SOME RESULTS ON METRIC TREES

ASUMAN GÜVEN AKSOY AND TIMUR OIKHBERG

Abstract. Using isometric embedding of metric trees into Banach spaces, this paper will investigate barycenters, type and cotype, and various measures of compactness of metric trees. A metric tree $(T,\, d)$ is a metric space such that between any two of its points there is an unique arc that is isometric to an interval in $\mathbb{R}$. We begin our investigation by examining isometric embeddings of metric trees into Banach spaces. We then investigate the possible images $x_0 = \pi((x_1 + \ldots + x_n)/n)$, where π is a contractive retraction from the ambient Banach space X onto T (such a π always exists), in order to understand the "metric" barycenter of a family of points $x_1, \ldots, x_n$ in a tree T. Further, we consider the metric properties of trees (such as their type and cotype). We identify various measures of compactness of metric trees (their covering numbers, ϵ-entropy and Kolmogorov widths) and the connections between them. Additionally, we prove that the limit of the sequence of the Kolmogorov widths of a metric tree is equal to its ball measure of non-compactness.

1. Introduction

The study of injective envelopes of metric spaces, also known as metric trees, (T-theory or $\mathbb{R}$-trees) began with J. Tits [46] in 1977 and since then, applications have been found within many fields of mathematics. For an overview of geometry, topology, and group theory applications, consult Bestvina [9]. For a complete discussion of these spaces and their relation to global NPC spaces we refer to [13]. Applications of metric trees in biology and medicine involve phylogenetic trees [43], and in computer science involve string matching [5]. Universal properties of "ℓ_1 trees" (this is a special class of separable metric trees) in the family of separable complete metrically convex metric spaces have been discovered, and used to investigate Lipschitz quotients of Banach spaces, in [29].

Since metric trees are described by three different names and several definitions, we start with the definition we will use.

Definition 1.1. Let $x, y \in M$, where $(M,\, d)$ is a metric space. A *geodesic segment from x to y*, is the image of an isometric embedding $\alpha : [a, b] \to M$ such that $\alpha(a) = x$ and $\alpha(b) = y$. The geodesic segment will be called a *metric segment* and denoted by $[x, y]$ throughout this paper.

Definition 1.2. A metric space (M, d), is a *metric tree* if and only if for all $x, y, z \in M$, the following holds:

(1) there exists a unique metric segment from x to y, and

(2) $[x,z] \cap [z,y] = \{z\} \Rightarrow [x,z] \cup [z,y] = [x,y]$.

Following are some examples of metric trees.

Example 1.3. (The Real Tree) Let $X_{\mathbb{R}}$ denote the set of all bounded subsets of $\mathbb{R}$ which contain their infimum. For all subsets x and y in $\mathbb{R}$, define a map $d : X_{\mathbb{R}} \times X_{\mathbb{R}} \to \mathbb{R}$ by

$$d(x,y) := 2\max\{\sup(x \triangle y), \inf x, \inf y\} - (\inf x + \inf y)$$

where by $x \triangle y$ we mean the symmetric difference of the sets x and y. Then d is a metric on $X_{\mathbb{R}}$, and $(X_{\mathbb{R}}, d)$ is a metric tree. For striking properties of this metric tree we refer to [21].

Example 1.4. (The Radial Metric, Spider Tree) Define $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_{\geqslant 0}$ by:

$$d(x,y) = \begin{cases} \|x-y\| & \text{if } x = \lambda\, y \text{ for some } \lambda \in \mathbb{R}, \\ \|x\| + \|y\| & \text{otherwise.} \end{cases}$$

We can observe the d is in fact a metric and that $(\mathbb{R}^2, d)$ is a metric tree.

Example 1.5. (The spider with n legs) In many of our examples, we shall consider a subtree of the radial tree described above. Fix $n \in \mathbb{N}$, and a sequence of positive numbers $(a_i)_{i=1}^n$, the *spider with n legs* is defined as a union of n intervals of lengths $a_1, \ldots, a_n$, emanating from the common center, and equipped with the radial metric. More precisely, our tree T consists of its center o, and the points (i,t), with $1 \leqslant i \leqslant n$ and $0 < t \leqslant a_i$. The distance d is defined by setting $d(o,(i,t)) = t$, and

$$d((i,t),(j,s)) = \left\{ \begin{array}{ll} |t-s| & i = j \\ t+s & i \neq j \end{array} \right. .$$

Abusing the notation slightly, we often identify o with $(i,0)$.

The simplest spider – that with three legs – is called a *tripod* (this terminology comes from [44]).

Example 1.6. (Non-simplicial Tree) In general metric trees are more complicated than metric graphs. Metric graphs are spaces obtained by taking connected graphs and metrizing nontrivial edges. Such a graph is a metric tree if the corresponding metric graph is connected and simply connected. For example, consider the set $[0,\infty) \times [0,\infty)$ with the distance $d : [0,\infty) \times [0,\infty) \to \mathbb{R}_{\geqslant 0}$ given by:

$$d(x,y) = \begin{cases} |x_1 - y_1| & \text{if } x_2 = y_2, \\ x_1 + y_1 + |x_2 - y_2| & \text{if } x_2 \neq y_2. \end{cases}$$

Set $X_n = (\mathbb{H}^n, \frac{1}{n}\, d)$, where $\mathbb{H}^n$ is a hyperbolic n-space. Then the ultraproduct $\prod X_n$ over some nontrivial ultrafilter $\mathcal{U}$ is the asymptotic cone $\mathbb{H}^n_{\mathcal{U}}$ of $\mathbb{H}^n$, an example of a non-simplicial tree. In this metric tree, the complement of every point has infinitely many connected components. For further discussion of this space and construction of metric trees related to the asymptotic geometry of hyperbolic metric spaces we refer to [13] and [22].

We refer the reader to [10] for the properties of metric segments and to [2], [3] and [24] for the basic properties of complete metric trees. Below we list some useful notation and results.

For x, y in a metric space M, we sometimes write $xy = d(x, y)$. For $x, y, z \in M$, we say y is *between x and z*, denoted xyz, if and only if $xz = xy + yz$. The following facts will be used throughout the paper:

(1) (Transitivity of betweenness [10]) Let M be a metric space and let $a, b, c, d \in M$. If abc and acd, then abd and bcd.

(2) (Three point property [2]) Let $x, y, z \in T$. There exists $w \in T$ such that

$$[x, z] \cap [y, z] = [w, z] \text{ and } [x, y] \cap [w, z] = \{w\}$$

Consequently,

$$[x, y] = [x, w] \cup [w, y] \ , [x, z] = [x, w] \cup [w, z], \text{ and } [y, z] = [y, w] \cup [w, z].$$

(3) (Uniform Convexity [2]) A metric tree T is uniformly convex. Although it is easy to prove uniform convexity by showing modulus of convexity $\delta_M(r, \epsilon) > 0$ for all $r, \epsilon > 0$,uniform convexity of metric trees also follows from CN-inequality for geodesic metric spaces [13]. Given points x, y_1, y_2 in a geodesic space where y_0 is the midpoint of the segment $[y_1, y_2]$, then CN-inequality implies that

$$d(x, y_0)^2 \leqslant \frac{1}{2} d(x, y_1)^2 + \frac{1}{2} d(x, y_2)^2 - \frac{1}{4} d(y_1, y_2)^2$$

In particular, $d(x, y_1) \leqslant R, \ \ d(x, y_2) \leqslant R$ and $d(y_1, y_2) \geqslant r$ imply

$$d(x, y_0) \leqslant (1 - \delta(\frac{r}{R}))R$$

where $\delta(\varepsilon) = \sqrt{1 - \frac{\varepsilon^2}{4}}$ and thus one has the usual Euclidean modulus of convexity.

(4) (Compactness, [2]) A metric tree T is compact if and only if

$$T = \bigcup_{f \in F} [a, f] \quad \text{for all } a \in T \text{ and } \overline{F} \text{ is compact}\,,$$

where F is the set of *final points of T (or leaves of T)*, defined as

$$F := \{f \in T \mid f \notin (x, y) \text{ for all } x, y \in T\}.$$

2. Hyperconvexity and Metric Trees

Definition 2.1. A metric space X is *hyperconvex* if

$$\bigcap_{i \in I} B_c(x_i; r_i) \neq \emptyset$$

for any collection $\{B_c(x_i; r_i)\}_{i \in I}$ of closed balls in X with $x_i x_j \leqslant r_i + r_j$.

The notion of a hyperconvex metric space was introduced by Aronszajn and Panitchpakdi [4]. They proved the following theorem, which is now well known.

Theorem 2.2 (Aronszajn and Panitchpakdi, [4]). *X is a hyperconvex metric space if and only if for all metric spaces D, if $C \subset D$ and $f : C \to X$ is a nonexpansive mapping, then f can be extended to the nonexpansive mapping $\tilde{f} : D \to X$.*

Hyperconvex spaces are complete and connected, the simplest example of hyperconvex space is the set of real numbers $\mathbb{R}$ or a finite dimensional real Banach space endowed with the maximum norm. While the Hilbert space ℓ^2 fails to be hyperconvex, the spaces ℓ^∞ and L^∞ are hyperconvex. The connection between hyperconvex metric spaces and metric trees is given in the following theorem:

Theorem 2.3 ([30], [1]). *A complete metric tree T is hyperconvex.*

3. Embeddings of Metric Trees into Banach spaces

Henceforth, we consider *isometric* embeddings of metric trees into Banach spaces. Note that there is a wealth of results concerning *Lipschitz* embeddings of trees (in a slightly different sense of the word "tree") into Banach spaces. In particular, the connections between Lipschitz embeddability of trees and superreflexivity were investigated in [11], and more recently, in [6]. The distortion necessary to embed a metric tree into a uniformly convex Banach space can be found in e.g. [37] (by [34], this problem is equivalent to computing the distortion of embedding the corresponding finite tree).

3.1. Embeddings into L_∞.

First we consider various embeddings into L_∞.

Theorem 3.1 (From [31], page 395). *Let X be a metric space and $a \in X$, then $J = J_a : X \to \ell^\infty(X) : x \mapsto (xm - am)_{m \in X}$ is an isometric embedding of X into $\ell^\infty(X)$.*

Proof. By the triangle inequality, we know that for all $m \in X$ that $xa \geqslant |xm - am|$. Therefore for all $x \in X$,

$$\|J(x)\|_\infty = \sup_{m \in X} |xm - am| \leqslant xa < \infty,$$

and hence in fact $J(x) \in \ell^\infty(X)$. It now suffices to show that $xy = \|J(x) - J(y)\|_\infty$ for all $x, y \in X$.

