Open problems on embeddings of finite metric spaces

edited by Jiří Matoušek

Department of Applied Mathematics and
Institute of Theoretical Computer Science (ITI)
Charles University
Malostranské nám. 25, 118 00 Praha 1
Czech Republic
e-mail: matousek@kam.mff.cuni.cz

since June 2010 maintained jointly with Assaf Naor

Last revised: August 2011

Contents

1 Introduction **3**

1.1 General references 3
1.2 Basic definitions 4

2 Problems from 2002 **4**

2.1 Subspaces of ℓ_1 into ℓ_1^n (Gideon Schechtman) 4
2.2 Squared ℓ_2 metrics into ℓ_1 (Nathan Linial) 5
2.3 Girth and ℓ_1/ℓ_2 embeddings (Nathan Linial) 5
2.4 Algorithmic difficulty of ℓ_1-embeddings (Chandra Chekuri; Anupam Gupta) 6
2.5 Planar-graph metrics into ℓ_1 (Nathan Linial) 6
2.6 Decision problem (Yuri Rabinovich) 7
2.7 How large graph? (Jiří Matoušek; probably folklore) 7
2.8 Planar into $\mathbb{R}^2$ (Jiří Matoušek) 7
2.9 Algorithmic complexity (Chandra Chekuri) 7
2.10 Bandwidth and ball growth (Nathan Linial) 8
2.11 Into $\mathbb{R}^3$ (Jiří Matoušek) 8
2.12 Lipschitz mapping of n grid points onto a square (Uriel Feige) 8
2.13 Path into $\mathbb{R}^3$, volume-respecting (Uriel Feige) 9
2.14 Path into ℓ_2 (Nathan Linial) 9
2.15 Levenstein metric into ℓ_1 (Piotr Indyk) 9
2.16 Fréchet metric into ℓ_∞ (Piotr Indyk) 10
2.17 Earth-mover distance (Piotr Indyk) 11
2.18 Convex extensions (Yuval Rabani) 11
2.19 Explicit graphs with high sphericity (Nathan Linial) 12

3 Problems from 2003 **13**

3.1 Expanders and ℓ_∞ (Nathan Linial) 13
3.2 Dimension reduction in ℓ_p with $p \notin \{1, 2, \infty\}$ (James R. Lee) . . 13
3.3 Doubling subsets of ℓ_2 and dimension reduction (James R. Lee) . 14
3.4 Linear dimension reduction (Assaf Naor) 14
3.5 Dimension reduction for cut metrics in some norm (Manor Mendel) 14
3.6 Lipschitz extension problems (James R. Lee and Assaf Naor) . . 14
3.7 Multiembedding into trees (Yair Bartal) 15
3.8 What metrics are planar? (Nathan Linial; Yuri Rabinovich) . . . 16
3.9 Order of congruence of ℓ_p^n (Reminded by Victor Chepoi) 16
3.10 m-simplex inequality (Michel Deza) 16

4 Solved or almost solved problems from 2002 **17**

4.1 Explicit embedding of ℓ_2 into ℓ_1 (Gideon Schechtman, Piotr Indyk; probably folklore) 17
4.2 ℓ_1 metrics into ℓ_2 (Nathan Linial) 17
4.3 Dimension reduction in ℓ_1 (Nathan Linial) 18
4.4 Pentagonal into ℓ_1 (Michel Deza) 18
4.5 Into ℓ_∞^d (Yuri Rabinovich) 18
4.6 Average distance for line embedding (Yuri Rabinovich) 19
4.7 Probabilistic into trees (Yair Bartal) 19
4.8 HST in a metric space (Yair Bartal) 19
4.9 Euclidean subspace in a metric space (Yair Bartal; Nathan Linial) 20
4.10 Assouad's problem (reminded by J. Matoušek) 20
4.11 More examples badly embeddable into ℓ_2 (Nathan Linial) 20
4.12 A variation on a theme by L. Levin (Nathan Linial) 20
4.13 Volume-respecting for large k (Uriel Feige) 21
4.14 Probabilistic into spanning trees (Yair Bartal; Nathan Linial) . . 21

5 Solved or almost solved problems from 2003 **21**

5.1 Realizability of graphs (Robert Connelly and Maria Sloughter) . 21
5.2 $\{0, 1, 2\}$-metrics into ℓ_p (Manor Mendel) 22
5.3 Nonlinear type and cotype (Assaf Naor) 23
5.4 Isometric vs. isomorphic L_2 squared (James R. Lee) 24
5.5 Decomposability of metrics (Kunal Talwar) 25
5.6 Trees into ℓ_1 (Anupam Gupta, James R. Lee, and Kunal Talwar) 26

1 Introduction

Finite metric spaces have emerged in recent years as a new and influential branch of discrete mathematics, with deep and surprising applications in Computer Science. This paper collects open problems presented by the participants of two workshops on discrete metric spaces and their algorithmic applications.

The first workshop was held in Haifa, Israel, in March 2002. The main organizer of the workshop was Yuri Rabinovich (University of Haifa), and the co-organizers were Uriel Feige (Weizmann Institute, Rehovot), Nathan Linial (Hebrew University of Jerusalem), Jiří Matoušek (Charles University, Prague), Ilan Newman (University of Haifa), and Alistair Sinclair (University of California at Berkeley). The financial support of the workshop by the US–Israel Binational Science Foundation (BSF), by the Caesaria Rothschild Foundation Institute, and by the Haifa University Research Counsel is gratefully acknowledged.

The second workshop took place in Princeton in August 2003. The main organizer was Moses Charikar (Princeton University), the co-organizers were Piotr Indyk (MIT), Nathan Linial (Hebrew University of Jerusalem), and Yuri Rabinovich (University of Haifa). We would like to thank DIMACS, NSF, the Institute for Advanced Study, and Princeton University for financial support of the workshop.

The open problems collected here have been presented during the talks at the workshops, during the problem sessions, or contributed later by the participants. The heading of each problem includes a short title (mostly assigned by the editor, in order to facilitate a quick orientation) and the name of the person presenting the problem. This is not necessarily the original author of the question; some of the problems seem to be folklore and it may be nontrivial to trace their origins.

Remarkably, about one fourth of the problem collection from Haifa were solved until the next workshop, in approximately one year. More problems have been solved or almost solved later. As a source of open problems, this collection has been badly damaged by this development, but on the other hand, it has become a document of an amazing progress in the area.

For historical interest and to complete the picture, the solved problems are included as well, in separate sections. The remaining problems are, to the best knowledge of the editor, still open at the time of the last revision.

1.1 General references

- P. Indyk: Algorithmic applications of low-distortion embeddings. In *Proc. 42nd IEEE Symposium on Foundations of Computer Science*, 2001, pages 10–33.
  (A survey with emphasis on algorithmic applications.)

- J. Matoušek: *Lectures on Discrete Geometry*, Chapter 15, Springer, New York, NY, 2002.
  (Proofs of several basic results and a survey; an updated version of the

chapter, reflecting some of the developments up until 2005, is available on the author's web page.)

- M. M. Deza and M. Laurent: *Geometry of Cuts and Metrics*, Springer, Berlin, 1997.
  (Isometric embeddings, mainly into ℓ_1.)

- P. Indyk and J. Matoušek: Low-distortion embeddings of finite metric spaces, Chapter 8 of *CRC Handbook of Discrete and Computational geometry*, 2nd edition, CRC Press, LLC, Boca Raton, FL, pages 177–196, 2004; also available on Matoušek's web page.

1.2 Basic definitions

Let (X, d_X) and (Y, d_Y) be metric spaces. The *Lipschitz norm* of a mapping $f\colon X \to Y$ is

$$\|f\|_{\mathrm{Lip}} = \sup_{x,y\in X, x\neq y} \frac{d_Y(f(x), f(y))}{d_X(x,y)}.$$

The *distortion* of such f is $\|f\|_{\mathrm{Lip}}\|f^{-1}\|_{\mathrm{Lip}}$. The notation $c_Y(X)$ stands for the infimum of distortions of all mappings $f\colon X \to Y$. Sometimes $c_p(X)$ abbreviates $c_{\ell_p}(X)$.

Let $\mathcal{G}$ be a class of (finite) graphs. Each graph $G \in \mathcal{G}$ with nonnegative real weights on edges defines the shortest-path metric on its vertex set. A metric space is called a *$\mathcal{G}$-metric* if it is isometric to a subspace of such a metric space for some $G \in \mathcal{G}$. For $\mathcal{G}$ being the class of all planar graphs, we speak of *planar metrics*, or perhaps more descriptively, of *planar-graph metrics*.

2 Problems from 2002

2.1 Subspaces of ℓ_1 into ℓ_1^n (Gideon Schechtman)

Given m points $y_i \neq 0$ in $\mathbb{R}^k$, let

$$K = \sum_{i=1}^{m} [-y_i, y_i] = \Big\{ a_1 + a_2 + \cdots + a_m : a_i \in [-y_i, y_i] \Big\}.$$

Problem: Does there exist a universal constant C such that given $x_1, \ldots, x_k \in K$, there are signs $\varepsilon_1, \ldots, \varepsilon_k \in \{-1, +1\}$ with

$$\sum_{i=1}^{k} \varepsilon_i x_i \in C\sqrt{k} \cdot K?$$

Known:

1. There exist $\varepsilon_1, \ldots, \varepsilon_k$ with $\sum_{i=1}^{k} \varepsilon_i x_i \in C\sqrt{k \log\log k} \cdot K$.

2. A negative solution implies

$$\lim_{k\to\infty} \frac{N_1(k,2)}{k} = \infty,$$

where

$$N_1(k,2) \quad = \sup_{\substack{X \text{ is a } k-\text{dim.} \\ \text{subspace of } L_1}} \inf\{n : X \text{ can be 2-embedded into } \ell_1^n\}.$$

2.2 Squared ℓ_2 metrics into ℓ_1 (Nathan Linial)

Can the cut cone be approximated well by a cone for which we have efficient membership/separation oracles? Specifically, how far is it from the cone of squared ℓ_2-metrics? In particular, is there a constant C such that every squared ℓ_2-metric embeds into ℓ_1 with distortion at most C? A metric space (X, ρ) is a *squared ℓ_2-metric* (also called a *metric of negative type*) if there is a map $f\colon X \to \ell_2$ such that $\rho(x,y) = \|x-y\|^2$ for all $x, y \in X$.

Progress report (January 2005): Khot and Vishnoi [KV05] gave a lower bound and proved that there is no constant C as above. A slightly better quantitative bound for the distortion in their example (of order $\log\log n$ for an n-point space) was given by Krauthgamer and Rabani [KR06]. On the other hand, Arora, Lee, and Naor [ALN05], improving on an earlier result by Chawla, Gupta, and Räcke [CGR05], proved that every n-point metric of negative type embeds into ℓ_2, and thus also in ℓ_1, with distortion $O(\sqrt{\log n}\,\log\log n)$.

Update (July 2010): Lee and Naor [LN06], using a result of Cheeger and Kleiner [CK06a] on a certain kind of weak differentiability of maps from the Heisenberg group to L_1, gave another example of squared ℓ_2-metrics (even a doubling one) whose embedding in L_1 requires arbitrarily large distortion. Cheeger, Kleiner and Naor [CKN09] showed that this example yields arbitrarily large n-point metric spaces whose ℓ_1 distortion is $(\log n)^{\Omega(1)}$.

