Discrete Comput Geom 29:19–21 (2003)
DOI: 10.1007/s00454-002-2827-z

Discrete & Computational
Geometry

The Euclidean Distortion of Complete Binary Trees*

Nathan Linial[1] and Michael Saks[2]

[1]Institute of Computer Science, Hebrew University,
Jerusalem 91904, Israel
nati@cs.huji.ac.il

[2]Department of Mathematics, Rutgers University,
Hill Center, 110 Frelinghuysen Road, Piscataway, NJ 08854, USA
saks@math.rutgers.edu

Abstract. Bourgain [1] showed that every embedding of the complete binary tree of depth n into l_2 has metric distortion $\geq \Omega(\sqrt{\log n})$. An alternative proof was later given by Matousek [3]. This note contains a short proof for this fact.

A mapping $\varphi\colon X \to Y$ between metric spaces (X, d) and (Y, ρ) has *distortion* $\leq \gamma$ if there is a real $a > 0$, such that

$$\forall x_1, x_2 \in X, \qquad \gamma a \cdot \rho(\varphi(x_1), \varphi(x_2)) \geq d(x_1, x_2) \geq a \cdot \rho(\varphi(x_1), \varphi(x_2)).$$

Every graph G induces a metric d_G on its vertex set, where $d_G(u, v)$ is the length of the shortest path in G joining u and v. In this note, $\|\cdot\|$ denotes the l_2 norm.

Here we give a short proof of:

Theorem 1 [1]. *Every mapping of T_n, the complete binary tree of depth n, into l_2 has distortion $\geq \Omega(\sqrt{\log n})$.*

This was previously proved (in a more general form) in [1] and [3]. The bound is tight; an l_2-embedding of T_n with distortion $O(\sqrt{\log n})$ appears in [1]. For a broader discussion of graph embeddings and distortion see, e.g., [2].

* This work was done in part while the authors were visiting Microsoft Research, Redmond, Washington. Nathan Linial was supported in part by grants from the US–Israel Binational Science Fund and from the Israel Science Foundation. Michael Saks was supported in part by NSF Grant CCR-9988526.

Our key tool is a simple geometric inequality. For a positive integer n, define $\Gamma = \Gamma_n$ to be the set $\{(p, q)\colon\ 0 \le p < q \le n, q - p = 2^i \text{ for some } i \ge 1\}$.

Lemma 1. *Let $x_0, \ldots, x_n$ be real vectors. Then*

$$\sum_{(p,q)\in\Gamma} \frac{\|x_p - 2x_{(p+q)/2} + x_q\|^2}{(p-q)^2} \le \sum_{p=0}^{n-1} \|x_{p+1} - x_p\|^2.$$

Proof. For real vectors, a, b, c the parallelogram identity $P(a, b, c)$ says $\|a - 2b + c\|^2 + \|a - c\|^2 = 2\|a - b\|^2 + 2\|b - c\|^2$. Summing $(1/(p-q)^2)P(x_p, x_{(p+q)/2}, x_q)$ over $(p, q) \in \Gamma$ yields

$$\sum_{(p,q)\in\Gamma} \left(\frac{\|x_p - 2x_{(p+q)/2} + x_q\|^2}{(p-q)^2} + \frac{\|x_p - x_q\|^2}{(p-q)^2} \right)$$
$$= \sum_{(p,q)\in\Gamma} \left(\frac{2\|x_p - x_{(p+q)/2}\|^2}{(p-q)^2} + \frac{2\|x_{(p+q)/2} - x_q\|^2}{(p-q)^2} \right).$$