Let $x, y \in X$ and first observe,

$$\begin{aligned} \|J(x) - J(y)\|_\infty &= \|\{xm - am\}_m - \{ym - am\}_m\|_\infty \\ &= \sup_{m \in X} |xm - am - ym + am| \\ &= \sup_{m \in X} |xm - ym|. \end{aligned}$$

We can see that $xy \leqslant \sup_{m\in X}|xm - ym|$, by letting $m = y$. By the triangle inequality, $xy \geqslant |xm - ym|$. Therefore, $xy = \sup_{m\in X}|xm - ym| = \|J(x) - J(y)\|_\infty$. ■

The embedding J_a defined above is called *canonical*. When the space X is bounded, we can also use the embedding $J(x)(y) = d(x, y)$.

We can also embed a metric space X into a larger L_∞ space. To this end, pick $t_0 \in X$, and denote by $\mathcal{L}_{X,t_0} = \mathcal{L}$ the space of 1-Lipschitz functions from X to $\mathbb{R}$, vanishing at t_0. Define the *universal embedding* of X into $\ell_\infty(\mathcal{L})$ by setting, $U(t) = (f(t))_{f\in\mathcal{L}}$ for $t \in X$. Below we show that U is indeed an isometric embedding, satisfying a certain "universal projective" property.

Theorem 3.2. (1) *The map U described above is an isometry.*

(2) *For any 1-Lipschitz function $g : X \to \mathbb{R}$, there exists a 1-Lipschitz affine functional $\tilde{g} : \ell_\infty(\mathcal{L}) \to \mathbb{R}$, such that $g = \tilde{g} \circ U$.*

(3) *For any 1-Lipschitz function $g : X \to Z$, where Z is a λ-injective Banach space, there exists a λ-Lipschitz affine map $\tilde{g} : \ell_\infty(\mathcal{L}) \to Z$, such that $g = \tilde{g} \circ U$.*

Proof. (1) Fix $x, y \in X$, and show that $\|U(x) - U(y)\| = xy$. As any $f \in \mathcal{L}$ is 1-Lipschitz, the definition of U yields

$$\|U(x) - U(y)\| = \sup_{f\in\mathcal{L}} |f(x) - f(y)| \leqslant xy.$$

To prove the reverse inequality, consider the function $f_x : X \to \mathbb{R} : t \mapsto xt - xt_0$. Clearly, $f_x \in \mathcal{L}$, hence $\|U(x) - U(y)\| \geqslant |f_x(x) - f_x(y)| = xy$. Thus, U is an isometry.

(2) By translation, we can assume that $g(t_0) = 0$, hence $g \in \mathcal{L}$. Set $\tilde{g}(a) = a_g$ for $a = (a_f)_{f\in\mathcal{L}} \in \ell_\infty(\mathcal{L})$, Then $\tilde{g}$ is a contractive linear functional. Moreover, for any $x \in X$, $\tilde{g}(U(x)) = (U(x))_g = g(x)$, as desired.

(3) Fix an isometric embedding $I : Z \to \ell_\infty(\Gamma)$. Let $P : \ell_\infty(\Gamma) \to Z$ be a projection of norm not exceeding λ. We can view $I \circ g$ as a collection of maps $h_\gamma : X \to \mathbb{R}$ ($\gamma \in \Gamma$). By Part (2), each of them admits a 1-Lipschitz extension $\tilde{h}_\gamma$. This results in a 1-Lipschitz map $\tilde{h} = (\tilde{h}_\gamma) : \ell_\infty(\mathcal{L}) \to \ell_\infty(\Gamma)$, extending $I \circ g$. We complete the proof by setting $\tilde{g} = P \circ \tilde{h}$. ■

Note that the canonical embedding of T into $\ell_\infty(T)$ need not share this property of the universal embedding. Indeed, suppose $T = [0, 1]$. Consider the function $g : T \to \mathbb{R}$, defined by setting $g(0) = 0$, $g(1/n) = 0$, $g((2n + 1)/(2n(n + 1))) = 1/(2n(n + 1))$ ($n \in \mathbb{N}$), and letting g be piecewise-linear on $(0, 1]\backslash\{1/n : n \in \mathbb{N}\}$. We claim that there is no affine bounded map $\tilde{g} : \ell_\infty(T) \to \mathbb{R}$ such that $\tilde{g} \circ J = g$, where J is the canonical embedding. To show this, recall the definition of J (with $x^* = 0$): for $x, y \in T$,

$$h_x(y) = J(x)(y) = |x - y| - |y| = \begin{cases} x - 2y & y \leqslant x \\ -x & y \geqslant x \end{cases} .$$

For $n \in \mathbb{N}$, set $a_n = (1/n + 1/(n+1))/2 = (2n+1)/(2n(n+1))$, $b_n = (1/n - 1/(n+1))/2 = 1/(2n(n+1))$, $F_n = h_{1/n}$, and $G_n = h_{a_n}$. By definition, $g(F_n) = 0$, and $g(G_n) = b_n$. Furthermore, $h_0 = 0$, and $g(h_0) = 0$. Therefore, the extension $\tilde{g} : \ell_\infty(T) \to \mathbb{R}$, if it exists, is a linear functional.

Let $H_n = F_{n+1} + F_n - 2G_n$. A simple computation shows that

$$H_n(y) = \begin{cases} 2(1/n - y) & a_n \leqslant y \leqslant 1/n \\ 2(y - 1/(n+1)) & 1/(n+1) \leqslant y \leqslant a_n \\ 0 & \text{otherwise} \end{cases},$$

hence $\|H_n\| = \sup_y |H_n(y)| = 2b_n$. By linearity, $\tilde{g}(H_n) = -2b_n$.

For $N \in \mathbb{N}$, let $H = \sum_{n=1}^N (2b_n)^{-1} H_n$. Then $\|H\| = 1$, and $\tilde{g}(H) = -N$. As N is arbitrary, there is no $\tilde{g}$ with the desired properties.

Note also that there need not be a "injective" counterpart of the "projective" universal embedding U. More precisely, suppose T is the "tripod" tree T, described in Example 1.5. There is no isometric embedding $A : T \to X$ with the property that, for any isometric embedding $B : T \to Y$ (X and Y are Banach spaces), there exists a contractive affine map $V : Y \to X$ satisfying $V \circ B = A$. Indeed, suppose, for the sake of contradiction, that there exists an A with this property. Consider $B_1 : T \to \ell_\infty^2$, taking $(1,t)$ to $t(1,1)$, $(2,t)$ to $t(1,-1)$, and $(3,t)$ to $-t(1,-1)$. We can assume that $A(o) = 0$ (as before, o denotes the "root" of T). Suppose $V_1 \circ B_1 = A$, for some V_1. Then $A(2,1) = -A(3,1)$. Modifying B_1 to obtain the "right" B_2 and B_3, we show that $A(1,1) = -A(3,1)$, and $A(1,1) = -A(2,1)$. But these three equalities cannot hold simultaneously.

3.2. **Embeddings into L_1.** Next we define the "semicanonical" embedding of T into ℓ_1 (cf. Section 2.5 of [24]). We say that a complete metric tree is *finitely generated* if it is a finite union of closed metric segments. Equivalently, a complete metric tree is finitely generated if it is compact, and has finitely many leaves. Such a tree has finitely many *branching points* (points whose removal breaks T into more than two connected components). A segment I of such a T is called *elementary* if its endpoints are leaves or branching points, and it contains no branching points in its interior. Write T as above as a union of elementary segments I_i ($1 \leqslant i \leqslant N$). Changing the enumeration if necessary, we can assume that $\bigcup_{i=1}^k I_k$ is connected for any k. To construct an isometry $U : T \to \ell_1^N$, pick an arbitrary $f_1 \in \ell_1^N$, and define U on I_1 by setting

$$U(I_1) = \{te_1 + f_1 : 0 \leqslant t \leqslant |I_1|\}$$

(here $(e_k)_{k=1}^N$ is the canonical basis for ℓ_1^N). Now suppose U has already been defined on $\bigcup_{i=1}^k I_k$ to ℓ_1^N in such a way that $U(I_j) = \{te_j + f_j : 0 \leqslant t \leqslant |I_j|\}$, where $f_j - f_1 \in \mathrm{span}[e_1, \ldots, e_{j-1}]$. To describe $U(I_{k+1})$, note that exactly one of the endpoints (say a) of I_{k+1} belongs to $\bigcup_{i=1}^{k} I_k$. Let $f_{k+1} = U(a)$. Then $f_{k+1} = |I_k| e_k + f_k$, hence $f_{k+1} - f_1 \in \text{span}[e_1, \ldots, e_k]$. For $x \in I_{k+1}$ define $U(x) = f_{k+1} + d(x,a)e_{k+1}$. Then

$$U(I_{k+1}) = \{te_{k+1} + f_{k+1} : 0 \leqslant t \leqslant |I_{k+1}|\}.$$

Clearly, this U describes an isometric embedding.

Now suppose T is an arbitrary tree. We can represent it as a union of finitely generated trees, each of which embeds into ℓ_1^N for some N. An ultraproduct argument (cf. [26]) implies that T embeds into an L_1 space.

3.3. **Embeddings into other Banach spaces.** We have not been able to construct explicit embeddings of metric trees into other Banach spaces. However, we can show:

Theorem 3.3. *Suppose X is a non-superreflexive Banach space, T is a finitely generated metric tree, and $\epsilon > 0$. Then there exists a Banach space Y, $(1+\epsilon)$-isomorphic to X, such that T embeds into Y isometrically.*

Remark 3.4. The renorming of the space X is essential, at least when it is separable. Indeed, consider the "tripod" T, with three "legs" of length 1 (described in Example 1.5). It turns out that any separable Banach space Z has a renorming not containing an isometric copy of T. To see this, recall that, by Section 1 of [25], Z has a renorming – call it X – which is strictly convex: that is, the equality $2(\|x\|^2 + \|y\|^2) = \|x-y\|^2$ implies $x = -y$. Now suppose $U : T \to X$ is an isometric embedding. Without loss of generality, we can assume $U(o) = 0$. Let $x_i = U((i,1))$ $(i \in \{1,2,3\})$. Then $\|x_i\| = 1$ for each i, and $\|x_i - x_j\| = 2$ if $i \neq j$. By the strict convexity of X, $x_3 = -x_1 = x_2$, which is impossible.

Proof of Theorem 3.3. Denote by $\mathcal{E}$ the set of all branching points and leaves of T. $e_1, e_2 \in \mathcal{E}$ are called *adjacent* if $[e_1, e_2]$ is an elementary segment of T – that is, a segment not containing any branching points. Select $e_\emptyset \in \mathcal{E}$, and call it the *root*. Enumerate the immediate descendants of $e_\emptyset$ (that is, the members of $\mathcal{E}$ which are connected to $e_\emptyset$ by elementary segments) by $e_1, \ldots, e_{n_\emptyset}$. For $1 \leqslant i \leqslant n_\emptyset$, let $a_i = d(e_\emptyset, e_i)$. For each i, enumerate its own immediate descendants $e_{11}, \ldots, e_{1n_1}$, and set $a_{ij} = d(e_i, e_{ij})$. Proceeding further in the same manner, we write $\mathcal{E}$ as the collection of points e_S, for a finite collection $\mathcal{S}$ of finite strings S. Then $e_{S'}$ is a descendant of e_S if and only if $S' = S \smile j$, for some $j \in [1, n_S]$. Set $a_{S'} = d(e_S, e_{S'})$, where S is the unique immediate predecessor of S'.