2.3 Girth and ℓ_1/ℓ_2 embeddings (Nathan Linial)

In this question all vertices in the graph G have all degrees at least 3 (actually, it suffices to have the average degree bounded away from 2; see [BLMN03]). Linial, London, and Rabinovich [LLR95] conjectured that $c_2(G) \geq \Omega(\text{girth}(G))$. Linial, Magen, and Naor [LMN02] proved $c_2(G) \geq \Omega(\sqrt{\text{girth}(G)}\,)$ for regular graphs. What is the true bound? Also, does $c_1(G)$ tend to infinity with girth(G)? The c_2 analogue of this weaker version is easy, since c_2 is unbounded for trees of large depth, but this is obviously not the case for c_1.

Solution of the second question: In August 2011 Ostrovskii [Ost11] constructed regular graphs G of constant degree and arbitrarily large girth with $c_1(G)$ bounded by a constant.

2.4 Algorithmic difficulty of ℓ_1-embeddings (Chandra Chekuri; Anupam Gupta)

Can the minimum distortion for embedding of a given finite metric space into ℓ_1 be approximated within a constant factor in polynomial time? (Testing isometric ℓ_1-embeddability is known to be NP-hard.)

What about embedding into $\mathbb{R}^1$ (or $\mathbb{R}^d$), even for a tree metric?

Update (Anupam Gupta): The problem of embedding into $\mathbb{R}^1$ is known to be NP-hard even for spiders (i.e., homeomorphs of stars). This can be shown by an easy reduction from knapsack, as was observed by Gupta and Nayak. The problem is also max-SNP hard for general graphs.

Update (July 2010): Bădoiu, Chuzhoy, Indyk, and Sidiropoulos [BCIS05] proved the following inapproximability result: it is NP-hard to distinguish n-point metric spaces that can be embedded in $\mathbb{R}^1$ with distortion $O(n^{4/12})$ from those that require distortion at least (roughly) $\Omega(n^{5/12})$. Matoušek and Sidiropoulos [MS08] proved an analogous results (with different exponents) for embeddings in $\mathbb{R}^2$, while for embeddings in $\mathbb{R}^d$, for every fixed d fixed, they showed a stronger result: it is NP-hard to distinguish n-point spaces that can be embedded in $\mathbb{R}^d$ with distortion at most C_d from those that require distortion at least $\Omega(n^{c/d})$; here c is an absolute positive constant and C_d depends only on d. For embeddings in $\mathbb{R}^2$, Edmonds, Sidiropoulos, and Zouzias [ESZ10] proved hardness of distinuishing 3-embeddable spaces from 3.49999-embeddable ones. Positive results (polynomial-time approximation algorithms) are known for some special classes of metric spaces—we refer to [MS08], [ESZ10] for citations.

2.5 Planar-graph metrics into ℓ_1 (Nathan Linial)

Is there a constant C such that the metric of every finite weighted planar graph embeds in ℓ_1 with distortion at most C? Even more generally, does the same hold for every nontrivial minor-closed family of graphs? Gupta, Newman, Rabinovich, and Sinclair [GNRS99] proved this for the class of graphs with no K_4-minor, as well as for the class of graphs with no $K_{2,3}$ minor.

Progress report (January 2003): Chekuri, Gupta, Newman, Rabinovich, and Sinclair [CGNRS03] proved this also for the class of k-*outerplanar graphs*. Roughly speaking, a planar graph is k-outerplanar if it has no $k+1$ disjoint nested cycles; see [CGNRS03] for a precise definition.

Update (July 2010): Chakrabarti, Jaffe, Lee and Vincent [CJLV08] proved that every graph that excludes a K_4-minor embeds into L_1 with distortion at most 2. This bound on the L_1 distortion of K_4-minor-free graphs is sharp, as shown by Lee and Raghavendra [LR07]. It is also shown by Chakrabarti, Jaffe, Lee and Vincent [CJLV08] that every graph that excludes the 4-wheel as a minor embeds into L_1 with $O(1)$ distortion. Lee and Sidiropoulos [LS09] showed that every minor-closed family $\mathcal{F}$ that does not contain every possible tree satisfies $\sup_{X\in\mathcal{F}} c_1(X) < \infty$. They also showed that if every planar graph emebds into L_1 with $O(1)$ distortion, and in addition a certain conjecture which they call "the k-sum conjecture" holds, then any forbidden minor family embeds into L_1 with distortion $O(1)$.

2.6 Decision problem (Yuri Rabinovich)

Given a metric, can one decide in polynomial time whether it is a planar metric?

2.7 How large graph? (Jiří Matoušek; probably folklore)

Let $\mathcal{G}$ be some class of graphs, say all planar graphs. What is the smallest $N = N(\mathcal{G}, n)$ such that every n-point $\mathcal{G}$-metric can be represented as a subspace of the metric of a graph $G \in \mathcal{G}$ with at most N vertices? (For $\mathcal{G}$ = all trees, $O(n)$ suffices.) What if we do not insist on an exact representation, but allow for some constant distortion, say?

2.8 Planar into $\mathbb{R}^2$ (Jiří Matoušek)

Is there any planar-graph metric on n vertices requiring distortion more than $O(\sqrt{n}\,)$ for embedding into $\mathbb{R}^2$? It is known, and not difficult to prove, that if every edge of a fixed non-planar graph is replaced by a path of length n, the resulting graph requires distortion $\Omega(n)$. On the other hand, every tree with unit-length edges can be embedded with $O(\sqrt{n}\,)$ distortion [BMMV02] and every n-point metric space can be $O(n)$-embedded even into $\mathbb{R}^1$ [Mat90].

Progress report. Bateni, Hajiaghayi, Demaine, and Moharrami [BHDM07] generalized the $O(\sqrt{n}\,)$ upper bound to outerplanar graphs with unit-length edges, and they found n-point planar graph metrics requiring $\Omega(n^{2/3})$ distortion for embedding into $\mathbb{R}^2$.

2.9 Algorithmic complexity (Chandra Chekuri)

Given a weighted graph G (or equivalently, a metric on $V(G)$), what is the complexity of finding the optimal embedding of G into a convex combination of dominating trees? We know that an $O(\log n)$ absolute distortion is achievable [FRT03] but this is true for all metrics on n vertices. We are interested in the optimal embedding for G. Can we obtain an $O(1)$ approximation for the optimal embedding? In Charikar et al. [CCGGP98] it is shown that this is related to the following optimization problem. Given a graph G and a weight function $w\colon V \times V \to V$, find a dominating tree metric for G such that the quantity $\sum_{u,v \in V(G)} w(u,v) d_T(u,v)$ is minimized. An approximation of α for this later problem implies an α approximation for the former problem and vice versa.

Comments and updates (Anupam Gupta): The problem of finding the best distortion (single) spanning tree for a given metric is known as the *tree spanner* problem, and it is known to be NP-hard [CC95]. The minimum spanning tree gives a trivial $n-1$ upper bound, and the presence of an isometric cycle of girth g in the given graph implies an $\Omega(g)$ lower bound. A result of

Emek and Peleg [EP04] gives an $O(\log n)$ approximation to find the best tree spanner for *unweighted* graphs.

2.10 Bandwidth and ball growth (Nathan Linial)

An easy lower bound on the bandwidth of any graph G is

$$\mathrm{bw}(G) \geq \Omega(\beta(G))$$

where

$$\beta(G) := \max_{x,r} \frac{|B_r(x)|}{r}.$$

Feige [Fei00] showed that these two parameters differ by a factor that is at most polylogarithmic in n. For expanders clearly a logarithmic gap exists. So we ask: Is it true that

$$\mathrm{bw}(G) \leq O(\beta(G) \log n)$$

always holds? This doesn't seem to be known even for trees.

2.11 Into $\mathbb{R}^3$ (Jiří Matoušek)

What is the worst-case distortion necessary to embed an n-point metric space into $\mathbb{R}^3$?

There is an $\Omega(\sqrt{n}\,)$ lower bound and an $O(n^{2/3} \log^{3/2} n)$ upper bound; see [Mat90]. Similarly, there is a significant gap between known lower and upper bounds for the worst-case distortion into $\mathbb{R}^{2k+1}$, $k \geq 1$ (while for even dimensions the upper and lower bounds match up to a logarithmic factor).

2.12 Lipschitz mapping of n grid points onto a square (Uriel Feige)

Consider the infinite integer two-dimensional lattice $\mathbb{Z}^2$ and number n distinct points in it from 1 to n. This induces an n point finite metric space (S, d) with ℓ_1 distances defined in a natural way. For what value of $c > 1$ (possibly depending on n) is it always possible to number the $\sqrt{n}$ by $\sqrt{n}$ grid G from 1 to n while ensuring that for every i and j, their ℓ_1 distance in G is at most c times their distance in S? (We allow here distances to shrink arbitrarily.)

This problem was motivated by generalizations of the bandwidth problem to two dimensions, with (pseudo) applications in VLSI layout.

Comment (J. Matoušek): This reminds of a beautiful (but not necessarily closely related) problem of Laczkovich in geometric measure theory: Does there exist, for every set $E \subseteq \mathbb{R}^d$ of positive Lebesgue measure, a Lipschitz map $f\colon \mathbb{R}^d \to \mathbb{R}^d$ with $[0,1]^d \subseteq f(E)$? A positive answer is known for $d = 2$, due to Preiss, while the question is wide open for all $d \geq 3$. A "discrete" version might perhaps go as follows: Does there exist a constant $C = C(d)$ such that every Cn^d-point set in $\mathbb{Z}^d$ can be mapped *onto* the "cube" $\{1, 2, \ldots, n\}^d$ by a C-Lipschitz map? For $d = 2$, this is a weakening of Feige's problem above, and it is known to hold, but nothing is known for $d = 3$.

2.13 Path into $\mathbb{R}^3$, volume-respecting (Uriel Feige)

Can the path P_n on n vertices be embedded via a contraction in $\mathbb{R}^3$ with only polylogarithmic distortion in the area of triangles? That is, for points $i < j < k$ we want the area of the triangle formed by the images of these points to be at least $(j-i)(k-j)/(\text{polylog } n)$. For embeddings in $\mathbb{R}^2$ the answer is negative, and for embeddings in $\mathbb{R}^{O(\log n)}$ the answer is positive.

Comment (July 2010): Some bounds for this problem can be found in Dehornoy [Deh08], although they are far from answering the question.

2.14 Path into ℓ_2 (Nathan Linial)

Feige [Fei00] introduced an interesting notion of volume to the field of discrete metric spaces, but other reasonable definition can be considered as well. Here is a problem illustrating one such alternative definition: Is there a map $\varphi\colon \{1,\ldots,n\} \to \ell_2$ such that

$$\text{area}(\varphi(i),\varphi(j),\varphi(k)) = \Theta((j-i)(k-j))$$

for every i,j,k, $n \geq k > j > i \geq 1$? (Unlike in Feige's definition, φ is *not* required to be non-expanding.)