For a, b with $1 \le a < b \le n$, the first summand on the right is $\|x_a - x_b\|^2/2(b-a)^2$ if and only if $(p, q) = (a, 2b - a) \in \Gamma$ and the second summand is $\|x_a - x_b\|^2/2(b-a)^2$ if and only if $(p, q) = (2a - b, b) \in \Gamma$. In each case, $b - a$ must be a power of 2. Therefore, the right-hand side does not exceed $\sum(\|x_a - x_b\|^2/(a-b)^2)$ where the sum is over pairs (a, b) such that $b - a$ is a power of 2. Separating the terms where $b = a + 1$, we bound the right-hand side from above by

$$\sum_{a=0}^{n-1} \|x_{a+1} - x_a\| + \sum_{(a,b)\in\Gamma} \frac{\|x_a - x_b\|^2}{(a-b)^2}.$$

Comparing this with the summation on the left-hand side yields the lemma. □

Proof of Theorem 1. Let f map $V(T_n)$ into l_2. We may assume f is nonexpansive, i.e., for every two vertices $\|f(x) - f(y)\| \le d_T(x, y)$. We seek a pair of vertices w, w' for which $\|f(w) - f(w')\|/d_T(w, w')$ is small. A *fork* in T is a quadruple of vertices $\Phi = (u, v, w, w')$, where v is a descendant of u, the least common ancestor of w, w' is v and $d_T(u, v) = d_T(v, w) = d_T(v, w')$. We let

$$\delta(\Phi) = \frac{\|f(u) - 2f(v) + f(w)\|}{d_T(u, w)} \quad \text{and} \quad \delta'(\Phi) = \frac{\|f(u) - 2f(v) + f(w')\|}{d_T(u, w')}.$$

By the triangle inequality:

$$\frac{\|f(w) - f(w')\|}{d_T(w, w')} \le \delta(\Phi) + \delta'(\Phi).$$

As in [1] and [3], the theorem follows by exhibiting a fork Φ for which $\delta(\Phi) + \delta'(\Phi) \le O(1/\sqrt{\log n})$. We do this by a simple averaging argument. We define a probability distribution over forks Φ and show that the expectation $\mathbf{E}[(\delta(\Phi))^2 + (\delta'(\Phi))^2] \le O(1/\log n)$. Hence, $\min(\|f(w) - f(w')\|/d_T(w, w')) \le O(1/\sqrt{\log n})$, as claimed.

As usual, we identify the vertices of T_n with binary strings of length $\leq n$. (The root is the empty string and the two children of vertex α are $\alpha 0$ and $\alpha 1$.) Let $\beta(j)$ denote the jth prefix of $\beta \in \{0, 1\}^n$, the jth node on the path from the root to the leaf β.

To select a fork randomly, independently choose β uniformly from $\{0, 1\}^n$ and (p, q) uniformly from Γ_n. Define the fork $\Phi = (\beta(p), \beta((p+q)/2), \beta(q), \beta'(q))$, where $\beta'(q)$ is obtained from $\beta(q)$ by complementing the bit indexed by $1 + (p+q)/2$. By symmetry, $\delta(\Phi)$ and $\delta'(\Phi)$ are identically distributed. For any $\alpha \in \{0, 1\}^n$, Lemma 1 with $x_i = f(\alpha(i))$ implies

$$\mathbf{E}[(\delta(\Phi))^2 \mid \beta = \alpha] \leq \frac{1}{|\Gamma_n|} \sum_{i=0}^{n-1} \|f(\alpha(i+1)) - f(\alpha(i))\|^2 \leq O\left(\frac{1}{\log n}\right).$$

The last inequality follows since f is nonexpansive and $|\Gamma_n| = \Omega(n \log n)$. Averaging over α gives $\mathbf{E}[(\delta(T))^2 + (\delta(T'))^2] = O(1/\log n)$, as required. $\square$

References

1. J. Bourgain, On Lipschitz embedding of finite metric spaces in Hilbert space, *Israel J. Math*. **52** (1985), 46–52.
2. N. Linial, E. London and Yu. Rabinovich, The geometry of graphs and some of its algorithmic applications, *Combinatorica*, **15** (1995), 215–245.
3. J. Matousek, On embedding trees into uniformly convex Banach spaces. *Israel J. Math*. **114** (1999), 221–237.

Received October 22, 2001. *Online publication July* 24, 2002.