For $S_1, S_2 \in \mathcal{S}$, write $S_1 \prec S_2$ if e_{S_1} is a predecessor of e_{S_2}, or equivalently, if $S_2 = S_1 \smile j_1 \smile \ldots \smile j_k$. For $S_\alpha = (i_{1\alpha} \ldots i_{k_\alpha \alpha})$ $(\alpha \in \{1,2\})$, set $S_1 \wedge S_2 = (i_{11} \ldots i_{k_0 1})$, where k_0 is the largest integer k with the property that $i_{k1} = i_{k2}$. If there are no such k, set $S_1 \wedge S_2 = \emptyset$. Then $e_{S_1 \wedge S_2}$ is the largest common predecessor of e_{S_1} and e_{S_2}.

It is easy to see that the distance d on the set $\mathcal{E}$ (inherited from the tree T) is given by the formula described below. For $S_\alpha = (i_{1\alpha} \dots i_{k_\alpha \alpha})$ ($\alpha \in \{1, 2\}$), let k_0 be the largest integer k such that $i_{k1} = i_{k2}$. Let $S = S_1 \wedge S_2 = (i_{11} \dots i_{k_0 1})$. Then

(3.1)
$$d(e_{S_1}, e_{S_2}) = d(e_{S_1}, e_S) + d(e_S, e_{S_2}) = \sum_{m=k_0+1}^{k_1} a_{(i_{11} \dots a_{m1})} + \sum_{m=k_0+1}^{k_2} a_{(i_{12} \dots a_{m2})}.$$

The main step is to renorm X (making it into Y) in such a way that there exists an isometry $J_\mathcal{E} : \mathcal{E} \to Y$. We then extend it to $J : T \to Y$ so that $J|_\mathcal{E} = J_\mathcal{E}$. For $t \in T \backslash \mathcal{E}$, there exist unique $e_1, e_2 \in \mathcal{E}$ such that t belongs to the elementary segment $[e_1, e_2]$. Let $\lambda = d(e_1, t)/d(e_1, e_2)$. Define $J(t) = \lambda J_\mathcal{E}(e_1) + (1 - \lambda) J_\mathcal{E}(e_2)$. Clearly, J is an isometry on any elementary segment. We claim that J is an isometry on T.

Indeed, consider $t_1, t_2 \in T$. For $\alpha \in \{1, 2\}$, there exists $S_\alpha \in \mathcal{E}$, for which $t_\alpha \in [e_{S_\alpha}, e_{S'_\alpha}]$, where S'_α is a successor of S_α (that is, $S'_\alpha = S_\alpha \smile j_\alpha$). Let $S = S_1 \wedge S_2$. We establish the desired equality $\|J(t_1) - J(t_2)\| = d(t_1, t_2)$ only in the case when S'_1 is not a successor of S'_2, and S'_2 is not a successor of S'_1 (other configurations are handled similarly). In this case, $[e_{S_1}, e_{S_2}] \subset [t_1, t_2] \subset [e_{S'_1}, e_{S'_2}]$. Note that

$$d(t_1, t_2) = d(t_1, e_{S_1}) + d(e_{S_1}, e_{S_2}) + d(t_2, e_{S_2}).$$

By the triangle inequality,

$$\begin{aligned} &\|J(t_1) - J(t_2)\| \\ &\leqslant \|J(e_{S_1}) - J(e_{S_2})\| + \|J(e_{S_1}) - J(t_1)\| + \|J(e_{S_2}) - J(t_2)\| \\ &= d(e_{S_1}, e_{S_2}) + d(t_1, e_{S_1}) + d(t_2, e_{S_2}) = d(t_1, t_2). \end{aligned}$$

On the other hand,

$$\begin{aligned} &\|J(t_1) - J(t_2)\| \\ &\geqslant \|J(e_{S'_1}) - J(e_{S'_2})\| - \|J(e_{S'_1}) - J(t_1)\| - \|J(e_{S'_2}) - J(t_2)\| \\ &= d(e_{S'_1}, e_{S_2}) - d(t_1, e_{S'_1}) - d(t_2, e_{S'_2}) = d(t_1, t_2). \end{aligned}$$

Together, these two inequalities yield the desired result.

So, it remains to exhibit the isometry $J_\mathcal{E} : \mathcal{E} \to Y$. As before, denote the cardinality of $\mathcal{E}$ by N. By [7] (p. 270), there exists, for every $M \in \mathbb{N}$, a family $(x_{iM})_{i=1}^N \subset B(0; 1)$ such that

$$\| \sum_{i=1}^N \alpha_i x_{iM} \| \geqslant (1 - 2^{-M}) \sum_{i=1}^N |\alpha_i| \tag{3.2}$$

for any sequence (α_i) with at most one change of signs. Introduce the lexicographic order $\ll$ on $\mathcal{S}$ as follows: if $S_1 \prec S_2$, then $S_1 \ll S_2$. Otherwise, let $S = S_1 \wedge S_2$, and write $S_\alpha = S \smile j_{1\alpha} \smile \dots \smile j_{m_\alpha \alpha})$ ($\alpha = 1, 2$). We say $S_1 \ll S_2$ if $j_{11} \leqslant j_{12}$.

Let $\phi : \mathcal{S} \to \{1, \ldots, N\}$ be the monotone increasing bijection with respect to the lexicographic order. Define $J_M : \mathcal{E} \to X$ by setting

$$J_M(e_S) = \sum_{R \prec S} a_R x_{\phi(e_R)M}.$$

More precisely, if $S = (i_1, \ldots i_k)$, then

$$J_M(e_S) = \sum_{j=1}^{k} d(e_{i_1 \ldots e_{j-1}}, e_{i_1 \ldots e_j}) x_{\phi(i_1 \ldots i_j)M}.$$

By the (3.1) and (3.2),

$$d(e_{S_1}, e_{S_2}) \geqslant \|J_M(e_{S_1}) - J_M(e_{S_2})\| \geqslant (1 - 2^{-M}) d(e_{S_1}, e_{S_2}).$$

Passing to a subsequence if necessary, we can assume that

$$\Phi(\alpha_1, \ldots, \alpha_N) = \lim_M \| \sum_i \alpha_i x_{iM} \|$$

exists for every sequence $(\alpha_i)_{i=1}^N$. Then $\Phi : \mathbb{R}^N \to \mathbb{R}$ is a quasinorm, and $\Phi(\alpha_1, \ldots, \alpha_N) = \sum_i |\alpha_i|$ if the sequence (α_i) changes sign at most once.

Let $\mathcal{I}$ be the largest subset of $\{1, \ldots, N\}$ such that $\Phi(\alpha_1, \ldots, \alpha_N) > 0$ whenever $\alpha_i = 0$ for any $i \notin \mathcal{I}$, and $\alpha_i \neq 0$ for some $i \in \mathcal{I}$ (such a $\mathcal{I}$ need not be unique, of course). Define a Banach space F with the basis $(f_i)_{i \in \mathcal{I}}$ by setting $\| \sum_{i \in \mathcal{I}} \alpha_i f_i \| = \Phi((\alpha_i)_{i \in \mathcal{I}})$. For $i \notin \mathcal{I}$, we can define $f_i \in F$ by setting $f_i = \sum_{j \in \mathcal{I}} \beta_{ij} f_j$, with the appropriate (β_{ij}); that is, we must have $\Phi((\alpha_k)) = 0$, with

$$\alpha_k = \begin{cases} -1 & k = i \\ \beta_{ik} & k \in \mathcal{I} \\ 0 & \text{otherwise} \end{cases}.$$

As $\lim_M \|x_{iM} - \sum_{j \in \mathcal{I}} \beta_{ij} x_{jM}\| = 0$, we conclude that, for the sequence (f_i) defined above,

$$\Phi(\alpha_1, \ldots, \alpha_N) = \| \sum_i \alpha_i f_i \|.$$

Furthermore, consider the map $J'_M : \mathcal{E} \to X$ by setting

$$J'_M(e_S) = \sum_{R \prec S} a_R x'_{\phi(e_R)M}.$$

Here, $x'_{iM} = x_{iM}$ if $i \in \mathcal{I}$, and $x'_{iM} = \sum_{j \in \mathcal{I}} \beta_{ij} x_{jM}$ otherwise. By the above,

$$\lim_M \|J'_M(e_{S_1}) - J'_M(e_{S_2})\| = d(e_{S_1}, e_{S_2}).$$

Therefore,

$$J' : \mathcal{E} \to F : e_S \mapsto \sum_{R \prec S} a_R f_{\phi(e_R)M}$$

is an isometry.

It remains to renorm X in the appropriate way. For a fixed $\epsilon > 0$, find M so large that

$$cd(e_{S_1}, e_{S_2}) \geqslant \|J'_M(e_{S_1}) - J'_M(e_{S_2})\| \geqslant c^{-1} d(e_{S_1}, e_{S_2}),$$

where $c^2 < 1 + \epsilon$. Consider the map

$$U : \text{span}[x_{iM} : i \in \mathcal{I}] \to F : x'_{iM} \to f_i.$$

For a sufficiently large M, $\|U\|, \|U^{-1}\| < c$. Embed F into ℓ_∞, and extend U to $\tilde{U} : X \to \ell_\infty$, with the norm $\|\tilde{U}\| < c$. Introduce an equivalent norm $\|\cdot\|_Y$ on X by setting $\|x\|_Y = \max\{c^{-1}\|x\|, \|\tilde{U}x\|\}$. Clearly, $c\|x\| \geqslant \|x\|_Y \geqslant c^{-1}\|x\|$, and $J = id \circ J'_M : \mathcal{E} \to Y$ is an isometry (here, $id : X \to Y$ is the formal identity). ■

4. Barycenters of trees

There have been numerous attempts to find an appropriate "non-linear" notion of the barycenter of a set (or of a measure) in a metric space. Several possible definitions are discussed in [44]. In this section, we approach this problem for metric trees, using their injectivity. More precisely: suppose U is an isometric embedding of a metric tree T into a Banach space X, equipped with the norm $\|\cdot\|$. Suppose $x_1, \ldots, x_n \in T$, and let $\tilde{x}_0 = (x_1 + \ldots + x_n)/n$ be their barycenter in X (we identify $x \in T$ with $U(x) \in X$). Let $\mathbf{P} = \mathbf{P}_{U,T,X}$ be the set of contractive retractions π from X onto $U(T)$ (it is non-empty since T is injective). We try to describe $\mathbf{P}(\tilde{x}_0)$. More generally, suppose $\alpha = (\alpha_i)_{i=1}^n$ is a sequence of positive numbers, with $\sum_{k=1}^n \alpha_k = 1$. Set $\tilde{x}^{(\alpha)} = \sum_{k=1}^n \alpha_k x_k$, and try to describe $\mathbf{P}(\tilde{x}^{(\alpha)})$.