2.15 Levenstein metric into ℓ_1 (Piotr Indyk)

Let Σ be an "alphabet" set, and let Σ^* denote all strings with symbols from Σ. Let $G = (\Sigma^*, E)$ be an infinite undirected graph, such that $\{s,s'\} \in E$ iff s' can be obtained from s by inserting one symbol into s, deleting one symbol from s, or substituting one symbol in s by another symbol. Let M_Σ (*the Levenstein metric*) be the shortest path metric of G, and let $M_\Sigma^{\leq d}$ denote M_Σ restricted to the set of strings of length at most d.

Problem: Find the smallest value of $c = c(d)$ such that $M_\Sigma^{\leq d}$ can be embedded into ℓ_1 with distortion c.

The ℓ_1 norm seems to be the best candidate for the "host" norm. This is due to the fact that $M_\Sigma^{\leq d}$ "contains" the Hamming metric $M_H^d = (\{0,1\}^d, D_H)$. Specifically, there exist an isometry from M_H into $M_\Sigma^{\leq d}$; e.g., for $\Sigma = \{1,\ldots,2d\}$ we can use the mapping f that for any argument $x \in \{0,1\}^d$ substitutes x_i by $2i - x_i$, for all $i = 1,\ldots,d$. Unbounded Σ is not crucial—one can easily obtain a constant distortion embedding of M_H into $M_{\{0,1\}}^{\leq d}$ as well.

However, a low-distortion embedding of $D_\Sigma^{\leq d}$ into $\ell_p^{d'}$ for $p \neq 1$, with $d' = 2^{o(d)}$ (ideally, $d' = d^{O(1)}$) would also be interesting.

It is known that if, in addition to insertions, deletions and substitutions, one allows certain additional operations, then the resulting metric can be embedded into ℓ_1 with distortion $O(\log d \log^* d)$. See Cormode and Muthukrishnan [CM02] for the exact statement of the result (in particular, for the exact definition of the metric).

A major implication of a low-distortion embedding would be an approximate nearest neighbor data structure for the Levenstein metric using small (polynomial) space. For this application, it is *not* crucial for the embedding to be deterministic. In particular, a randomized embedding with very low probability of contraction but only constant probability of expansion is sufficient (see the talk by Piotr Indyk). The only non-trivial result for this nearest neighbor problem is an $O(1)$-approximate data structure, using polylogarithmic query time and n^{d^ε} space, for any constant $\varepsilon > 0$ [Ind04]. The latter result does not use embeddings.

Progress report: Andoni, Deza, Gupta, Indyk, and Raskhodnikova [ADGIR03] showed a lower bound of $\frac{3}{2}$ for embedding the Levenstein metric into ℓ_1, as well as some other results related to this problem. Then Khot and Naor [KN05] proved the lower bound $c(d) = \Omega(\sqrt{\log d / \log\log d}\,)$. This was further improved to $\Omega(\log d)$ by Krauthgamer and Rabani [KR06]. Ostrovsky and Rabani [OR05] proved an upper bound of $2^{O(\sqrt{\log d \log\log d}\,)}$. Charikar and Krauthgamer [CK06] found an $O(\log d)$ embedding in L_1 for the metric space consisting of all permutations of d symbols with the Levenstein metric.

2.16 Fréchet metric into ℓ_∞ (Piotr Indyk)

Let C^d be the set of all polygonal chains in $\mathbb{R}^2$ consisting of d pieces. We represent each curve in C^d as a function $f\colon [0,1] \to \mathbb{R}^2$. For any two curves f, g, define

$$D_F(f,g) = \inf_{\pi:[0,1]\to[0,1]} \sup_{t\in[0,1]} \|f(\pi(t)) - g(t)\|_2$$

where π is continuous, monotone increasing, $\pi(0) = 0$, and $\pi(1) = 1$. (This is called the *Fréchet metric* or *dogkeeper's metric.*)

Problem I: Find the smallest $c = c(d)$ such that (C^d, D_F) can be embedded into $\ell_\infty^{d'}$ for a finite d' with distortion c.

Comments: Note that the set C^d is infinite. Thus the universality of ℓ_∞ results in an embedding into an infinite-dimensional space. The ℓ_∞ norm seems to be the best candidate for the host norm. This is because for any bounded set $S \subset \ell_\infty^d$, the metric (S, ℓ_∞) can be isometrically embedded into (C^{3d}, D_F) (easy to see).

For the practical purposes, we would like d' to be small (ideally, polynomial in d). Possibly, this could be helped by assuming that the curves are discrete. That is, one can consider the space C_I^d containing all polygonal chains with d segments, such that the endpoints of all segments are from $\{0, \ldots, I\}^2$. This leads to a discrete version of Problem I.

Discrete Problem I: Find the smallest $c = c(d, I)$ such that (C_I^d, D_F) can be embedded into $\ell_\infty^{d'}$ (for $d' = (d + \log I)^c$) with distortion c. (Of course, different (but small) c's bounding the distortion and the dimension are also interesting).

Comments: As in the case of the Levenstein metric, the approximate nearest neighbor problem is a major application of any embedding result for D_F. The only algorithm known for the latter problem has bounds similar to those for the Levenstein distance [Ind02].

The only previously known result with a similar flavor is the embedding of the *Hausdorff* metric over bounded subsets of ℓ_p^2, with constant distortion, into low-dimensional ℓ_∞ norm [FI99].

2.17 Earth-mover distance (Piotr Indyk)

Consider the metric M_E^d, the *earth-mover distance*, defined over d-subsets of $\mathbb{R}^2$. For any two sets A, B, the distance $D_E(A, B)$ is defined as the minimum-weight bipartite matching between A and B; i.e.,

$$\min_{\pi: A \to B, \pi \text{ is one-to-one}} \sum_{a \in A} \|a - \pi(a)\|_2$$

Problem: Find the smallest $c = c(d)$ such that M_E^d can be embedded into ℓ_1 with distortion c.

Comments: If the metric is defined over subsets of $\mathbb{R}$, then it can be isometrically embedded into ℓ_1 (and vice versa). This uses the fact that the optimal matching is non-crossing (see the talk by Chandra Chekuri). Moreover, the EMD over s-element subsets of $(\{1 \ldots \Delta\}^k, \ell_p)$ can be embedded into ℓ_1 with distortion $O(k \log \Delta)$; this follows from Charikar's paper [Cha02].

A variation of the above metric (with sum replaced by max) is also of interest; no non-trivial result for this case is known.

Progress report. Khot and Naor [KN05] proved a lower bound of $\Omega(d)$ for the distortion of an embedding of the space of all probability measures on $\{0, 1\}^d$ with the earth-mover (transportation) metric.

Naor and Schechtman [NS06] proved that the space of all probability measures on $\{0, 1, \ldots, n\}^2 \subset \mathbb{R}^2$ with the earth-mover metric requires $\Omega(\sqrt{\log n}\,)$ distortion for embedding in L_1.

2.18 Convex extensions (Yuval Rabani)

Let X be a set, $|X| = n$, and let $Y \subset X$, $|Y| = k$. Let d be a metric on Y.

A semi-metric space (X, δ) is a *metric extension* of (Y, d) if $\delta(x, y) = d(x, y)$ for all $x, y \in Y$. It is a 0-*extension* if for every $x \in X \setminus Y$ there is a $y \in Y$ with $\delta(x, y) = 0$. It is a *convex extension* if there is a mapping $\varphi\colon X \to \Delta^{k-1}$ (where Δ^{k-1} denotes a $(k-1)$-dimensional simplex) such that Y is mapped bijectively onto the vertices of Δ^{k-1} and $d(x, y)$ is the earth-mover distance between $\varphi(x)$ and $\varphi(y)$. Here each point $p \in \Delta^{k-1}$ can be uniquely written as a convex combination of the vertices of Δ^{k-1}, and we interpret it as a probability distribution $(p_i : i \in Y)$ over Y, where p_i is the coefficient of the vertex $\varphi(i)$ in that convex combination. In the particular case when $d(i, j) = 1$ for all $i, j \in Y$, we simply require $\delta(x, y) = \frac{1}{2}\|\varphi(x) - \varphi(y)\|_1$ for all $x, y \in X$. For (Y, d) arbitrary, $\delta(x, y)$ should be the minimum cost of converting the distribution $\varphi(x)$ to the distribution $\varphi(y)$ by moving mass between points, where a move between i and j costs $d(i, j)$ per unit of moved mass; that is,

$$\delta(x, y) = \min\left\{ \sum_{i,j \in Y} \varepsilon_{ij} d(i, j) : \varphi(y)_i = \varphi(x)_i - \sum_j \varepsilon_{ij} + \sum_j \varepsilon_{ji} \text{ for all } i,\ \varepsilon_{ij} \geq 0 \right\}.$$

Questions: Embed metric or convex extensions into convex combinations of 0-extensions. What is the best distortion? The minimum expansion in the worst case? For:

- *d uniform (all distances are 1), δ a convex extension.*
  Known: There is a tight bound of 12/11 on the minimum expansion for $k = 3$. No other tight bounds are known. Asymptotically in k, the bound is between 8/7 and 1.3438. There are better bounds for small values of k.

- *d arbitrary, δ a convex extension.*
  Known: There is a lower bound of $\Omega(\sqrt{\log k})$ and an upper bound of $O(\log k)$ on the minimum expansion. The upper bound is also the best known for convex extensions. (Notice that in the case of metric extensions, the question of distortion does not make sense, because it is easy to construct extensions that require arbitrarily large contraction.)
  Conjecture: For convex extensions, the minimum expansion is a constant. (The best known lower bound is that for uniform d.)

- *d arbitrary, δ a metric extension.*
  Known: There is an upper bound of $O(\log k \log\log k)$. The best lower bound is that for the uniform case. ‘

Embeddings of convex extensions where $x \in X \setminus Y$ can be mapped only to $y \in Y$ with $\varphi(x)_y > 0$ are also interesting.

Comment (C. Chekuri): Yuval Rabani's problem is essentially asking for the integrality gap of an LP relaxation in the paper [CKNZ01] for the 0-extension problem. There are several open questions related to the more general metric labeling problem in the slides of my talk on this at the workshop.

Progress report: In Fakcharoenphol, Harrelson, Rao, and Talwar [FHRT03] the upper bound for the minimum expansion for embedding a metric or a convex extension into a convex combination of 0-extensions has been improved slightly from $O(\log k)$ to $O(\log k/ \log\log k)$. Chuzhoy and Naor proved for arbitrary d and δ a metric extension, the gap can be as much as $\sqrt{\log k}$.

Several new results related to this problem, and in particular, an $\Omega((\log n)^{1/4-\varepsilon})$ inapproximability bound for the algorithmic 0-extension problem, were proved by Karloff, Khot, Mehta, and Rabani [KKMR06].

2.19 Explicit graphs with high sphericity (Nathan Linial)

An embedding $\varphi\colon V(G) \to \ell_2^d$ is called a *proximity map* for G if

$$\|\varphi(x) - \varphi(y)\| \le 1 \text{ if and only if } x \text{ and } y \text{ are adjacent.}$$

(The smallest d such that G has a proximity map into ℓ_2^d is also called the *sphericity* of G in the literature.) It is known that every graph has a proximity map in $d = n-1$ dimensions, but most graphs do not have such a map unless $d \ge \Omega(n)$. In particular, the complete bipartite graph $K_{n,n}$ requires $d \ge n$; see Reiterman, Rödl, and Šiňajová [RRS89] for references and related results. Can you find other explicit families of graphs that require $d = \Omega(n)$?

