$ and then the assumption $k \leq \ln n$:

$$\ln(\binom{n}{2}.(1 - P_{n,k}(b))) \;\leq\; 2\ln n + \ln C_{n,k} + \ln g(b) \leq$$

$$2\ln n + \frac{k}{2}(\ln n - \ln k + \ln 2e) + \ln k + \frac{k-1}{2}(\ln\ln n - \ln n + 2\ln 5) - \frac{n-k-2}{2}25\frac{\ln n}{n}$$

$$\leq \; -\ln n + k/2.\ln(\ln n/k). \tag{5}$$

The function $k \mapsto k/2.\ln(\ln n/k)$ has a maximum for $k = \ln n/e$, and for this value of k it has a value $\ln n/2e$. Thus, for all k in question, the expression (5) tends to $-\infty$ as n tends to ∞.

Hence, the numbers a, b can be chosen in such a way that

$$b/a = O(n^{2/k}\sqrt{\log n/k})$$

and (4) holds, and this proves Theorem 2.2. □

5 Lower bound for distortion

In this section we prove Theorem 2.3; we start by some definitions and lemmas.

A subset N of a metric space (X, ρ) is called an *ε-net*, if

(i) For every $x \in X$ there exists $y \in N$ with $\rho(x, y) < \varepsilon$, and

(ii) Every pair of points in N has distance at least ε.

Let s_m denote the regular m-dimensional simplex (in E^m) with side length 1. By Σ_m we shall denote the m-skeleton (the collection of all at most m-dimensional faces) of s_{2m+2}. By a classical result of R. van Kampen, Σ_m cannot be topologically embedded into E^{2m} (for $m = 1$ we get just one of Kuratowski graphs - K_5). Further let δ_m denote the minimum distance of a pair of disjoint simplices in Σ_m.

Lemma 5.1 *(i) For every $m \geq 1$, $\delta_m > 1/2$.*

(ii) For every fixed m, the following holds: for every $\varepsilon > 0$ there exists an ε-net in Σ_m, consisting of $O(\varepsilon^{-m})$ points.

Proof: (i) Let h_m denote the height of s_m. It is not difficult to compute (by induction) that $h_m = \sqrt{(m+1)/(2m)} > 1/2$ (we omit this elementary calculation here). We will now show that $\delta_m \geq h_{2m+2}$. Let $s, s' \in \Sigma_m$ be two disjoint simplices (sharing no vertex). Since a simplex is a convex hull of its vertices, some of the shortest segments joining s to s' starts in some vertex of s or s', say in a vertex x of s. Now the simplex s' must be completely contained in the facet of the simplex s_{2m+2} not containing x, and this whole facet has distance h_{2m+2} from x.

(ii) It is sufficient to prove the statement for a m-dimensional simplex instead of Σ_m, since Σ_m consists of a constant number of m-dimensional simplices. It suffices to use a standard volume argument: For every $\varepsilon > 0$ the ε-neighborhood of a regular unit simplex in E^{2m+2} has volume $O(\varepsilon^{m+2})$. If X is a 2ε-net in this simplex, the ε-balls centered in points of X are disjoint and contained in the ε-neighborhood of the simplex. Each of these balls has volume $\Omega(\varepsilon^{2m+2})$, and this gives $|X| = O(\varepsilon^{-m})$.□

In the proof of Theorem 2.3 we use a result of T.Ganea [Gan59]. Let X, Y be metric spaces and $\varepsilon > 0$. A continuous mapping $f : X \to Y$ is called an ε-mapping, when $diam(f^{-1}(y)) < \varepsilon$ for every $y \in f(X)$. Ganea proved the following theorem:

Theorem 5.2 ([Gan59], Theorem 2) *For every metric compact polyhedron X of dimension $m \neq 2$ there exists an $\varepsilon = \varepsilon(X)$, such that if X admits a ε-mapping into E^{2m}, then X can be topologically embedded into E^{2m}.* □

Let us remark that a straightforward generalization of this result (for a polyhedron of dimension m, embedded into E^k, $m \leq k \leq 2m$) does not hold; see e.g. [Gan59, MaS66, MaS67].

Ganea's proof implies the following (here the assumption $m \neq 2$ is already unnecessary):

Lemma 5.3 [Gan59] *For $\varepsilon \geq \delta_m$ there exists no ε-mapping of Σ_m into E^{2m}.*□

Proof of Theorem 2.3 : Let us start by a remark on extensions of Lipschitz mappings. It is known that for every n the following holds ([WW75]): When Y is a metric space, $X \subseteq Y$ and $f : X \to \ell_\infty^n$ is a Lipschitz mapping (here ℓ_∞^n denotes the n-dimensional real vector space, equipped with the L_∞-norm), then there exists an extension $\bar{f} : Y \to \ell_\infty^n$ of f with $\|\bar{f}\|_{Lip} = \|f\|_{Lip}$.

Let now f be a mapping from X to E^k, and let $X \subseteq Y$. If we formally identify E^k with ℓ_∞^k, we get that f can be extended to $\bar{f} : Y \to \ell_\infty^k$, with $\|\bar{f}\|_{Lip} \leq \|f\|_{Lip}$ (with respect to the L_∞-norm). If we now consider $\bar{f}$ as a mapping into E^k, we have $\|\bar{f}\|_{Lip} \leq \sqrt{k}\|f\|_{Lip}$ (with respect to the Euclidean norm).