Proposition 4.1. *Suppose T is a complete metric tree, embedded isometrically into a normed space X. For $x_0 \in T$ and $\tilde{x} \in X$, the following are equivalent:*

(1) $x_0 \in \mathbf{P}(\tilde{x})$.
(2) *For any $x \in T$, $d(x_0, x) \leqslant \|\tilde{x} - x\|$.*

If, in addition, T is compact, then the two statements above are equivalent to:

(3) *For any leaf (final point) $y \in T$, $d(x_0, y) \leqslant \|\tilde{x} - y\|$.*

In the proofs below, we sometimes identify T with its image in the ambient Banach space, and $d(\cdot, \cdot)$ with $\| \cdot - \cdot \|$.

Proof. By the injectivity of T, (1) holds if and only if there exists a contractive map

$$\pi : T \cup \{\tilde{x}\} \to T \text{ such that } \pi|_T = I_T, \text{ and } \pi(\tilde{x}) = x_0.$$

This, in turn, is equivalent to (2). Clearly, (2) implies (3). To show that, for a compact T, the converse is true, recall the "Krein-Milman Theorem for metric trees" (Statement (4) in the end of Section 1, proved in [2]), which asserts that $T = \bigcup_{y \in \mathcal{L}} [x_0, y]$, where $\mathcal{L}$ is the set of leaves of T. For $x \in T$, find $y \in \mathcal{L}$ such that $x \in [x_0, y]$. If $\|x_0 - y\| \leqslant \|\tilde{x} - y\|$, then

$$\|x_0 - x\| = \|x_0 - y\| - \|x - y\| \leqslant \|\tilde{x} - y\| - \|x - y\| \leqslant \|\tilde{x} - x\|,$$

thus (3) implies (2). ∎

Corollary 4.2. *If then $x_0 \in \mathbf{P}(\tilde{x}^{(\alpha)})$, then $d(x_0, x) \leqslant \sum_k \alpha_k d(x_k, x)$ for any $x \in T$.*

Proof. By the Proposition 4.1(2)

$$\|x_0 - x\| \leqslant \Big\| \sum_k \alpha_k x_k - x \Big\| = \Big\| \sum_k \alpha_k (x_k - x) \Big\| \leqslant \sum_k \alpha_k \|x_k - x\|$$

for any $x \in T$, whenever $\pi(\tilde{x}^{(\alpha)}) = x_0$. ∎

In certain cases, the converse to this corollary is also true: this is shown by the following two theorems. However, in general, the converse implication does not hold (Example 4.10).

Theorem 4.3. *Suppose T is a complete metric tree, embedded into $\ell_\infty(T)$ in the canonical way. For $x_0 \in T$, the following are equivalent:*

1. $x_0 \in \mathbf{P}(\tilde{x}^{(\alpha)})$.
2. $d(x_0, x) \leqslant \sum_k \alpha_k d(x_k, x)$ *for any* $x \in T$.

Proof. The implication (1) ⇒ (2) follows from Corollary 4.2. We establish the converse. Recall that the canonical embedding takes $x \in T$ to $h(x) \in \ell_\infty(T)$, where $h(x)(y) = d(x, y) - d(x^*, y)$. Suppose $d(x_0, x) \leqslant \sum_k \alpha_k d(x_k, x)$ for any $x \in T$. By Proposition 4.1, we have to show that $d(x_0, x) \leqslant \|\tilde{x}^{(\alpha)} - h(x)\|$ for any $x \in T$. We identify $\tilde{x}^{(\alpha)}$ with the function $\phi : T \to \mathbb{R}$, defined by

$$\phi(y) = \sum_k \alpha_k h(x_k)(y) = \sum_k \alpha_k \|x_k - y\| - \|x^* - y\|.$$

Then

$$\|\tilde{x}^{(\alpha)} - h(x)\| = \sup_{y \in T} |\phi(y) - h(x)(y)| = \sup_{y \in T} \Big| \sum_k \alpha_k (\|x_k - y\| - \|x - y\|) \Big|$$
$$\geqslant \Big| \sum_k \alpha_k (\|x_k - x\| - \|x - x\|) \Big| = \sum_k \alpha_k \|x_k - x\| \geqslant \|x_0 - x\|.$$

∎

Theorem 4.4. *Suppose T is a finitely generated tree, embedded into ℓ_1^N in the semicanonical way. For $x_0 \in T$, the following are equivalent:*

1. $x_0 \in \mathbf{P}(\tilde{x}^{(\alpha)})$.
2. $d(x_0, x) \leqslant \sum_k \alpha_k \|x_k - x\|$ *for any* $x \in T$.

Proof. By Proposition 4.1, we have to show that, if $x_0 \in T$ is such that $d(x_0, x) \leqslant \sum_k \alpha_k d(x_k, x)$ for any $x \in T$, then $\|x_0 - x\| \leqslant \|\tilde{x}^{(\alpha)} - x\|$ for any leaf $x \in T$. By translation, we can identify x_0 with $0 \in \ell_1^N$. For each k, let $u_k \in T$ satisfy $[x_0, x_k] \cap [x_0, x] = [x_0, u_k]$. Let $y_k = x_k - u_k$ and $z_k = x - u_k$. By the construction of the semicanonical embedding of T into ℓ_1^N, y_k and z_k have disjoint support, and therefore,

$$\|x_k - x\| = \|y_k - z_k\| = \|y_k\| + \|z_k\| \tag{4.1}$$

Furthermore, $\|x_0 - x\| = \|u_k\| + \|z_k\|$. Thus,

$$\sum_k \alpha_k \|x_k - x\| = \sum_k \alpha_k (\|y_k\| + \|z_k\|) \geqslant \sum_k \alpha_k \|x_0 - x\| = \sum_k \alpha_k (\|u_k\| + \|z_k\|),$$

which is equivalent to

$$\sum_k \alpha_k \|y_k\| \geqslant \sum_k \alpha_k \|u_k\|. \tag{4.2}$$

We have to show that

$$\|\tilde{x}^{(\alpha)} - x\| = \|\sum_k \alpha_k (x_k - x)\| \geqslant \sum_k \alpha_k (\|u_k\| + \|z_k\|) = \|x\|.$$

In view of (4.2) and (4.1), it is enough to prove that, for any leaf $x \in T$,

$$\|\sum_k \alpha_k (x_k - x)\| = \sum_k \alpha_k \|x_k - x\|. \tag{4.3}$$

Thus, it suffices to establish that, for any $1 \leqslant j \leqslant N$, the sign of $\langle e_j^*, x_k - x \rangle$ is independent of k (here, $e_j^* \in \ell_1^{N*}$ is defined by setting $\langle e_j^*, \sum_s \gamma_s e_s \rangle = \gamma_j$). If $0 = x_0$ is not in the interior of I_j, then, by changing the orientation of the j-th coordinate axis, we can assume that, for any $y \in I_j$, $\langle e_j^*, y \rangle = \|y - a_j\|$, where a_j is the endpoint of I_j which is nearest to x_0. Moreover, $\langle e_j^*, y \rangle = 0$ if $[x_0, y] \cap I_j = \emptyset$, and $\langle e_j^*, y \rangle = |I_j|$ if $[x_0, y] \cap I_j = I_j$.

Now suppose $0 = x_0$ belongs to the interior of I_j. By relabeling, we can assume that $j = 1$, and $I_1 = \{te_1 : \alpha \leqslant t \leqslant \beta\}$, where $\alpha < 0 < \beta$, and $\beta e_1 \in [x_0, x]$. If $y \notin I_1$, then $\langle e_j^*, y \rangle = \alpha$ if $[y, x_0] \cap [x_0, x] = \{x_0\}$, and $\langle e_j^*, y \rangle = \alpha$ if $[y, x_0] \cap [x_0, x] \supset [x_0, \beta e_1]$.

Now consider $\langle e_j^*, x_k - x \rangle$. If $j = 1$, then $\langle e_1^*, x \rangle = \beta$, and $\langle e_1^*, x_k \rangle \leqslant \beta$, hence $\langle e_1^*, x_k - x \rangle \leqslant 0$. For $j > 1$, $\langle e_j^*, x \rangle = |I_j|$ if $I_j \subset [x_0, x]$, and $\langle e_j^*, x \rangle = 0$ otherwise. Moreover, $0 \leqslant \langle e_j^*, y \rangle \leqslant |I_j|$ for any $y \in T$. Thus, for a fixed j, the quantities $\langle e_j^*, x_k - x \rangle$ are either all non-positive, or all non-negative. This establishes (4.3). ■

Remark 4.5. The sets $\{x_0 \in T : d(x_0, x) \leqslant \sum_k \alpha_k d(x_k, x) \,\forall\, x \in T\}$ were briefly discussed in Remark 7.2(iii) of [44]. Namely, consider the probability measure $q = \sum_k \alpha_k \delta_{x_k}$. The set of points described above was denoted by $C^*(q)$.

As shown by the following example, this set need not be contained in the metric or linear convex hull of $x_1, \dots, x_n$. Recall that $S \subset T$ is *metric convex* if S contains the metric segment connecting any two of its points. The *metric convex hull* of S (the smallest metric convex set containing S) is denoted by $\mathrm{con}S$.

Example 4.6. As an example, consider the points $x_i = (i, 1)$ $(1 \leqslant i \leqslant 3)$ in a spider with four legs (defined in Example 1.5). If T is embedded into $\ell_\infty(T)$ (ℓ_1^4) in the canonical (respectively semicanonical) way, then $\mathbf{P}(\tilde{x}_0)$ consists of o, as well as of all (j, t) with $1 \leqslant j \leqslant 4$ and $t \leqslant 1/3$. In particular, $(4, 1/3)$ or rather, its canonical or semicanonical image belongs to neither the metric nor linear convex hull of $\{x_1, x_2, x_3\}$.

Certain information about $\mathbf{P}(\tilde{x}_0)$ may be extracted from the following results.