3 Problems from 2003

3.1 Expanders and ℓ_∞ (Nathan Linial)

In simple terms I am asking whether there is a nontrivial way in which expanders can be embedded in a normed space. A daring version of the question is: Let X be an infinitedimensional Banach space. Suppose there is a family of constant-degree expanders G_k on n_k vertices, $n_k \to \infty$ as $k \to \infty$, that embed in X with distortion $o(\log n_k)$. Does this imply that X contains copies of ℓ_∞^d with distortion $1+\varepsilon$ for every $\varepsilon > 0$? (By a theorem of Maurey this entails that X has no cotype.)

Of course, if this is not the case, it is interesting to understand what happens if the distortion is smaller. Likewise, suppose that X contains with positive probability graphs in the space $G(n, 1/2)$ with distortion bounded away from 2. Does this imply the same conclusion for X?

An even more daring version of the question would be: There is an explicit family of graphs of diameter 2 (e.g., Paley graphs), such if X hosts them with distortion bounded away from 2, then X has no cotype.

Comments. At the meeting, I. Newman and others have raised the question of whether large expanders are "metrically universal", that is, necessarily contain (near-isometrically) an epsilon-net of the unit ball in a high-dimensional ℓ_∞^d. As pointed out by J. Matoušek, this is not true—consider expanders whose diameter is at most constant multiple of the girth (that is, with logarithmically large girth; note that the diameter of an expander is always logarithmic).

As pointed out by A. Naor, phenomena in the spirit alluded to in the question are already known. If X is infinitedimensional, then the distortion of a complete binary tree of depth d is either at most $1+\varepsilon$ for every $\varepsilon > 0$ or it is at least $(\log d)^\alpha$ for some $\alpha \in (0, \frac{1}{2})$. (Apparently this can be read out from Bourgain's proof and other stuff.) Likewise, the distortion of a d-dimensional cube must be either at most $1+\varepsilon$ or at least d^α. This should follow from [BMW86].

3.2 Dimension reduction in ℓ_p with $p \notin \{1, 2, \infty\}$ (James R. Lee)

Consider the smallest d for which every n-point subset of L_p can be $O(1)$-embedded into ℓ_p^d. For $p \in \{1, \infty\}$, $d = n^{\Theta(1)}$, while for $p = 2$, we have $d = \Theta(\log n)$. (The case $p = 1$ is due to [BC03], $p = 2$ is from [JL84], and $p = \infty$ is from [Mat96]).

Somewhat surprisingly, none of these cases seem to shed light on the problem for other values of p. The geometric technique of Lee and Naor [LN04], originally conceived for the case $p = 1$, extends to prove lower bounds for ℓ_∞ (via either expanders or cubes), but the natural extension of the argument to ℓ_p for $p \in (1, 2)$ fails (see [LMN05]). Finally, note that for $p \neq 2$, a simple argument rules out any non-trivial *linear* dimension reduction; see the [LMN05].

Question: Does there exist $p \neq 2$ with $d(n) = n^{o(1)}$?

3.3 Doubling subsets of ℓ_2 and dimension reduction (James R. Lee)

The well-known Johnson–Lindenstrauss flattening lemma tells us that every n-point subset of ℓ_2 can be $O(1)$-embedded into ℓ_2^d with $d = O(\log n)$, and up to a constant factor, this bound is tight by simple volume arguments (consider any orthonormal basis). Can better bounds be achieved when volume arguments do not apply?

A metric space (X, d) is called *doubling* if there exists a constant λ such that every ball in X can be covered by λ balls of half the radius. More specifically, can every subset $X \subseteq \ell_2$ be D-embedded into ℓ_2^k, where k and D depend only on λ? It would be natural to conjecture that $k = O(\log \lambda)$ suffices for sufficiently large constant D.

This question was asked by Lang and Plaut [LP01], and independently by Gupta, Krauthgamer, and Lee [GKL03].

3.4 Linear dimension reduction (Assaf Naor)

We know amazingly little about linear dimension reduction beyond the Johnson-Lindenstrauss lemma. For $1 \leq p \leq \infty$ there are arbitrarily large n-point subsets A_n of L_p such that any linear mapping $T\colon A_n \to \ell_p^d$ incurs distortion $D \geq \Omega[(n/d)^{|1/p-1/2|}]$ [LMN05]. Is this the best possible?

We do not even know if it is possible to linearly embed every n-point subset of L_p in $\ell_p^{n/4}$ (say) with $O(1)$ distortion. In general, is it true that any n-points in a Banach space X can be linearly embedded in a $n/4$ dimensional subspace of X with $O(1)$ distortion?

3.5 Dimension reduction for cut metrics in some norm (Manor Mendel)

Given a constant D, what is the smallest $d = d(D, n)$ such that for any n-point cut metric M, there exists a d-dimensional normed space $X = X(M)$, satisfying $c_X(M) \leq D$? Here a *cut metric* on X is a pseudometric induced by a partition of X into two disjoint sets A, B; points in the same set have distance 0, while every point of A has distance 1 to every point of B.

Remarks:

1. The Newman–Rabinovich diamond graphs (in fact, any planar graph) are $O(1)$ embeddable in $\ell_\infty^{O(\log n)}$ [KLMN03].
2. By John's theorem, if $d(D, n)$ were at most $O(\log n)$ for some constant D, it would imply that n-point cut metrics are $O(\sqrt{\log n})$ embeddable in ℓ_2; also see Problem 4.2.

3.6 Lipschitz extension problems (James R. Lee and Assaf Naor)

For metric spaces $X \subseteq Y$ and Z, let $f\colon X \to Z$ be a Lipschitz function. The Lipschitz extension problem asks, When can f be extended to a Lipschitz function

$\tilde{f}: Y \to Z$ on all of Y? Let us define

$$e(X,Y,Z) = \sup_{f:X\to Z} \inf \Big\{ K : f \text{ can be extended to a function } \tilde{f}: Y \to Z$$
$$\text{with } \|\tilde{f}\|_{\mathrm{Lip}} \le K \, \|f\|_{\mathrm{Lip}} \Big\}.$$

Finally, let us write $e_n(Y,Z) = \sup\{e(X,Y,Z) : X \text{ is an } n\text{-point subset of } Y\}$. For more background to the following problems and results see [LN04a].

A classical result of Kirszbraun shows that $e_n(L_2, L_2) = 1$, and the seminal paper [Bal92] extends Kirszbraun's theorem to show that for any fixed $p \in (1,2)$, $e_n(L_2, L_p) \le \frac{6}{\sqrt{p-1}}$. Ball poses the intriguing open problem:

Question 1. Can $e_n(L_2, L_1)$ also be bounded by some constant independent of n?

Perhaps this community's familiarity with finite subsets of L_1 can help to answer the question negatively (as in the case of dimension reduction). One can also ask:

Question 2. For $1 < q < 2 < p < \infty$, is $e_n(L_p, L_q)$ bounded?

Answer (September 2004): The positive solution to Ball's Markov type 2 problem (see Problem 5.3) implies boundedness here [NPSS04].

While $e_n(L_2, L_2)$ and $e_n(L_2, L_\infty)$ both equal 1, there is a lower bound $e_n(L_2, L_p) = \Omega((\log n / \log\log n)^{(p-2)/p^2})$ for every fixed p, $2 < p < \infty$.

Question 3. What is the correct order of magnitude of $e_n(L_2, L_p)$, $2 < p < \infty$?

For an arbitrary metric space X and Banach space Z, the best known general upper bound is $e_n(X,Z) = O(\log n / \log\log n)$. But all lower bounds so far fail to beat the $O(\sqrt{\log n})$ barrier (Johnson and Lindenstrauss proved $e_n(L_1, L_2) = \Omega(\sqrt{\log n / \log\log n}))$.

Question 4. What is the largest possible value of $e_n(X,Z)$, for an arbitrary metric space X and an arbitrary Banach space Z?

3.7 Multiembedding into trees (Yair Bartal)

A *multiembedding* of a metric space (M, d_M) in a metric space (N, d_N) is a surjective mapping $f: N \to M$ (this is not a typo!). That is, each $x \in M$ is embedded into a set of "representatives" $f^{-1}(x)$. We also require a noncontraction property: for all $x, y \in N$ we have $d_N(x,y) \ge d_M(f(x), f(y))$.

The *path distortion* of a multiembedding f of M in N is the infimum of numbers α such that for every sequence $p = (v_1, v_2, \ldots, v_m)$ of points $v_i \in M$ there exists a sequence $p' = (v'_1, v'_2, \ldots, v'_m)$ of points $v'_i \in N$ with $f(v'_i) = v_i$ such that $\ell(p') \le \alpha \cdot \ell(p)$, where $\ell(p) = \sum_{i=1}^{m-1} d_M(v_i, v_{i+1})$, and similarly for $\ell(p')$.

Question 1: What is the smallest $\alpha = \alpha(n)$ such that any n-point metric space M has a multiembedding into a tree metric N with path distortion at most α, where $|N| \le n^{O(1)}$? In particular, is $\alpha(n) = o(\log n)$?

Question 2: What about other target metrics, such as ℓ_2?

Remarks:

1. This type of embedding suffices to reduce metrical task systems and group Steiner problems to the same problems but on trees.

2. Probabilistic embedding bounds imply multiembedding bounds, and therefore, $\alpha(n) = O(\log n)$.

3.8 What metrics are planar? (Nathan Linial; Yuri Rabinovich)

For which graphs G is the shortest-path metric (with unit-weight edges, say) planar, that is, given by the metric of some planar graph, possibly with weighted edges, on a superset of the vertex set of G? For example, the metric of K_5 is planar, while that of $K_{3,3}$ is not.

3.9 Order of congruence of ℓ_p^n (Reminded by Victor Chepoi)

Let $c_p(n)$ be the smallest integer k such that a (finite or countable) metric space X embeds isometrically into ℓ_p^n whenever each subspace $Y \subseteq X$ with at most k points does.

Questions: Find $c_1(3)$ and $c_\infty(3)$.

Known results:

(i) (K. Menger) $c_2(n) = n + 3$.

(ii) (Bandelt & Chepoi) $c_1(2) = 6$.

(iii) (Bandelt & Chepoi) $c_1(n) \geq n^2$ and $c_1(3) \geq 10$.

3.10 m-simplex inequality (Michel Deza)

For a set $X \subset \mathbb{R}^n$, we define the function $\mu_m\colon X^{m+1} \to [0, \infty)$ by letting $\mu_m(x_1, x_2, \ldots, x_{m+1})$ be the m-dimensional volume of $\mathrm{conv}(x_1, x_2, \ldots, x_{m+1})$. Let $s_m = s_m(X)$ be the maximum s such that

$$s \cdot \mu_m(x_1, x_2, \ldots, x_{m+1}) \leq \sum_{i=1}^{m+1} \mu_m(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{m+1}, y)$$

whenever $x_1, x_2, \ldots, x_{m+1}, y \in X$ are all distinct.