Let n be large enough compared to m. Our metric space X will be chosen as an ε-net in Σ_m, where $\varepsilon = \Omega(n^{-1/m})$, according to Lemma 5.1(ii).

Let us assume that there exists a non-contracting mapping $f : X \to E^{2m}$ with Lipschitz constant D. Let $\bar{f}$ be an extension of f with domain Σ_m with (by the above remark) $\|\bar{f}\|_{Lip} = O(\|f\|_{Lip}) = O(D)$. By Lemma 5.3 the mapping $\bar{f}$ is not a δ_m-mapping, which means that there exist two points $x, y \in \Sigma_m$, $\|x - y\| \geq \delta_m$, with $\bar{f}(x) = \bar{f}(y)$. Let us choose $x', y' \in X$ in such a way that $\|x - x'\| \leq \varepsilon$, $\|y - y'\| \leq \varepsilon$. Then $\|x' - y'\| \geq \delta_m - 2\varepsilon \geq 1/4$, hence (since f is non-contracting) also $\|f(x') - f(y')\| \geq 1/4$. But at the same time

$$\|f(x') - f(y')\| \leq \|f(x') - f(x)\| + \|f(y) - f(y')\| = O(D\varepsilon),$$

so $D = \Omega(1/\varepsilon) = \Omega(n^{1/m})$.

So far our metric space X lies in E^{2m+2}. Let us choose a homeomorphic embedding $h : \Sigma_m \to E^{2m+1}$ with distortion bounded by a constant. The space $X' = h(X)$ has a bounded Lipschitz distance from X, and so it must also be $dist(X', \subseteq E^{2m}) = \Omega(n^{1/m})$.□

Remark: It would be also possible to choose X as the vertices of a triangulation of Σ_m by simplices of diameter $O(\varepsilon)$, and then take $\bar{f}$ as the simplicial extension of f. However, there are some technical difficulties in the construction of the appropriate triangulation, and the above proof is simpler. For small dimensions ($m = 1, 2$) it is easy to find optimal triangulations of Σ_m by regular simplices, and we get a slightly better constants of proportionality.

Let us also remark that for the embeddings into E^1 and E^2, it is possible to give completely elementary proofs (without the rather heavy machinery hidden behind the result of Ganea) using the non-realizability of a circle in the real line or of K_5 in the plane.

References

[AM83] N.Alon, J.Milman: Embeddings of l_∞^k in finite dimensional Banach spaces, Israel J. Math 45(1983) p. 265-280

[Bou85] J.Bourgain: On Lipschitz embedding of finite metric spaces in Hilbert space, Israel J. Math. 52(1985) p. 46-52

[BFM86] J.Bourgain, T.Figiel, V.Milman: On Hilbertian subspaces of finite metric spaces, Israel J. Math. 55(1986) p. 147-152

[BMW86] J.Bourgain, V.Milman, H.Wolfson: On type of metric spaces, Trans. Am. Math. Soc. 294(1986) p. 295-317

[Enf69a] P.Enflo: On a problem of Smironov, Ark. Mat. 8(1969), p. 107-109

[Enf69b] P.Enflo: On the nonexistence of uniform homeomorphisms between L_p-spaces, Ark. Mat. 8(1969), p. 103-105

[FLM77] T.Figel, J.Lindenstrauss, V.Milman: The dimension of almost spherical sections of convex bodies, Acta Math. 59(1977) p. 53-94

[Gan59] T.Ganea: Comment on embedding of polyhedra in Euclidean spaces, Bull. Acad. Polon. Sci. S0r. Sci. Math. Astronom. Phys 7(1959) p. 27-32

[Hei88] W.J.Heiser: Multidimensional scaling with least absolute residuals, in: Classification and related methods of data analysis, H.H. Bock ed., North Holland 1988, p. 455-472

[JL84] W.Johnson, J.Lindenstrauss: Extensions of Lipschitz maps into a Hilbert space, Contemp. Math. 26(1984) p. 189-206

[JLS87] W.B.Johnson, J.Lindenstrauss, G.Schechtman: On Lipschitz embedding of finite metric spaces in low dimensional normed spaces, in:*Geometrical aspects of functional analysis* (J.Lindenstrauss, V.D.Milman eds.), LNM 1267, Springer-Verlag 1987

[Kru64] J.B.Kruskal: Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis, Psychometrika 29(1964a) p. 1-27

[Mat85] R.Mathar: The best Euclidian fit to a given distance matrix in prescribed dimension, Lin. Alg. Appl 67(985) p. 1-6

[Ma89] J.Matoušek: Lipschitz distance of metric spaces, CSc. degree thesis (in Czech), Charles University 1989

[MaS66] S.Mardešić, J.Segal:A note on polyhedra embeddable in the plane, Michig. Math. J. 33(1966) p. 633-638

[MaS67] S.Mardešić, J.Segal: ε-mappings and generalized manifolds, Michig. Math. J. 14(1967) p.171-182

[MiS86] V.D.Milman, G.Schechtman: *Asymptotic theory of finite dimensional normed spaces*, LNM 1200, Springer-Verlag 1986

[Scho38] I.J.Schoenberg: Metric spaces and positive definite functions, Trans. Amer. Math. Soc. 44(1938), p. 522-536

[WW75] J.H.Wells, L.R.Williams: *Embeddings and extensions in analysis*, Springer-Verlag 1975