Proposition 4.7. *Suppose a complete metric tree T is embedded isometrically into a normed space X, and $\tilde{x}$ is a point of X. Then $\mathbf{P}(\tilde{x})$ is a closed, metric convex subset of T.*

Proof. Proposition 4.1 implies that $x_0 \in \mathbf{P}(\tilde{x})$ if and only if $d(x, x_0) \leqslant \|x - \tilde{x}\|$ for any $x \in T$. This implies that $\mathbf{P}(\tilde{x})$ is closed. Furthermore, suppose $x_1, x_2, \in \mathbf{P}(\tilde{x})$, and $x_0 \in [x_1, x_2]$. Then, by Section 2 of [44],

$$d(x, x_0) \leqslant \max\{d(x, x_1), d(x, x_2)\} \leqslant \|x - \tilde{x}\|$$

for any $x \in T$, which implies $x \in \mathbf{P}(\tilde{x})$. ∎

In certain cases, when the structure of $x_1, \dots, x_n$ in T is simple, we can describe $\mathbf{P}(\tilde{x}_0)$ explicitly. For instance, if $n = 2$, then $\mathbf{P}(\tilde{x}_0) = \{x_0\}$, where $x_0 \in [x_1, x_2]$ satisfies $d(x_1, x_0) = d(x_1, x_2)/2$ (equivalently, $d(x_2, x_0) = d(x_1, x_2)/2$). Indeed,

$$d(x_0, x_1) = \|(x_1 + x_2)/2 - x_1\| = d(x_1, x_2)/2,$$

and similarly, $d(x_2, x_0) = d(x_1, x_2)/2$. If $x_0 \notin [x_1, x_2]$, then there exists $y \in [x_1, x_2]$ such that $[x_0, x_s] = [x_0, y] \cup [y, x_s]$ for $s = 1, 2$. Then $d(x_0, x_1) + d(x_0, x_2) > d(x_1, x_2)$, which contradicts the contractiveness of the map taking $\tilde{x}_0$ to x_0. Thus, x_0 is the unique point of $[x_1, x_2]$ satisfying $d(x_1, x_0) = d(x_1, x_2)/2$.

In a more complex situation, consider the "tripod" T, with limbs of length 1 (described in Example 1.5). For $0 \leqslant \alpha \leqslant \beta \leqslant 1$, we define $(i, [\alpha, \beta]) = \{(i, t) : \alpha \leqslant t \leqslant \beta\}$.

Theorem 4.8. *Consider the points $x_i = (i, 1)$ $(i = 1, 2, 3)$ in the tripod T described above. Suppose S is a subset of T. Then there exists an embedding of T into a Banach space X such that $S = \mathbf{P}(\tilde{x}_0)$ if and only if there exist $i_0 \in \{1, 2, 3\}$ and $0 \leqslant \alpha \leqslant \beta \leqslant 1/3$, such that either (i) $S = (i_0, [\alpha, \beta])$, or (ii) $S = (i_0, [0, \beta]) \cup \left(\cup_{i \neq i_0} (i, [0, \alpha]) \right)$.*

Proof. First suppose T is embedded in a normed space X, and show that $\mathbf{P}(\tilde{x}_0)$ is of the form described in the theorem. For $1 \leqslant i \leqslant 3$, let $d_i = \|x_i - \tilde{x}_0\|$. By relabeling, we can assume that $d_1 \leqslant d_2 \leqslant d_3$. Note that $d_3 \leqslant 4/3$. Indeed,

$$\begin{aligned} d_3 &= \Big\| x_3 - \frac{x_1 + x_2 + x_3}{3} \Big\| \\ &= \frac{1}{3} \|(x_3 - x_1) + (x_3 - x_2)\| \leqslant \frac{1}{3}\big(d(x_3, x_1) + d(x_3, x_2)\big) = \frac{4}{3}. \end{aligned}$$

Furthermore, $d_1 + d_2 \geqslant d(x_1, x_2) = 2$, hence in particular, $d_1 \geqslant 2/3$, and $d_2 \geqslant 1$. Let $\beta = d_2 - 1$, and $\alpha = |d_1 - 1|$.

By Proposition 4.1, $x_0 = (i, t) \in T$ belongs to $\mathbf{P}(\tilde{x}_0)$ if and only if $d(x_i, x_0) \leqslant d_i$ for each i. Thus, $x_0 = (1, t) \in \mathbf{P}(\tilde{x}_0)$ if and only if two conditions are satisfied:

(1) $d(x_1, x_0) = 1 - t \leqslant d_1$, or in other words, $t \geqslant 1 - d_1$, which translates to $t \geqslant \alpha$ or $t \geqslant 0$, depending on whether $1 - d_1$ is positive or negative.
(2) $d(x_2, x_0) = 1 + t \leqslant d_2$, or in other words, $t \leqslant d_2 - 1 = \beta$.

The third condition, $d(x_3, x_0) = 1 + t \leqslant d_3$, is subsumed in the second one, as $d_3 \geqslant d_2$. Thus, $(1, t) \in \mathbf{P}(\tilde{x}_0)$ if and only if $1 - \beta \leqslant t \leqslant 1 - \alpha$.

Similarly, $x_0 = (2, t) \in \mathbf{P}(\tilde{x}_0)$ if and only if two conditions are satisfied:

(1) $d(x_1, x_0) = 1 + t \leqslant d_1$, or in other words, $t \leqslant d_1 - 1$, which means either $t \leqslant \alpha$ ($d_1 \geqslant 1$), or there are no suitable t's ($d_1 < 1$).
(2) $d(x_2, x_0) = 1 - t \leqslant d_2$, which is always true, since $d_2 \geqslant 1$.

Thus, the set of t for which $(2, t) \in \mathbf{P}(\tilde{x}_0)$ is either $[0, \alpha]$, or $\emptyset$. The set $\{t : (3, t) \in \mathbf{P}(\tilde{x}_0)\}$ is described in a similar fashion.

Next we construct an embedding of T into a Banach space X, for which $\mathbf{P}(\tilde{x}_0) = S$. Suppose first $0 \leqslant \alpha \leqslant \beta < 1/3$, and construct an embedding of T into $L_1(0, 2)$ with the property that $\mathbf{P}(\tilde{x}_0) = (1, [\alpha, \beta])$. Let $c = 1 - 3\beta$, and $a = (1 - 3\alpha)/c$. Define the functions f_1, f_2, f_3 as follows:

$$f_1(u) = \begin{cases} 1 & 0 \leqslant u \leqslant 1 \\ 0 & 1 < u \leqslant 2 \end{cases}, \quad f_2(u) = \begin{cases} -a & 0 \leqslant u \leqslant c/2 \\ 0 & c/2 < u \leqslant 1 \\ 1 - ac/2 & 1 < u \leqslant 2 \end{cases},$$

$$f_3(u) = \begin{cases} 0 & 0 \leqslant u \leqslant 1 - c/2 \\ -a & 1 - c/2 < u \leqslant 1 \\ -(1 - ac/2) & 1 < u \leqslant 2 \end{cases}.$$

Note that $f_i f_j \leqslant 0$ for $i \neq j$, hence $\|t f_i - s f_j\| = t + s$ for positive t and s. Therefore, the mapping $(i, t) \mapsto t f_i$ describes an embedding of T into $L_1(0, 2)$.

The barycenter $\tilde{x}_0$ corresponds to the function g, given by

$$g(u) = \begin{cases} -(a - 1)/3 & u \in [0, c/2] \cup (1 - c/2, 1] \\ 1/3 & c/2 < u \leqslant 1 - c/2 \\ 0 & 1 < u \leqslant 2 \end{cases}.$$

Then

$$\|f_1 - g\| = \Big(1 + \frac{a-1}{3}\Big)c + \frac{2}{3}(1-c) = 1 - \frac{1-ac}{3} = 1 - \alpha,$$

and

$$\|f_2 - g\| = \|f_3 - g\| = \frac{c}{2}\Big(a - \frac{a-1}{3}\Big) + \frac{1}{3}(1-c) + \frac{c}{2}\cdot\frac{a-1}{3} + \Big(1 - \frac{ac}{2}\Big)$$
$$= 1 + \frac{1-c}{3} = 1 + \beta.$$

By Proposition 4.1(3), $\mathbf{P}(\tilde{x}_0)$ consists of all points $x_0 \in T$ such that $d(x_i, x_0) \leqslant \|x_i - \tilde{x}_0\|$ for $i \in \{1,2,3\}$; that is, of all the points $(1,t)$ with $\alpha \leqslant t \leqslant \beta$.

To obtain $S = (i_0, [0,\beta]) \cup \big(\cup_{i \neq i_0} (i,[0,\alpha])\big)$ we modify the above construction, by setting $c = 1 - 3\beta$, and $a = (1+3\alpha)/c$. Then $\|f_2 - g\| = \|f_3 - g\| = 1 + \beta$, and $\|f_1 - g\| = 1 + (ac-1)/3 = 1 + \alpha$.

A modification of this construction works in the "limit" case of $\beta = 1/3$. In this case embed T into $M([0,2])$ (the space of regular Radon measures on $[0,2]$). As before, let $\mu_1 = f_1 = \chi_{(0,1)}$, and set

$$\mu_2 = a\delta_0 + (1-a)\delta_2, \;\; \mu_2 = a\delta_1 - (1-a)\delta_2,$$

with $a \in [0,1]$ to be determined later (here, δ_x is the Dirac measure supported by x). Once again, it is easy to check that the map $(i,t) \mapsto t\mu_i$ defines an embedding of T to $M([0,2])$. The barycenter $\tilde{x}_0$ corresponds to the measure

$$\nu = \frac{1}{3}\Big(a(\delta_0 + \delta_1) + \chi_{(0,1)}\Big)$$

hence

$$\|\mu_1 - \nu\| = \frac{2}{3}(1+a) = 1 - \frac{1-2a}{3}, \;\text{ and }\; \|\mu_2 - \nu\| = \|\mu_3 - \nu\| = \frac{4}{3}.$$

To obtain $\mathbf{P}(\tilde{x}_0) = (1, [\alpha, 1/3])$, set $a = (1-3\alpha)/2$ (then $(1-2a)/3 = \alpha$). To end up with $S = (1,[0,1/3]) \cup \big(\cup_{i \neq 1} (i,[0,\alpha])\big)$, set $a = (3\alpha+1)/2$. ■

Proposition 4.9. *Suppose a metric tree T is embedded isometrically into $L_1(\mu)$, $x_1, \ldots, x_n$ are points of T, and $\alpha_1, \ldots, \alpha_n$ are positive numbers, satisfying $\sum_k \alpha_k = 1$. If $x_0 \in T$ belongs to $\mathbf{P}(\tilde{x}^{(\alpha)})$, then the unique point of* $\mathrm{con}(x_1, \ldots, x_n)$*, nearest to x_0, also belongs to $\mathbf{P}(\tilde{x}^{(\alpha)})$.*

Proof. Let $S = \mathrm{con}(x_1, \ldots, x_n)$. Suppose $x_0 \in \mathbf{P}(\tilde{x}^{(\alpha)})$, or equivalently (Proposition 4.1), $\|y - x_0\| \leqslant \|y - \tilde{x}^{(\alpha)}\|$ for any $y \in T$. Only the case of $x_0 \notin S$ needs to be studied. Pick $x \in S$, and let x' be the point of $[x_0, x]$ with the property that $d(x_0, x') = \inf\{d(x_0, y) : y \in [x_0, x] \cap S\}$. In other words, x' is the point of $[x_0, x] \cap S$, farthest from x. The set S is closed, hence $x' \in S$.