- Generally $s_m \leq m + 1$.
- If X is the vertex set of β_n (the n-dimensional crosspolytope), then $s_m = 3$ for all $n \geq m \geq 3$, and $s_m = 1 + \sqrt{3}$ for $n \geq m = 2$.
- If X is the vertex set of γ_n (the n-dimensional cube), then $s_m \to 1$ as $n \to \infty$ for every $m \geq 0$.

- If X is the vertex set of α_n (the n-dimensional simplex), then $s_m = m+1$ for every $m \geq 1$.

Problem: For X of size $k \geq 5$ and $m \geq 4$, does the equality $s_m = m+1$ imply that X is the vertex set of α_n?

We checked it for $s = 1, 2$. It does not hold for $k = 4$ (a parallelogram) and it does not hold for $m = 3$ (the set $X = \{\pm e_1, e_2, \ldots, e_n\}$ is only known non-simplex with $s_3 = 4$).

For more information on this problem see a manuscript by M. Deza, M. Dutour and H. Maehara, available, e.g., from the proposer.

4 Solved or almost solved problems from 2002

4.1 Explicit embedding of ℓ_2 into ℓ_1 (Gideon Schechtman, Piotr Indyk; probably folklore)

It is well known that ℓ_2^d embeds into $\ell_1^{d'}$ with constant distortion and with $d' = O(d)$. In particular, for any $c = 1+\varepsilon > 1$, one can achieve $d' = O(d/\varepsilon^2)$ (this bound is due to Gordon; a slightly weaker one was proved earlier by Figiel, Lindenstrauss, and Milman). However, the proof uses probabilistic method, and, as a result, the embedding construction is not "explicit".

Problem: Give an "explicit" construction that matches or almost matches the above bound.

Remarks. There are several known explicit constructions of embeddings of ℓ_2 into ℓ_1. One uses a matrix with rows consisting of elements of a 4-wise independent family of functions and gives $d' = O(d^2)$ and $c = \sqrt{3}$ [LLR95]; a similar result can also be achieved using an embedding due to König (see [KK01], or an unpublished thesis of L. Rabinovich for similar results). Another construction uses a matrix obtained from Nisan's pseudorandom generator and gives $\varepsilon = 1/d^{\Theta(1)}$ and $d' = 2^{O(\log^2 d)}$ [Ind00].

A solution to this problem would have consequences for constructing a deterministic algorithm for approximate nearest neighbor in high dimensions.

Progress report. Indyk [Ind07] constructed an explicit $(1+1/\log n)$-embedding of ℓ_2^n in ℓ_1^m with $m \leq n 2^{O((\log\log n)^2)}$. The proof combines explicit constructions of "highly orthogonal" configurations of vectors with extractors of randomness—two ingredients from rather remote-looking areas.

Another interesting construction, with $m = (1+o(1))n$ but with distortion more than logarithmic, namely, $(\log n)^{O(\log\log\log n)}$, was provided by Guruswami, Lee, and Razborov [GLR10].

4.2 ℓ_1 metrics into ℓ_2 (Nathan Linial)

Is it true that every n point metric X in ℓ_1 embeds into ℓ_2 with distortion $O(\sqrt{\log n})$? If not, how large can the required distortion be?

Progress report. Lee, Mendel, and Naor [LMN05] proved that any n-point subspace of ℓ_1 can be embedded into ℓ_2 with *average* distortion $O(\sqrt{\log n})$.

More precisely, there is a map that contracts every distance by at most $O(\sqrt{\log n}\,)$, and the average of the expansions of the $\binom{n}{2}$ distances is $O(1)$.

The result of Arora, Lee, and Naor [ALN05] mentioned in Problem 2.2 nearly settles the problem, implying that every n-point subspace of ℓ_1 embeds into ℓ_2 with distortion $O(\sqrt{\log n}\,\log\log n)$.

4.3 Dimension reduction in ℓ_1 (Nathan Linial)

Given an integer n and $\gamma > 1$, what is the least $k = k(n, \gamma)$ such that every n-point metric in ℓ_1 can be embedded in ℓ_1^k with distortion at most γ? Specifically, is $k \leq O(\log n)$ for a bounded γ?

Solution. Essentially answered by Brinkman and Charikar [BC03]. It was shown that a certain n-point space that embeds into ℓ_1 with a constant distortion (namely, the recursive diamond graph, previously used by Newman and Rabinovich as an example of a planar graph requiring $\Omega(\sqrt{\log n})$ distortion to embed into ℓ_2) requires dimension at least $k = n^{\Omega(1/\gamma^2)}$ for embedding into ℓ_1^k with distortion γ. A considerably simpler proof of the Brinkman–Charikar result was obtained by Lee and Naor [LN04] in June 2003. Another elegant and even simpler proof using entropy was provided by Regev [Reg11].

Newman and Rabinovich [NR09b] showed, based on deep work of Batson, Spielman and Srivastava on graph sparsifiers [BSS09], that for every $\varepsilon > 0$ we have $k(n, 1+\varepsilon) \leq c(\varepsilon)n$, improving on the previously best known bound, due to Schechtman [Sch87], of $k(n, 1+\varepsilon) \leq c(\varepsilon)n \log n$

4.4 Pentagonal into ℓ_1 (Michel Deza)

Is any (finite) 5-gonal metric space embeddable into ℓ_1 with distortion bounded by a constant? Here we call a metric space $(2k+1)$-gonal if for every multiset A of k points and every multiset B of $k+1$ points in the space, the sum of all pairwise distances in A plus the sum of all pairwise distances in B does not exceed the sum of all distances between a point of A and a point of B. The case $k = 1$ corresponds to the triangle inequality, and a 5-gonal metric space corresponds to the case $k = 2$. We know, for any $k > k'$, examples of $(2k'+1)$-gonal, but not $(2k+1)$-gonal, metric spaces. Any metric subspace of ℓ_1 is hypermetric (i.e., $(2k+1)$-gonal for every natural k), but not vice versa (for example, the shortest path metric of $K_7 - C_5$ and $K_7 - P_3$).

Solution. Newman, Rabani, and Rabinovich constructed n-point pentagonal metrics (even hypermetrics) requiring distortion at least of order $(\log n)^{0.6}$ for embedding into ℓ_1. This and many other results can be found in [ALN+05].

4.5 Into ℓ_∞^d (Yuri Rabinovich)

Can every planar metric on n points be embedded into $\ell_\infty^{c_1 \log n}$ with distortion at most c_2, for some universal constants c_1, c_2?

Solution. Answered positively by Krauthgamer, Lee, Mendel, and Naor [KLMN03].

4.6 Average distance for line embedding (Yuri Rabinovich)

Is there a constant $c > 0$ such that for any planar metric (X, ρ), a mapping $f\colon X \to \mathbb{R}^1$ exists with

- $|f(x) - f(y)| \le \rho(x, y)$; that is, f is non-expanding, and
- $\sum_{x,y \in X} |f(x) - f(y)| \ge c \cdot \sum_{x,y \in X} \rho(x, y)$; that is, the average distance decreases at most by a constant?

This is known to hold for the metric of a graph of bounded treewidth (with c depending on the treewidth). On the other hand, for the metric of an n-vertex constant-degree expander, any non-expanding embedding into $\mathbb{R}^1$ decreases the average distance by at least $\Omega(\log n)$.

Solution. Answered positively by the proposer [Rab03], even for an arbitrary class of graphs with a fixed excluded minor.

4.7 Probabilistic into trees (Yair Bartal)

Prove or disprove the following conjecture: any metric space on n points can be $O(\log n)$-probabilistically embedded into hierarchically well-separated trees (equivalently, embedded into a convex combination of dominating HSTs). The known lower and upper bounds are $\Omega(\log n)$ and $O(\log n \log\log n)$.

The following definition of HST follows [BBM01], and it slightly differs from the ones in earlier papers, although it is essentially equivalent for $k > 1$. A 1-HST is exactly an *ultrametric*; that is, the metric on the leaves of a rooted tree T (with weighted edges) such that all leaves have the same distance from the root. For a k-HST with $k > 1$ we require that, moreover, $\Delta(v) \le \Delta(u)/k$ whenever v is a child of u in T, where $\Delta(v)$ denotes the diameter of the subtree rooted at v (w.l.o.g. we may assume that each non-leaf has degree at least 2, and so $\Delta(v)$ equals the distance of v to the nearest leaves).

Solution. Resolved positively by Fakcharoenphol, Rao, and Talwar [FRT03], whose construction also significantly simplified previous results on probabilistic embeddings into trees.

4.8 HST in a metric space (Yair Bartal)

Let $R_{\mathrm{HST}}(c, n)$ be the largest number k so that any metric space on n points contains a subspace of size at least k that is c-embeddable into an HST (see problem 4.7). Prove or disprove the following conjecture: There exist constants c and $\varepsilon > 0$ such that $R_{\mathrm{HST}}(c, n) \ge n^{\varepsilon}$. In [BBM01] it is shown that for every $c \ge 3$, $R_{\mathrm{HST}}(c, n) \ge 2^{\Omega(\log^{1-1/c} n)}$.

Solution. A substantial progress in Ramsey-type questions for metric spaces was made by Bartal, Linial, Mendel, and Naor [BLMN03]. They proved that every n-point metric space contains a subspace of size $n^{1-\varepsilon}$ that is embeddable in an ultrametric (1-HST) with $O(\varepsilon^{-1} \log(1/\varepsilon))$ distortion, and this is tight in the worst case up to the $\log(1/\varepsilon)$ factor, as well as many other results for special classes of metric spaces.

The $\log(1/\varepsilon)$ factor in the upper bound was removed by Mendel and Naor [MN07].

4.9 Euclidean subspace in a metric space (Yair Bartal; Nathan Linial)

Let $R_p(c, n)$ be the largest number k so that any metric space on n points contains a subspace of size at least k that is c-embeddable into ℓ_p. Since every HST is an ultrametric, and hence it embeds isometrically into ℓ_2, we have $R_p(c, n) \geq R_{\mathrm{HST}}(c, n)$. Can one get better lower bounds?

Specifically, prove or disprove the following conjecture: For every $\delta > 0$ there exists $c > 0$ such that $R_2(c, n) \geq n^{1-\delta}$. Or, a weaker one: There exist constants c and $\alpha > 0$ such that $R_2(c, n) \geq n^{\alpha}$.

Solution. Also answered in [BLMN03].

4.10 Assouad's problem (reminded by J. Matoušek)

Let X be a separable metric space such that any ball of radius $2r$ can be covered by at most L balls of radius r, for all $r > 0$. Do there exist numbers $k = k(L)$ and $C = C(L)$ such that any such X can be embedded into k-dimensional Euclidean space with distortion at most C? (It is known that for any fixed $p < 1$, the metric space arising from X by raising all distances to power p can be so embedded, with k and C depending on p as well; see Assouad [Ass83].)

Solution. Assouad's question was answered negatively by Semmes [Sem96], based on a theorem of Pansu (thanks to Manor Mendel for letting me know). New strong results around Assouad's problem were obtained by Gupta, Krauthgamer, and Lee [GKL03].

4.11 More examples badly embeddable into ℓ_2 (Nathan Linial)

Bourgain's basic theorem says that $c_2(X) \leq O(\log n)$ for every n-point metric space X. This bound is known to be tight, but so far the only known extreme examples are (the metrics of) constant-degree expander graphs. Can you find other families of explicit examples?

Solution. Khot and Naor [KN05] constructed a new type of example based on good binary codes. The lower bound argument uses Fourier analysis on the Boolean cube.

Update (July 2010): A new, more general construction was found by Newman and Rabinovich [NR09a].

4.12 A variation on a theme by L. Levin (Nathan Linial)

Given a graph G we seek an embedding $\varphi\colon V(G) \to \ell_2^d$ such that

(i) $\|\varphi(x) - \varphi(y)\| \geq 1$ for every two distinct vertices in G, and

(ii) $\|\varphi(x) - \varphi(y)\| \leq 2$ for every two adjacent vertices. (The constant 2 is arbitrary).

Our goal is to minimize the dimension d. By simple volume considerations $d \geq \Omega(\rho(G))$, where

$$\rho(G) := \max_{x \in G, r \geq 1} \frac{\log(1 + |B_r(x)|)}{\log(r+1)}$$

is G's growth rate. (Essentially, the smallest t such that every ball of radius r in G has no more than r^t vertices). Can the smallest d be bounded in terms of ρ? Perhaps even $d \leq O(\rho)$?

Solution. Resolved by Krauthgamer and Lee [KL03], who show that $d = O(\rho \log \rho)$ can be achieved, while there are cases requiring $d = \Omega(\rho \log \rho)$.

4.13 Volume-respecting for large k (Uriel Feige)

The following question was asked in my talk on volume respecting embeddings. For the relevant definitions see Feige [Fei00]. The question is whether every finite metric space has a volume respecting embedding with distortion $O(\log n)$. Specifically, the problem is that when we consider sets of k vertices and k is large, then the proof no longer shows a distortion of $O(\log n)$. The current proof gives upper bounds on the distortion that grow at a rate of $\sqrt{k}$. Another related result is that of Satish Rao that can be used to give a different embedding that has $O(\log k)$ dependence on k, but a worse dependence on n.

Solution. Answered positively by Krauthgamer, Lee, Mendel, and Naor [KLMN03].

4.14 Probabilistic into spanning trees (Yair Bartal; Nathan Linial)

What is the best statement of the form: Given any graph G with nonnegative edge-weights, there is a distribution π on G's spanning trees such that for every two vertices x and y,

$$\mathbf{E}_\pi d_T(x,y) \leq O(d_G(x,y) f(n)).$$

Here d_G and d_T are distances in G and in the tree T, and $\mathbf{E}_\pi$ stands for expectation w.r.t. the distribution π. Alon et al. [AKPW95] showed this with $f(n) \leq \exp(O(\sqrt{\log n \log \log n}))$, but this might be true even with $f(n) \leq O(\log n)$. What is the truth?

Progress report (January 2005): A great improvement, to $O((\log n \log \log n)^2)$, was made by Emek, Elkin, Spielman, and Teng [EEST05].

Update (July 2010): Abraham, Bartal and Neiman [ABN08] obtained a further (almost sharp) improvement to $f(n) = O\left(\log n (\log \log n)(\log \log \log n)^3\right)$.

5 Solved or almost solved problems from 2003

5.1 Realizability of graphs (Robert Connelly and Maria Sloughter)

A finite graph G, without loops or multiple edges, is d-realizable if for every map, a realization, of its vertices in $\mathbb{R}^N$, $\mathbf{p} = (p_1, \ldots, p_n)$, $p_i \in \mathbb{R}^N$, there is another realization $\mathbf{q} = (q_1, \ldots, q_n)$, $q_i \in \mathbb{R}^d$ such that for each edge $\{i, j\}$ of G, $|p_i - p_j| = |q_i - q_j|$.

The general problem is to characterize those graphs that are d-realizable and to find an algorithm that determines whether a graph is d-realizable in polynomial time. The most interesting case is $d = 3$.

A *minor* H of a graph G is obtained by successively contracting and/or deleting edges of G. The graph K_k is the complete graph on k vertices (every vertex is joined to all the others); V_8 is the graph obtained by joining opposite vertices of a cycle of length 8; and $C_5 \times C_2$ is two copies of a cycle of length 5, where corresponding vertices are joined by edges.

The class of the 3-realizable graphs is minor-closed, and the list of minimal forbidden minors consists of K_5, the octahedron, and the graphs among V_8 and $C_5 \times C_2$ that are not 3-realizable. In any case there is a linear time algorithm to determine 3-realizability.

The main question is to decide whether V_8 and $C_5 \times C_2$ are 3-realizable.

Solution. Belk [Bel04] proved that both V_8 and $C_5 \times C_2$ are 3-realizable.

5.2 $\{0,1,2\}$-metrics into ℓ_p (Manor Mendel)

A $\{0,1,2\}$-metric is a metric in which all distances are 0, 1, or 2. We denote by $A_{\ell_p}(2)$ the supremum over all finite $\{0,1,2\}$-metrics M of $c_{\ell_p}(M)$. Question: What is $A_{\ell_p}(2)$ for $p > 2$?

Remarks (see [BLMN03] for more details):

1. Clearly, $1 \le A_{\ell_p}(2) \le 2$.
2. For $p \le 2$, $A_{\ell_p}(2) = 2$.
3. For $p \ge 2$, $A_{\ell_p}(2) \ge 4^{1/p}$.
4. $A_{\ell_p}(2) = 2$ implies that the metric Ramsey function $R_{\ell_p}(\alpha, n) \le C(\alpha) \log n$ for $\alpha < 2$.
5. $A_{\ell_p}(2) < 2$ implies the following impossibility of dimension reduction in ℓ_p: There exist constants $\varepsilon > 0$ and $c > 0$ such that for all sufficiently large n, there exists n-point $X \subset \ell_p$ such that any embedding $X \to \ell_p^d$ with $d \le cn$ has distortion larger than $1+\varepsilon$. (This follows from a Bourgain-style counting/volume argument).

Another (related) natural question: Let $B_{\ell_p}(\Phi)$ be the supremum of $c_{\ell_p}(M)$ over all finite metric spaces M with aspect ratio at most Φ, where the aspect ratio of a metric space is the diameter divided by the minimum interpoint distance. What can be said about $B_{\ell_p}(\Phi)$? Clearly $B_{\ell_p}(\Phi) \le \Phi$, and a result of Matoušek implies that $B_{\ell_p}(\Phi) \ge 1 + c\Phi/p$ for a universal positive constant c.

Solution. Charikar and Karagiozova [CK05] proved that $A_p(2) = 2$ for all $p > 2$.

5.3 Nonlinear type and cotype (Assaf Naor)

One of the deepest aspects of the local theory of Banach spaces is the theory of type and cotype (see e.g. [MS86]). Various problems that have to do with the development of nonlinear analogs of these concepts have remained open for many years.

To motivate things, we note that Hilbert space is characterized among all normed spaces X by the parallelogram identity, which states that for every $x_1, \ldots, x_n \in X$,

$$\mathbb{E}\left\|\sum_{i=1}^{n} \varepsilon_i x_i\right\|^2 = \sum_{i=1}^{n} \|x_i\|^2,$$

the expectation being over a random choice of a sign vector $\varepsilon \in \{-1, 1\}^n$. The power of this identity has led researchers to study the following weakenings of it. A normed X spaces is said to have *type p with constant T* if

$$\mathbb{E}\left\|\sum_{i=1}^{n} \varepsilon_i x_i\right\|^p \leq T^p \sum_{i=1}^{n} \|x_i\|^p,$$

and it has *cotype q with constant C* if

$$\mathbb{E}\left\|\sum_{i=1}^{n} \varepsilon_i x_i\right\|^q \geq \frac{1}{C^q} \sum_{i=1}^{n} \|x_i\|^q.$$

for all $x_i \in X$ and for all n. The least T above is denoted $T_p(X)$ and the least C above is denoted $C_q(X)$. In particular, L_p has type p and cotype 2 for $1 \leq p \leq 2$ and type 2 and cotype p for $p \geq 2$. No Banach space has type greater than 2 and cotype less than 2.

These parameters control numerous geometric properties of X. Most notably, Kwapien (see e.g. [T-J89]) proved that the Euclidean distortion satisfies $c_2(X) = O(T_2(X) \cdot C_2(X))$. An analog of this result for metric spaces is still elusive and would probably be very useful.

Enflo type. This is one of the possible definitions of a "nonlinear type": A metric space (X, d) is said to have Enflo-type p with constant K if for every n and every $f : \{-1, 1\}^n \to X$,

$$\mathbb{E}\left[d(f(\varepsilon), f(-\varepsilon))^p\right] \leq K^p \sum_{i=1}^{n} \mathbb{E}\left[d(f(\varepsilon), f(\varepsilon \oplus i))^p\right],$$

where $\varepsilon \oplus i$ is obtained from ε by flipping the ith coordinate. The question is whether the Enflo type coincides with the (linear) type for every normed space. Pisier [Pis86] proved that if a normed space has type p then it has Enflo-type $p - \varepsilon$ for every $\varepsilon > 0$, and Naor and Schechtman [NS02] showed that the loss in ε is unnecessary for a wide class of Banach spaces (the so called UMD spaces), in particular for L_p spaces.

Markov type. Following Ball [Bal92], a metric space (X, d) is said to have Markov type $p > 0$ if there exists a constant $K > 0$ such that for every symmetric stochastic $n \times n$ matrix A, every $x_1, \ldots, x_n \in X$, and every integer

$t > 0$,

$$\sum_{i,j}(A^t)_{ij}d(x_i,x_j)^p \le K^p t \sum_{i,j} A_{ij}d(x_i,x_j)^p.$$

This notion is useful in various embedding problems, as well as in the Lipschitz extension problem (see Problem 3.6 below). Hilbert space is easily seen to have Markov type 2 (with $K = 1$), but not much more is known about this notion. When restricted to normed spaces, the relation between Markov type and linear type is unclear. In particular, we have the following long-standing open problem:

K. Ball's Markov type 2 problem. Does L_p, $p > 2$, have Markov type 2?

Solution (September 2004): This has been answered affirmatively by Naor, Peres, Schramm, and Sheffield [NPSS04].

Another question is:

Problem. Do trees have Markov type 2?

Solution (September 2004): This is also true by [NPSS04].

When (X, d) is a graph metric, we can naturally define a related notion to Markov type which we call *edge Markov type:* simply demand that in the definition above, A_{ij} is non-zero only if $\{i, j\}$ is an edge. The relationship between these two notions is unclear; perhaps they are equivalent (up to a universal loss in the constant K).

Nonlinear cotype. Is there a nonlinear version of cotype? By that we mean a metric property (i.e. a property which involves only distances) that coincides (or nearly coincides) with the linear notion of cotype when restricted to normed spaces. In particular, it would be desirable for L_1 to have nonlinear cotype 2, since there are no known nontrivial bi-Lipschitz invariants of L_1 (apart from the Poincaré inequality for expanders). After such a notion is defined, the hope is that, with both nonlinear type and nonlinear cotype at hand, we can transfer some of the linear theory to the setting of arbitrary metric spaces.

Solution (2005): Mendel and Naor [MN05] suggested a definition of metric cotype and proved several satisfactory properties for it.