We claim that, for any $u \in S$, $x' \in [x_0, u]$, and consequently, $d(x_0, u) = d(x_0, x') + d(x', u)$. Indeed, suppose, for the sake of contradiction, $x' \notin [x_0, u]$. Then there exists $z \in [x_0, x'] \backslash \{x'\}$ such that $[x', u] = [x', z] \cup [z, u]$. By convexity, $z \in S$, which is impossible, by the definition of x'.

Next, we show that $x' \in \mathbf{P}(\tilde{x}^{(\alpha)})$. By Proposition 4.1, it suffices to show that, for any $y \in T$, $\|y - x'\| \leqslant \|y - \tilde{x}^{(\alpha)}\|$. We consider two cases:

(1) $[x', y] \cap S$ is strictly larger than $\{x'\}$. As S is closed and metric convex, $[x', y] \cap S = [x', z]$, for some z. We know that $[x_0, z] = [x_0, x'] \cup [x', z]$, hence $[x_0, y] = [x_0, x'] \cup [x', y]$. Then $d(x', y) = d(x_0, y) - d(x_0, x')$, and therefore,

$$d(x', y) \leqslant d(x_0, y) \leqslant \|\tilde{x}^{(\alpha)} - y\|.$$

(2) $[x', y] \cap S = \{x'\}$. In this case, note first that, for any $u \in S$, $x' \in [y, u]$, and consequently, $d(y, u) = d(y, x') + d(x', u)$. Indeed, if $x' \notin [y, u]$, then there exists $z \in [y, x'] \backslash \{x'\}$ such that $[x', u] = [x', z] \cup [z, u]$. Then $z \in S$, which contradicts our assumptions about y.

Now recall that the ambient space is $L_1(\Omega, \mu)$. We can assume that $x' = 0$. Then, for any $u \in S$, $\|y - u\| = \|y\| + \|u\|$, hence $yu \leqslant 0$ μ-a.e. (we view y and u as functions on Ω). As $x_1, \ldots, x_n \in S$, we also have $\tilde{x}^{(\alpha)} y \leqslant 0$ μ-a.e.. Therefore, $\|y - \tilde{x}^{(\alpha)}\| \geqslant \|y\| = d(x', y)$, which is what we need. ■

Example 4.10. The above result doesn't hold for embeddings into arbitrary spaces. Consider the "spider" $T = \{(i, t) : i \in \{1, 2, 3, 4\}, 0 \leqslant t \leqslant 1\}$, as in Example 4.6. Embed T into ℓ_∞^3 by setting $(1, t) \mapsto (-e_1 + e_2 + e_3)t$, $(2, t) \mapsto (e_1 - e_2 + e_3)t$, $(3, t) \mapsto (e_1 + e_2 - e_3)t$, and $(4, t) \mapsto (e_1 + e_2 + e_3)t$, ($e_1, e_2, e_3$ denote the canonical basis in ℓ_∞^3). For $i = 1, 2, 3$, let $x_i = (i, 1)$. Then the "linear" barycenter of $\{x_1, x_2, x_3\}$ is $\tilde{x}_0 = (e_1 + e_2 + e_3)/3$. As this point lies on the image of T in ℓ_∞^3, $\mathbf{P}(\tilde{x}_0) = \{(4, 1/3)\}$. This example also shows that the converse to Corollary 4.2 doesn't hold.

Finally, we present an example suggesting that nothing non-trivial can be said about the distance from the "linear" barycenter of a tree to the tree itself.

Example 4.11. Consider a "spider" T with n limbs of length 1, that is, the set of points (i, t), with $1 \leqslant i \leqslant n$ and $0 \leqslant t \leqslant 1$, with the usual radial metric. For $1 \leqslant i \leqslant n$ let $x_i = (i, 1)$. Then there exists an embedding of T into $L_1(1, n+1)$ such that $\|x - \tilde{x}_0\| \geqslant 1$ for any $x \in T$. Indeed, the embedding taking (i, t) to $\chi_{(i, i+t)}$ has the desired properties.

5. Type and cotype of Metric Trees

In this section we consider the metric type and cotype of metric spaces. The notion of metric type was introduced in [12], see also [41]. More recently, metric cotype was defined in [38].

We say that a metric space (X, d) satisfies *the four-point inequality* if, for any x_1, x_2, x_3, x_4 in X,

$$d(x_1, x_2) + d(x_3, x_4) \leqslant \max\{d(x_1, x_3) + d(x_2, x_4), d(x_1, x_4) + d(x_2, x_3)\}. \tag{5.1}$$

Moreover X is said to satisfy *Reshetnyak's inequality* if, for any $x_1, x_2, x_3, x_4 \in X$,

$$d(x_1, x_2)^2 + d(x_3, x_4)^2 \leqslant d(x_1, x_3)^2 + d(x_2, x_4)^2 + d(x_1, x_4)^2 + d(x_2, x_3)^2. \tag{5.2}$$

Lemma 5.1. *The four-point inequality implies Reshetnyak's inequality.*

Proof. Suppose x_1, x_2, x_3, x_4 satisfy

$$d(x_1, x_2) + d(x_3, x_4) \leqslant \max\{d(x_1, x_3) + d(x_2, x_4), d(x_1, x_4) + d(x_2, x_3)\},$$

and show that

$$d(x_1, x_2)^2 + d(x_3, x_4)^2 \leqslant d(x_1, x_3)^2 + d(x_2, x_4)^2 + d(x_1, x_4)^2 + d(x_2, x_3)^2.$$

By scaling and relabeling, we can assume that

$$d(x_1, x_2) + d(x_3, x_4) = 1 \leqslant d(x_1, x_3) + d(x_2, x_4).$$

Let $a = d(x_1, x_2)$, $b = d(x_1, x_3)$. Then $d(x_3, x_4) = 1 - a$, $d(x_2, x_4) = 1 - b$, and furthermore,

$$d(x_1, x_4) \geqslant |d(x_1, x_3) - d(x_3, x_4)| = |a + b - 1|, \text{ and}$$
$$d(x_2, x_3) \geqslant |d(x_1, x_2) - d(x_1, x_3)| = |a - b|.$$

Thus, it suffices to show that, for any $a \in [0, 1]$ and $b \geqslant 0$,

$$a^2 + (1 - a)^2 \leqslant b^2 + (1 - b)^2 + (a + b - 1)^2 + (a - b)^2.$$

The last inequality is easily verified. ■

Remark 5.2. By Chapter 3 of [24], any metric tree satisfies (5.2). By [44], a more general result is true: any global NPC space satisfies (5.1).

Definition 5.3. A metric space (X, d) is said to have *metric*(or *BMW*) *type* p $(1 \leqslant p \leqslant 2)$ with constant K if, for any $n \in \mathbb{N}$, and any function $f : \{-1, 1\}^n \to X$, we have

$$\sum_{\epsilon \in \{-1,1\}^n} d(f(\epsilon), f(-\epsilon))^2 \leqslant K n^{1/p-1/2} \sum_{\epsilon \in \{-1,1\}^n} \sum_{i=1}^{n} d(f(\epsilon), f(\epsilon^{[i]}))^2, \tag{5.3}$$

where $(\epsilon_1, \ldots, \epsilon_n)^{[i]} = (\epsilon, \ldots, \epsilon_{i-1}, -\epsilon_i, \epsilon_{i+1}, \ldots, \epsilon_n)$. On an intuitive level, we can think of the points $f(\epsilon)$ as vertices of a "cube." Then the left hand side of (5.3) is the sum of the squares of the "diagonals" of this cube, while the right hand side involves its "edges."

We do not quote the definition of metric cotype, due to space constraints. Instead, we refer the reader to [38].

Theorem 5.4.
(1) *Any metric space satisfying the four-point inequality has metric type 2, with constant* 1*. In particular, this result holds for metric trees.*
(2) *Any complete metric tree has metric cotype 2, with a universal constant.*

Proof. We handle the cotype first. Recall that any L_1 space has cotype 2, with the constant $\sqrt{2}$ (this classical fact can be seen, for instance, by combining the Khintchine constant from [45] with the basic properties of cotype, described in e.g. [28]). Therefore, by Theorem 1.2 of [38], L_1 has metric cotype 2, with the constant $90\sqrt{2}$. We have seen that any finitely generated metric tree embeds isometrically into ℓ_1^N, for some N. As the cotype passes to subspaces, any finitely generated tree must have metric cotype 2, with constant $90\sqrt{2}$. Finally, metric cotype is a "local" property, hence any complete metric tree must possess it.

To tackle the type, we need to show that, for any $f : \{-1,1\}^n \to X$, we have

$$\sum_{\epsilon \in \{-1,1\}^n} d(f(\epsilon), f(-\epsilon))^2 \leqslant \sum_{\epsilon \in \{-1,1\}^n} \sum_{i=1}^n d(f(\epsilon), f(\epsilon^{[i]}))^2. \tag{5.4}$$

We shall prove this by induction on n. Indeed, for $n = 1$ we trivially have an equality, while the case for $n = 2$ is the Reshetnyak inequality. Suppose (5.4) has been established for n, and prove it for $n+1$. Consider $f : \{-1,1\}^{n+1} \to X$. By the induction hypothesis, for any $\epsilon \in \{-1,1\}^n$,

$$\sum_{\epsilon \in \{-1,1\}^n} d(f(\epsilon,1), f(-\epsilon,1))^2 \leqslant \sum_{\epsilon \in \{-1,1\}^n} \sum_{i=1}^n d(f(\epsilon,1), f(\epsilon^{[i]},1))^2,$$

and

$$\sum_{\epsilon \in \{-1,1\}^n} d(f(\epsilon,-1), f(-\epsilon,-1))^2 \leqslant \sum_{\epsilon \in \{-1,1\}^n} \sum_{i=1}^n d(f(\epsilon,-1), f(\epsilon^{[i]},-1))^2,$$

hence

$$\begin{aligned} &\sum_{\epsilon \in \{-1,1\}^n} \big(d(f(\epsilon,1), f(-\epsilon,1))^2 + d(f(\epsilon,-1), f(-\epsilon,-1))^2 \big) \\ &\leqslant \sum_{\epsilon \in \{-1,1\}^n} \sum_{i=1}^n \big(d(f(\epsilon,1), f(\epsilon^{[i]},1))^2 + d(f(\epsilon,-1), f(\epsilon^{[i]},-1))^2 \big). \end{aligned} \tag{5.5}$$