5.4 Isometric vs. isomorphic L_2 squared (James R. Lee)

For a metric space (X, d), let $(X, \sqrt{d})$ be the metric space with distance function $\sqrt{d}(a, b) = \sqrt{d(a, b)}$. We define three classes of metric spaces.

$$\begin{aligned}
\mathsf{NEG} &= \{(X,d) : c_2(X,\sqrt{d}) = 1\},\\
\mathsf{NEG}_D &= \{(X,d) : c_{\mathsf{NEG}}(X) \le D\},\\
\widetilde{\mathsf{NEG}}_D &= \{(X,d) : c_2(X,\sqrt{d}) \le D\}.
\end{aligned}$$

The metrics in NEG are called *metrics of negative type* or sometimes *squared ℓ_2 metrics.* It is clear that if $X \in \mathsf{NEG}_D$, then $X \in \widetilde{\mathsf{NEG}}_{D'}$ for some $D' = D'(D)$, but what about the converse?

Question: If $X \in \widetilde{\mathsf{NEG}}_{O(1)}$, does it follow that $X \in \mathsf{NEG}_{O(1)}$? In words, if X with the square root of the metric can be embedded into ℓ_2 with distortion bounded by a constant, can X with the original metric be approximated by a metric of negative type with a constant-bounded distortion?

We note that if X admits β-padded decompositions (see Problem 5.5), then $X \in \widetilde{\mathsf{NEG}}_{O(\beta)}$ [LMN05], so a positive answer would expand widely the class of metrics known to be "L_2 squared." It is also interesting to note that the recent paper [ARV04] uses only the $\widetilde{\mathsf{NEG}}$ property, and not the apparently stronger NEG property in showing that the integrality gap of the Sparsest Cut Semidefinite Program is $O(\sqrt{\log n})$.

Solution (2010). Solved negatively by Lee and Moharrami [LM10], who showed that there exist arbitrarily large n-point metric spaces X_n in $\widetilde{\mathsf{NEG}}_{O(1)}$, yet $c_{\mathsf{NEG}}(X_n) = \Omega((\log n)^{1/4}/\log\log n)$.

5.5 Decomposability of metrics (Kunal Talwar)

Given a metric space (X, d), a partition $\mathcal{P} = \{X_1, X_2, \ldots, X_k\}$ of X is called a Δ-*decomposition* if each part X_i has diameter at most Δ. We say a metric (X, d) is α-decomposable if given any $\Delta > 0$, there exists a distribution $\mathcal{D}$ over Δ-decompositions of X such that for any $u, v \in X$,

$$\Pr_{\mathcal{D}}\Big[u \text{ and } v \text{ fall in different parts}\Big] \quad \le \quad \alpha \cdot \frac{d(u,v)}{\Delta}.$$

A somewhat stronger and often useful notion is that of *padded decomposition*. For a partition $\mathcal{P} = \{X_1, X_2, \ldots, X_k\}$ of X, and for $u \in X_i$, define $d(u, \mathcal{P}) = d(u, X \setminus X_i)$. We say a distribution $\mathcal{D}$ over Δ-decompositions is β-padded if for any $r \in (0, 1)$ and any $u \in X$,

$$\Pr_{\mathcal{D}}\Big[d(u,\mathcal{P}) \le r\Delta\Big] \le r\beta.$$

Such decompositions have several applications to approximation algorithms and embeddings.

Let α_X and β_X be the best possible parameters for a metric X in the above definitions. It is easy to show that $\alpha_X \le \beta_X$. Any n-point metric satisfies $\beta_X = O(\log n)$. Metrics derived from graphs that exclude a K_r minor are known to satisfy $\beta_X = O(r^2)$. The best lower bound known is $\Omega(\log r)$. What is the right answer for such metrics?

How does the decomposability of a metric relate to its other properties? In particular, we can ask:

Question: Can $c_1(X)$, the smallest distortion needed for embedding of a metric space X into ℓ_1, be bounded above by a function of β_X?

Solution. Cheeger and Kleiner [CK06a] proved that the Heisenberg group $\mathbb{H}$, which is a doubling metric space and hence $\beta_{\mathbb{H}} < \infty$, satisfies $c_1(\mathbb{H}) = \infty$. Cheeger, Kleiner and Naor [CKN09] showed that there exist arbitrarily large n-point metric spaces X_n with $\beta_{X_n} = O(1)$, yet $c_1(X_n) = (\log n)^{\Omega(1)}$.

5.6 Trees into ℓ_1 (Anupam Gupta, James R. Lee, and Kunal Talwar)

Given a constant D, what is the smallest value $d = d(D, n)$ such that every tree T on n vertices embeds into ℓ_1^d with distortion at most D? Charikar and Sahai [CS02] show that $d(1+\varepsilon, n) \le O(\varepsilon^{-2}\log^2 n)$, and a lower bound of $\Omega(\log n)$ can be shown for $d(O(1), n)$ using elementary volume arguments. What is the correct trade-off?

Can we even embed the complete binary tree on n vertices into $O(\log n)$-dimensional ℓ_1 with $O(1)$ distortion?

Bartal and Mendel [BM04] give a $(1+\varepsilon)$-embedding of any *ultrametric* on n vertices into ℓ_1^k with $k = O(\varepsilon^{-2}\log n)$.

Solution. Lee, de Mesmay, and Moharrami [LMM11] proved that every n-point metric tree admits a $(1+\varepsilon)$-embedding in ℓ_1 of dimension at most $C(\varepsilon)\log n$ for every $\varepsilon > 0$, with $C(\varepsilon) = O((\frac{1}{\varepsilon})^4 \log \frac{1}{\varepsilon})$. For complete binary trees with unit-length edges they obtain $C(\varepsilon) = O((\frac{1}{\varepsilon})^2)$.

References

[ABN08] I. Abraham, Y. Bartal and O. Neiman: Nearly Tight Low Stretch Spanning Trees, in *49th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2008, pages 781–790.

[ADGIR03] A. Andoni, M. Deza, A. Gupta, P. Indyk, and S. Raskhodnikova: Lower bounds for embedding of edit distance into normed spaces, *Proc. of the 14th ACM-SIAM Sympos. Discrete Algorithms*, 2003, pages 523–526.

[AKPW95] N. Alon, R. M. Karp, D. Peleg, and D. West: A graph-theoretic game and its application to the k-server problem, *SIAM J. Computing*, 24(1):78–100, 1995.

[ALN05] S. Arora, J. R. Lee, and A. Naor: Euclidean distortion and the sparsest cut, *Proc. 37th Annual ACM Symposium on Theory of Computing* 553–562, 2005.

[ALN+05] S. Arora, L. Lovász, I. Newman, Y. Rabani, Yu. Rabinovich, and S. Vempala: Local versus global properties of metric spaces, manuscript, 2005.

[ARV04] S. Arora, S. Rao, and U. Vazirani: Expander flows and a $\sqrt{\log n}$-approximation to sparsest cut, in *Proc. 36th Annual ACM Symposium on Theory of Computing*, Chicago, IL, 2004, pages 222–231.

[Ass83] P. Assouad: Lipschitz embeddings into $\mathbb{R}^n$ (in French), *Bull. Soc. Math. Fr.* 111:429–448, 1983.

[Bal92] K. Ball: Markov chains, Riesz transforms and Lipschitz maps, *Geom. Funct. Anal.*, 2(2):137–172, 1992.

[BBM01] Y. Bartal, B. Bollobás, and M. Mendel: A Ramsey-type theorem for metric spaces and its applications for metrical task systems and related problems, *Proc. 42nd IEEE Symposium on Foundat. Comput. Sci.*, 2001, pages 396–405.

[BC03] B. Brinkman and M. Charikar: On the impossibility of dimension reduction in ℓ_1, *Proc. 44th Annual IEEE Symposium on Foundations of Computer Science*, 2003, pages 514–523.

[BCIS05] M. Bădoiu, J. Chuzhoy, P. Indyk, and A. Sidiropoulos. Low-distortion embeddings of general metrics into the line. In *Proc. 37th ACM Symposium on Theory of Computing*, 2005.

[Bel04] M. Belk: Relaizability of graphs in three dimensions, *Discr. Comput. Geometry* 37:139–162, 2007.

[BHDM07] M. Bateni, M. Hajiaghayi, E. D. Demaine, and M. Moharrami: Plane embeddings of planar graph metrics, *Discr. Comput. Geometry* 38:615–637, 2007.

[BLMN03] Y. Bartal, N. Linial, M. Mendel, and A. Naor: On some low-distortion metric Ramsey problems, *Ann. Math.* (2) 162,2:643–709, 2005.

[BMMV02] R. Babilon, J. Matoušek, J. Maxová, and P. Valtr: Low-distortion embeddings of trees, *Proc. Graph Drawing 2001*, Springer, 2002, pages 343–351.

[BM04] Y. Bartal and M. Mendel: Randomized algorithms for the k-server problem in growth rate bounded metric spaces, In *Proc. of the 15th ACM–SIAM Sympos. on Discrete Algorithms*, 2004, pages 666–671.

[BMW86] J. Bourgain, V. Milman, H. Wolfson: On type of metric spaces, *Trans. Amer. Math. Soc.* 294:295–317, 1986.

[BSS09] J. D. Batson, D. A. Spielman and N. Srivastava: Twice-Ramanujan sparsifiers, in *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, 2009, pages 255–262.

[CC95] L. Cai and D. G. Corneil: Tree spanners, *SIAM J. Discrete Math.* 8:359–387, 1995.

[CCGGP98] M. Charikar, C. Chekuri, A. Goel, S. Guha, and S.A. Plotkin: Approximating a finite metric by a small number of tree metrics, *Proc. 39th IEEE Sympos. Foundat. Comput. Sci.*, 1998, pages 379–388.

[CGNRS03] C. Chekuri, A. Gupta, I. Newman, Yu. Rabinovich, and A. Sinclair: Embedding k-outerplanar graphs into L_1, *Proc. 14th ACM–SIAM Sympos. Discr. Algo.* 2003.

[CGR05] S. Chawla, A. Gupta, and H. Räcke: Embeddings of negative-type metrics and an improved approximation to generalized sparsest cut, in *Proc. 16th ACM–SIAM Sympos. Discr. Algo.*, 2005.

[Cha02] M. Charikar: Similarity estimation techniques from rounding, *Proc. 34th Annu. ACM Sympos. on Theory of Computing*, 2002.

[CJLV08] A. Chakrabarti, A. Jaffe, J. R. Lee and J. Vincent: Embeddings of Topological Graphs: Lossy Invariants, Linearization, and 2-Sums, *49th Annual IEEE Symposium on Foundations of Computer Science*, 2008, pages 761-770.

[CK05] M. Charikar and A. Karagiozova: A tight threshold for metric Ramsey phenomena, in *Proc. 16th ACM–SIAM Sympos. Discr. Algo.*, 2005.

[CK06] M. Charikan and R. Krauthgamer: Embedding the Ulam metric into ℓ_1, *Theory of Computing* 2:207–224, 2006.

[CK06a] J. Cheeger and B. Kleiner: Differentiating maps into L^1 and the geometry of BV functions, preprint, ArXiv:math/0611954, 2006.

[CKN09] J. Cheeger, B. Kleiner and A. Naor: A $(\log n)^{\Omega(1)}$ integrality gap for the Sparsest Cut SDP, *Proceedings of 50th Annual IEEE Symposium on Foundations of Computer Science* 2009, pages 555–564.

[CKNZ01] C. Chekuri, S. Khanna, J. Naor, and L. Zosin: Approximation algorithms for the metric labeling problem via a new linear programming formulation, *Proc. 12th ACM–SIAM Sympos. Discr. Algo.* 2001, pages 109–118.

[CM02] G. Cormode and S. Muthukrishnan: The string edit distance matching problemwith moves, *Proc. 13th ACM–SIAM Sympos. Discr. Algo.*, 2002.

[CS02] M. Charikar and A. Sahai: Dimension reduction in the ℓ_1 norm, *Proc. 43rd Annual IEEE Symposium on Foundations of Computer Science* 2002, pages 551–560.

[Deh08] P. Dehornoy: On the 3-distortion of a path. *Eur. J. Comb.* 29(1):171–178, 2008.

[EEST05] M. Elkin, Y. Emek, D. A. Spielman, and S.-H. Teng: Lower-stretch spanning trees, preprint, `http://arxiv.org/abs/cs.DS/0411064`.

[EP04] Y. Emek and D. Peleg: Approximating Minimum Max-Stretch spanning Trees on unweighted graphs, in *Proc. 15th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2004, pages 261–270.

[ESZ10] J. Edmonds, A. Sidiropoulos, A. Zouzias, Inapproximability for planar embedding problems, in *Proc. 21st ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2010.

[Fei00] U. Feige: Approximating the bandwidth via volume respecting embeddings, *J. Comput. Syst. Sci*, 60:510–539, 2000.

[FHRT03] J. Fakcharoenphol, C. Harrelson, S. Rao, and K. Talwar: An improved approximation algorithm for the 0-extension problem, *Proc. of the 14th ACM-SIAM Symposium on Discrete Algorithms*, 2003.

[FI99] M. Farach-Colton and P. Indyk: Approximate nearest neighbor algorithms for Hausdorff metrics via embeddings, *Proc. 40th IEEE Symposium on Foundations of Computer Science*, 1999, pages 171–179.

[FRT03] J. Fakcharoenphol and S. Rao and K. Talwar: A tight bound on approximating arbitrary metrics by tree metrics, *Proc. ACM Symposium on Theory of Computing*, 2003, pages 448–455.

[GKL03] A. Gupta, R. Krauthgamer, and J. R. Lee: Bounded geometries, fractals, and low-distortion embeddings, *Proc. 44th IEEE Symposium on Foundations of Computer Science*, 2003, pages 534–543.

[GLR10] V. Guruswami, J. R. Lee, and A. Razborov, Almost Euclidean subspaces of ℓ_1^N via expander codes, *Combinatorica* 30,1:P47–68, 2010.

[GNRS99] A. Gupta, I. Newman, Yu. Rabinovich, and A. Sinclair: Cuts, trees and ℓ_1-embeddings of graphs, *Proc. 40th IEEE Symposium on Foundations of Computer Science,* 1999, pages 399–409.

[Ind00] P. Indyk: Stable distributions, pseudorandom generators, embeddings and data stream computation, *Proc. 41st IEEE Symposium on Foundations of Computer Science*, 2000, pages 189–197.

[Ind02] P. Indyk: Approximate nearest neighbor algorithms for Fréchet distance via product metrics, *Proc. 18th ACM Symposium on Computational Geometry*, 2002.

[Ind04] P. Indyk: Approximate Nearest Neighbor under Edit Distance via Product Metrics, *Proc. 15th ACM-SIAM Symposium on Discrete Algorithms*, 2004.

[Ind07] P. Indyk: Uncertainty principles, extractors, and explicit embeddings of L2 into L1, *Proc. 39th Annual ACM Symposium on Theory of Computing*, 2007.

[JL84] W. Johnson, J. Lindenstrauss: Extensions of Lipschitz maps into a Hilbert space, *Contemporary Math. 26 (Conference in modern analysis and probability)*, pages 189–206, Amer. Math. Soc. 1984.

[KK01] A. Koldobsky and H. König: Aspects of the isometric theory of Banach spaces, in Johnson, W. B. (ed.) et al., *Handbook of the geometry of Banach spaces*, Vol. 1, Elsevier, Amsterdam, 2001, pages 899–939.

[KKMR06] H. Karloff, S. Khot, A. Mehta, and Y. Rabani: On earthmover distance, metric labeling, and 0-extension, In *Proc. 38th Annual ACM Symposium on Theory of Computing (STOC)*, 2006, pages 547–556.

[KL03] R. Krauthgamer, J. R. Lee: The intrinsic dimensionality of graphs, in *Proc. 35th Annual ACM Symposium on Theory of Computing*, 2003.

[KLMN03] R. Krauthgamer, J. R. Lee, M. Mendel, A. Naor: Measured descent: A technique for optimal embeddings, in *Proc. 45th IEEE Symposium on Foundations of Computer Science (FOCS)*, 2004, pages 434–443.

[KN05] S. Khot and A. Naor: Nonembeddability theorems via Fourier analysis, *Mathematische Annalen* 334,4:821–852, 2006.

[KV05] S. Khot and N. Vishnoi: The unique games conjecture, integrality gap for cut problems and the embeddability of negative type metrics into ℓ_1, in *Proc. 46th IEEE Symposium on Foundations of Computer Science (FOCS)*, 2005.

[KR06] R. Krauthgamer and Y. Rabani: Improved lower bounds for embeddings into L_1, In *17th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2006.

[LLR95] N. Linial, E. London, and Yu. Rabinovich: The geometry of graphs and some its algorithmic applications, *Combinatorica*, 15:215–245, 1995.

[LMM11] J. R. Lee, A. de Mesmay, and M. Moharrami: Dimension reduction for finite trees in ℓ_1, arXiv preprint 1108.2290v2, 2011.

[LMN02] N. Linial, A. Magen, and N. Naor: Euclidean embeddings of regular graphs—the girth lower bound, *Geometric and Functional Analysis* 12:380–394,2002.

[LMN05] J. R. Lee, M. Mendel, and A. Naor: Metric structures in L_1: Dimension, snowflakes, and average distortion, *European Journal of Combinatorics* 26,8:1180–1190, 2005.

[LM10] J. R. Lee and M. Moharrami: Bilipschitz snowflakes and metrics of negative type, in *Proceedings of the 42nd ACM Symposium on Theory of Computing*, 2010, pages 621–630.

[LN04] J. R. Lee and A. Naor: Embedding the diamond graph in L_p and dimension reduction in L_1, *Geom. Funct. Anal.* 14 (2004), no. 4, 745–747.

[LN04a] J. R. Lee and A. Naor: Absolute Lipschitz extendability. *C. R. Math. Acad. Sci. Paris* 338(2004), no. 11, 859–862.

[LN06] J. R. Lee and A. Naor: L_p metrics on the Heisenberg group and the Goemans–Linial conjecture, In *Proc. 47th IEEE Symposium on Foundations of Computer Science*, 2006, pages 99–108.

[LP01] U. Lang and C. Plaut: Bilipschitz embeddings of metric spaces into space forms, *Geom. Dedicata* 87:285–307, 2001.

[LR07] J. R. Lee and P. Raghavendra: Coarse Differentiation and Multi-flows in Planar Graphs, in *proc. of APPROX*, 2007, pages 228-241.

[LS09] J. R. Lee and A. Sidiropoulos: On the geometry of graphs with a forbidden minor, in *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, 2009, pages 245–254.

[Mat96] J. Matoušek: On the distortion required for embedding finite metric spaces into normed spaces, *Israel J. Math.* 93:333–344, 1996.

[Mat90] J. Matoušek: Bi-Lipschitz embeddings into low-dimensional Euclidean spaces, *Comment. Math. Univ. Carolinae*, 31:589–600, 1990. Also available at `kam.mff.cuni.cz/~matousek`.

[MN05] M. Mendel and A. Naor: Metric cotype, submitted, extended abstract to appear in *17th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2006.

[MN07] M. Mendel and A. Naor: Ramsey partitions and proximity data structures, *J. European Math. Soc.* 9(2):253–275, 2007.

[MS86] V. D. Milman and G. Schechtman: *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, 1986.

[MS08] J. Matoušek, A. Sidiropoulos, Inapproximability for metric embeddings into $\mathbb{R}^d$, *Trans. Amer. Math. Soc.*, in press. Extended abstract: Proc. IEEE Symposium on Foundations of Computer Science (FOCS), 2008, pages 405–413.

[NPSS04] A. Naor, Y. Peres, O. Schramm and S. Sheffield: Markov chains in smooth Banach spaces and Gromov hyperbolic metric spaces, *Duke Mathematical Journal* 134,1:165–197, 2006.

[NR09a] I. Newman and Yu. Rabinovich: Hard metrics from Cayley graphs of Abelian groups, *Theory of Computing* 5, Article 6:125-134, 2009.

[NR09b] I. Newman and Yu. Rabinovich: On Cut Dimension of ℓ_1 Metrics and Volumes, and Related Sparsification Techniques. Preprint, 2009. Available at http://arxiv.org/abs/1002.3541.

[NS02] A. Naor and G. Schechtman: Remarks on non linear type and Pisier's inequality, *J. Reine Angew. Math.*, 552:213–236, 2002.

[NS06] A. Naor and G. Schechtman: Planar earthmover is not in L_1, In *Proc. 47th IEEE Symposium on Foundations of Computer Science*, 2006, pages 655–666.

[Ost11] M. I. Ostrovskii: Low-distortion embeddings of graphs with large girth, preprint, arXiv 1108.2542v2, 2011.

[OR05] R. Ostrovsky and Y. Rabani: Low distortion embeddings for edit distance, in *Proc. 37th Annual ACM Symposium on Theory of Computing*, 2005.

[Pis86] G. Pisier: Probabilistic methods in the geometry of Banach spaces, In *Probability and analysis (Varenna, 1985)*, volume 1206 of *Lecture Notes in Math.*, pages 167–241. Springer, Berlin, 1986.

[Rab03] Yu. Rabinovich: On Average Distortion of Embedding Metrics into l_1 and into the Line, Proc. 35th Annual ACM Symposium on Theory of Computing, 2003, pages 456–462.

[Reg11] O. Regev: Entropy-based bounds on dimension reduction in L_1, preprint, arXiv:1108.1283, 2011.

[RRS89] J. Reiterman, V. Rödl, and E. Šiňajová: Embeddings of graphs in Euclidean spaces, *Discrete Comput. Geom.* 4:349–364, 1989.

[Sch87] G. Schechtman: More on embedding subspaces of L_p in l_r^n. *Compositio Math.* 61 (1987), no. 2, 159–169.

[Sem96] S. Semmes: On the nonexistence of bilipschitz parameterization and geometric problems about A_{oo} weights, *Rev. Mat. Iberoamericana* 12: 337–410, 1996.

[T-J89] N. Tomczak-Jaegermann: *Banach-Mazur distances and finite-dimensional operator ideals*, volume 38 of *Pitman Monographs and Surveys in Pure and Applied Mathematics*, Longman Scientific & Technical, Harlow, 1989.