Applying Reshetnyak inequality to the 4-tuple $(f(\epsilon,1),\ f(-\epsilon,1),\ f(\epsilon,-1),\ f(-\epsilon,-1))$ $(\epsilon \in \{-1,1\}^n)$, we see that

$$\begin{aligned} &d(f(\epsilon,1), f(-\epsilon,-1))^2 + d(f(\epsilon,-1), f(-\epsilon,1))^2 \\ &\leqslant d(f(\epsilon,1), f(-\epsilon,1))^2 + d(f(\epsilon,-1), f(-\epsilon,-1))^2 \\ &+ d(f(\epsilon,1), f(\epsilon,-1))^2 + d(f(-\epsilon,1), f(-\epsilon,-1))^2. \end{aligned}$$

Combining this inequality with (5.5), we obtain

$$\begin{aligned}
&\sum_{\delta\in\{-1,1\}^{n+1}} d(f(\delta), f(-\delta))^2 \\
&= \sum_{\epsilon\in\{-1,1\}^n} \big(d(f(\epsilon,1), f(-\epsilon,-1))^2 + d(f(\epsilon,-1), f(-\epsilon,1))^2\big) \\
&\leqslant \sum_{\epsilon\in\{-1,1\}^n} \big(d(f(\epsilon,1), f(-\epsilon,1))^2 + d(f(\epsilon,-1), f(-\epsilon,-1))^2\big) \\
&+ \sum_{\epsilon\in\{-1,1\}^n} \big(d(f(\epsilon,1), f(\epsilon,-1))^2 + d(f(-\epsilon,1), f(-\epsilon,-1))^2\big) \\
&\leqslant \sum_{\epsilon\in\{-1,1\}^n} \sum_{i=1}^{n} \big(d(f(\epsilon,1), f(\epsilon^{[i]},1))^2 + d(f(\epsilon,-1), f(\epsilon^{[i]},-1))^2\big) \\
&+ \sum_{\epsilon\in\{-1,1\}^n} \big(d(f(\epsilon,1), f(\epsilon,-1))^2 + d(f(-\epsilon,1), f(-\epsilon,-1))^2\big) \\
&= \sum_{\delta\in\{-1,1\}^{n+1}} \sum_{i=1}^{n+1} d(f(\delta), f(\delta^{[i]}))^2,
\end{aligned}$$

which is what we need. ∎

We next tackle the negative type of metric trees. Recall that a metric space X has *negative type* p ($p > 0$) if, for any $x_1, \dots, x_n \in X$, the $n \times n$ matrix $(d(x_i, x_j)^p)$ is conditionally negative definite. Here, we say that a Hermitian matrix $A = (a_{ij})_{i,j=1}^n$ is conditionally negative definite if $\sum_{i,j=1}^n a_{ij}\xi_i\overline{\xi_j} \leqslant 0$ whenever the vector $\xi = (\xi_i)_{i=1}^n$ satisfies $\sum_i \xi_i = 0$. The notion of p-negative type is equivalent to p-roundness, see e.g. [20, 35]. Negative type is strongly related to positive definiteness of kernels, and to embeddability into L_p-spaces (see e.g. Section 8.1 of [8]).

It was shown in [27] that any metric tree has negative type 1. Therefore, it has negative type p for any $p \in (0, 1]$. We shall show that a metric tree need not have negative type p for $p > 1$. More precisely, consider the "spider" T_n, consisting of a central point, and n limbs of length 1.

Proposition 5.5. *If $p > 1$, then T_n fails to have negative type p for n large enough.*

Proof. Suppose $n > c/(c-2)$, where $c = 2^p$. Consider the subset of T_n, consisting of the central point t_0, and the endpoints $t_1, \dots, t_n$. The corresponding $(n+1) \times (n+1)$ distance matrix is

$$C = \begin{pmatrix} 0 & 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & c & c & \dots & c \\ 1 & c & 0 & c & \dots & c \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & c & c & c & \dots & 0 \end{pmatrix}.$$

We shall show the existence of $\xi = (\xi_0, \xi_1, \ldots, \xi_n)$ such that $\xi_1 + \ldots + \xi_n = -\xi_0$, and $\langle C\xi, \xi\rangle > 0$. Note that, for ξ as above, $D\xi = 0$, where the D is a matrix of all whose entries equal 1. Let

$$A = -\frac{1}{c-1}(C - cD) = \begin{pmatrix} a & 1 & 1 & 1 & \ldots & 1 \\ 1 & a & 0 & 0 & \ldots & 0 \\ 1 & 0 & a & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 0 & 0 & 0 & \ldots & a \end{pmatrix},$$

where $a = c/(c-1) < 2$. It suffices to find $\xi = (\xi_0, \xi_1, \ldots, \xi_n)$ such that $\xi_1 + \ldots + \xi_n = -\xi_0$, and $\langle A\xi, \xi\rangle < 0$.

An induction argument yields the determinant of this $(n+1) \times (n+1)$ matrix: $\det A = a^{n+1} - na^{n-1}$. Thus, A has $n-1$ eigenvalues equal to 0, as well as non-zero eigenvalues $\lambda_1 = a - \sqrt{n}$ and $\lambda_2 = a + \sqrt{n}$. The corresponding normalized eigenvectors are $\eta^1 = (-\sqrt{n}, 1, \ldots, 1)/\sqrt{2n}$ and $\eta^2 = (\sqrt{n}, 1, \ldots, 1)/\sqrt{2n}$. Then, for any $\xi \in \ell_2^n$,

$$\langle A\xi, \xi\rangle = \lambda_1 |\langle \xi, \eta^1\rangle|^2 + \lambda_2 |\langle \xi, \eta^2\rangle|^2.$$

Now consider $\eta = (\sqrt{2n}, -\sqrt{2n}, 0, \ldots, 0)$. Then $\langle \xi, \eta^1\rangle = -\sqrt{n} - 1$, and $\langle \xi, \eta^1\rangle = \sqrt{n} - 1$. Therefore,

$$\langle A\xi, \xi\rangle = (a - \sqrt{n})(\sqrt{n} + 1)^2 + (a + \sqrt{n})(\sqrt{n} - 1)^2 = 2a(n+1) - 4n,$$

which is negative, by our choice of n. ∎

Finally, we note that all metric trees have Markov type 2 [39].

6. Entropy Quantities and Other Measures of Compactness

6.1. **ϵ-entropy and related quantities.** Kolmogorov introduced the notion of ϵ-entropy as a measure of the massiveness of sets [32]. This notion has been useful in function spaces (see [23]), especially with asymptotic distribution of eigenvalues of elliptic operators, or as a way of measuring the sizes of spaces of solutions to PDE's [16]. Recently entropy and n-widths has been utilized as a measure of efficiency in the task of data compression (see [19], [42], [18]). In this section we examine the notion of entropy and its connection to the fact that complete metric trees are centered (Theorem 6.2). This proves the useful fact that ϵ-entropy of a bounded subset A is equal to ϵ-entropy of A relative to (T, d). We also connect the covering numbers of a compact subset of a tree with these of its convex hull (see e.g. [14] for some Banach space results in the same vein).

Definition 6.1. Let M be a metric space and $A \subset M$

- $\{x_i\}_{i \in I} \subset M$ is an *ϵ-net for A in M* if

$$A \subset \bigcup_{i \in I} B(x_i; \epsilon)$$

- $\{U_\alpha\}_{\alpha\in I}$, where $U_\alpha \subset M$, is an *ϵ-cover for A* if $\operatorname{diam}(U_\alpha) \leqslant 2\epsilon$ and
$$A \subset \bigcup_{\alpha\in I} U_\alpha.$$
- A is *centered* if for all $U \subset A$ such that $\operatorname{diam}(U) = 2r$, there exists $a \in A$ such that $U \subset B_c(a;r)$. By $B_c(a;r)$ we mean the closed ball centered at a radius r.
- $U \subset A$ is a *ϵ-separated subset of A* if

$$\epsilon \leqslant x_i x_j \quad \text{for all } i,j \in I \text{ with } i \neq j,\ x_i \neq x_j \in U$$

Let $\mathcal{N}_\epsilon(A)$ be the cardinality of a minimal ϵ-cover of A. Then define the *ϵ-entropy of A* as
$$\mathcal{H}_\epsilon(A) := \log_2 \mathcal{N}_\epsilon(A).$$

Similarly, let $\mathcal{K}_\epsilon^M(A)$ be the cardinality of a minimal ϵ-net for A in M. Then define the *ϵ-entropy of A relative to M* as
$$\mathcal{E}_\epsilon^M(A) := \log_2 \mathcal{K}_\epsilon^M(A).$$

Furthermore if we set $\mathcal{M}_\epsilon(A)$ as the maximal number of elements in ϵ-separated set $U \subset A$, we define *ϵ-capacity of A* as
$$\mathcal{C}_\epsilon(A) := \log_2 \mathcal{M}_\epsilon(A).$$

Theorem 6.2. *Every complete metric tree T is centered.*

Very few spaces are centered. A typical example of a space which is not centered is $\mathbb{R}^2$. This can be seen if one tries to locate a center for an equilateral triangle of side length $2r$ so that its distance to all points is at most r.

For the proof we require a lemma.

Lemma 6.3. *Let A be a subset of metric tree T with* $\operatorname{diam}(A) = 2r$. *Then for all $\epsilon > 0$ exists $m \in con(A)$ such that $A \subset B(m; r+\epsilon)$.*

Proof. For all $\epsilon > 0$, there exists $x, y \in A$ such that $xy > 2r - 2\epsilon$ and let m be the midpoint of $[x,y]$.

Let $z \in A$, then by the three-point property of metric trees, there exists $w \in [x,y]$ such that $[z,x] \cap [z,y] = [z,w]$. Without loss of generality we can assume that $m \in [w,x]$ and hence $w \in [z,x]$ by transitivity. Next $\operatorname{diam}(A) = 2r$ and $w \in [z,x]$ imply that
$$2r \geqslant zx = zm + mx = zm + \frac{xy}{2} > zm + (r-\epsilon),$$
which implies $r + \epsilon > zm$. Therefore $A \subset B(m; r+\epsilon)$. ■

Proof of 6.2. Let U be a bounded subset of a metric tree T, and let $\operatorname{diam}(U) = 2r$. For all $n \in \mathbb{N}$, there exists $x_n, y_n \in U$ such that $x_n y_n > 2(r - n^{-1})$. Let $z_n \in T$ be the midpoint of $[x_n, y_n]$, and we claim $\{z_n\}$ is a Cauchy sequence.

Let $0 < 2N^{-1} < \epsilon$ with $N \in \mathbb{N}$, and then let $n, m \geqslant N$. Let $u \in [x_n, y_n]$ be such that $[z_m, u] = [z_m, x_n] \cap [z_m, y_n]$, by swapping x_n and y_n we can claim without loss of generality that $u \in [z_n, y_n]$. Therefore $z_n \in [x_n, z_m]$.

Since metric segments are closed under intersections, $z_m \in [x_m, y_m]$, and z_m is an end point of $[x_n, z_m]$, we have that $[x_n, z_m] \cap [x_m, y_m] = [z_m, v]$ where $v \in [x_m, y_m]$. Hence by switching x_m and y_m we can claim without loss of generality that $v \in [x_m, z_m]$. Therefore $z_m \in [x_n, y_m]$.

Since $\operatorname{diam}(U) = 2r$, $x_n, y_m \in U$, $z_n \in [x_n, z_m]$ and $z_m \in [x_n, y_m]$, we have

$$\begin{aligned} 2r &\geqslant x_n y_m = x_n z_n + z_n z_m + z_m y_m \\ &> (r - n^{-1}) + z_n z_m + (r - m^{-1}) \geqslant 2r - 2N^{-1} + z_n z_m. \end{aligned}$$

Therefore, $z_n z_m < 2N^{-1} < \epsilon$ and $\{z_n\}$ is Cauchy. M is complete so let $\lim_n z_n = z$.

Suppose that there exists $u \in U$ such that $zu > r + 2\epsilon$ for some $\epsilon > 0$. Since $\lim_n z_n = z$, we can find a n such that $n^{-1} < \epsilon$ and $zz_n < \epsilon$. Furthermore, by the proof of Lemma 6.3 we know that $z_n u < r + n^{-1} < r + \epsilon$. Hence by the triangle inequality, $zu \leqslant zz_n + z_n u < r + 2\epsilon$, which contradicts that $zu > r + 2\epsilon$. Hence $U \in B_c(z; r)$ and therefore T is centered. ■

Remark 6.4. The above theorem can be proved using the definition of hyperconvexity ; and since a complete metric tree T is hyperconvex, T is centered follows easily. However our proof relies only on the properties of the metric segments and thus sheds more light on the local property of trees.

Theorem 6.5. *If A is a subset of a complete metric tree T, then $\mathcal{E}^T_\epsilon(A) = \mathcal{H}_\epsilon(A)$.*

Proof. It follows from the fact that every complete metric tree is centered and this implies that an ϵ-net for A is equivalent to an ϵ-cover. ■

Remark 6.6. Clearly $\mathcal{N}_\epsilon(A) = \mathcal{K}^M_\epsilon(A) \leqslant \mathcal{M}_\epsilon(A)$ implies that $\mathcal{E}^T_\epsilon(A) = \mathcal{H}_\epsilon(A) \leqslant \mathcal{C}_\epsilon(A)$. Thus if one fixes a "special" subset $A \subset T$ and wants to quantify the "massiveness" of A, one needs to find a lower bound for $\mathcal{H}_\epsilon(A) = \mathcal{H}^T_\epsilon(A)$ and an upper bound for $\mathcal{C}_\epsilon(A)$ which are sufficiently close to each other.

Next we connect the covering numbers $\mathcal{N}_\epsilon(S)$ of a compact subset S of a tree T with those of its convex hull (see e.g. [14] for some Banach space results).

Theorem 6.7. *Suppose S is a compact subset of a complete metric tree T, and ϵ_1, ϵ_2 are positive numbers. Then*

$$\mathcal{N}_{\epsilon_1+\epsilon_2}(\mathrm{con}(S)) \leqslant \mathcal{N}_{\epsilon_1}(S) \lceil \operatorname{diam} S/(4\epsilon_2) \rceil.$$

Proof. For the sake of brevity, set $N = N_{\epsilon_1}(S)$, $D = \operatorname{diam}(S)$, and $S' = \mathrm{con}(S)$. Convexity of the norm (see [44]) implies that the diameter of S'

equals D. By Theorem 6.2, there exists $x_0 \in S'$ such that for any $y \in S'$, $d(x_0, y) \leqslant \operatorname{diam}(S')/2 = D/2$, and moreover, $S' = \bigcup_{x \in S} [x_0, x]$.

Find $x_1, \ldots, x_N \in T$ such that for any $x \in S$ there exists i with the property that $d(x, x_i) \leqslant \epsilon_1$. Let x_i' be the point of S' which is closest of x_i. Then $d(x_i', y) \leqslant d(x_i, y)$ for any $y \in S'$. Indeed, there exists $z \in [x_i', y]$ such that $[x_i, x_i'] = [x_i, z] \cup [z, x_i']$. By convexity, $z \in S'$, hence $z = x_i'$, which is what we need.

Now let $K = \lceil D/(4\epsilon_2) \rceil$. For each i, find the points $(y_{ij})_{j=1}^K$ on $[x_0, x_i']$ in such a way that $d(x_0, y_{i1}) \leqslant \epsilon_2$, $d(x_i', y_{iK}) \leqslant \epsilon_2$, and $d(y_{ij}, y_{i,j+1}) \leqslant 2\epsilon_2$ for $1 \leqslant j \leqslant K$. In total, we have NK points y_{ij}. It remains to show that, for any $y \in S$, $d(y, y_{ij}) \leqslant \epsilon_1 + \epsilon_2$ for some (i, j).

As we have observed, there exists $x \in S$ such that $y \in [x_0, x]$. Find i such that $d(x, x_i') \leqslant \epsilon_1$. By Corollary 2.5 of [44], there exists $z \in [x_0, x_i']$ such that $d(z, y) \leqslant \epsilon_1$. Furthermore, there exists j such that $d(z, y_{ij}) \leqslant \epsilon_2$. By the triangle inequality,
$d(y, y_{ij}) \leqslant \epsilon_1 + \epsilon_2$. ■

6.2. **Kolmogorov numbers.** Kolmogorov introduced the notion of diameters (or widths) to generalize many of our intuitive ideas about "nice" normed linear spaces. Since then, Kolmogorov diameters have been widely used in approximation theory (see [40] and references therein). On the other hand, the notion of the measure of non-compactness of a subset of a metric space was introduced by Kuratowski [33] as a way to generalize Cantor's intersection theorem. In 1955, Darbo [17] applied measures of non-compactness to prove a powerful fixed point theorem. Since then measures of non-compactness have been a standard notion in fixed point theory. In the following, we define these two concepts and then show that even if Kolmogorov diameters (or n-widths) depends on n-dimensional subspaces, one can still define a n-width of a set in a metric tree and obtain a relationship between the measure of non-compactness of a subset and its n-widths.

Definition 6.8. Given a normed linear space X and subset A, define the *nth Kolmogorov diameter (n-width) of A in X* as:

$$\begin{aligned} &\delta_n(A, X) = \delta_n(A) \\ &:= \inf \left\{ \sup_{a \in A} d(a, M_n) \mid M_n \text{ is a } n\text{-dimensional subspace of } X \right\}. \end{aligned}$$

Observe that $\{\delta_n(A)\}_{n=1}^{\infty}$ forms a monotone decreasing sequence of non-negative numbers, provided for some n that $\delta_n(A) < \infty$.

Given metric space M and bounded subset A, define the *ball (Hausdorff) measure of non-compactness* as

$$\beta(A) := \inf \left\{ b > 0 \mid A \subset \bigcup_{j=1}^{n} B(x_j; b) \text{ for some } x_j \in M \right\}.$$

Using Theorem 3.1, we can extend the definition of Kolmogorov diameters to arbitrary metric spaces.

Definition 6.9. Let J be an embedding of a metric tree T into a normed space X, and $A \subset T$. Define the *nth Kolmogorov diameter of A in $\ell^\infty(T)$* as:

$$\delta_n^{(J)}(A) := \delta_n(J(A), \ell^\infty(T))$$

where $J : T \to \ell^\infty(T)$ is the isometric embedding from Theorem 3.1.

When there is no confusion about J, we simply write $\delta_n(A)$ instead of $\delta_n^{(J)}(A)$.

Theorem 6.10. *If T is a complete metric tree, embedded isometrically into a Banach space X, and $A \subset T$ is bounded, then*

$$\lim_{n\to\infty} \delta_n(A) = \beta(A).$$

Proof of $\beta(A) \leqslant \lim_{n\to\infty} \delta_n(A)$. Suppose $c > b > \lim_n \delta_n(A)$, which exists because $\{\delta_n(A)\}_{n=1}^\infty$ forms a monotone decreasing sequence of nonnegative numbers, and the fact that $\operatorname{diam}(A) < \infty$. Then by the monotonicity of $\{\delta_n(A)\}$, there exists n such that $\delta_n(A) < b$. This means there there exists an n-dimensional subspace T_n of X such that $\sup_{a\in A} d(J(a), T_n) < b$. Hence for all $x \in J(A)$, there exists $y \in M_n$, such that $\|x - y\|_\infty < b$.

Define $Q = B(J(A); b) \cap T_n$. For each $x \in J(A)$, by the above there exists $y \in T_n$ with $\|x - y\|_\infty < b$. So $y \in B(x; b) \subset B(J(A); b)$, and hence $y \in Q$. So $x \in B(y, b) \subset B(Q; b)$, and therefore $J(A) \subset B(Q; b)$.

Now Q is totally bounded, for it is a bounded subset of a finite dimensional subspace of $\ell^\infty(T)$. So there exists a finite $(c - b)$-net, $\{q_n\}$, for Q. Then since $J(A) \subset B(Q; b)$, we know that $\{q_n\}$ is a finite c-net for $J(A)$. Using the fact that P is nonexpansive and $P \circ J = \mathrm{id}_T$, we see that $\{P(q_n)\}$ is a finite c-net for A.

Hence $\beta(A) \leqslant c$, and therefore, $\beta(A) \leqslant \lim_n \delta_n(A)$. ■

Proof of $\beta(A) \geqslant \lim_{n\to\infty} \delta_n(A)$. Let $b > \beta(A)$. Then we have a finite b-net $\{x_j\}_{j=1}^n \subset T$ for A. Let $T_n = \operatorname{span}(\{J(x_j)\}_{j=1}^n)$, which will have dimension no greater than n. For all $x \in J(A)$, there exists $J(x_j) \in T_n$ such that $\|x - J(x_j)\|_\infty < b$. Hence we have that

$$\sup_{x\in J(A)} d(x, T_n) \leqslant b,$$

and therefore $\delta_n(A) = \delta_n(J(A), \ell^\infty(T)) \leqslant b$. Thus, $\beta(A) \geqslant \lim_{n\to\infty} \delta_n(A)$. ■

Corollary 6.11. *Suppose A is a bounded subset of a complete metric tree T. Then $\lim_{n\to\infty} \delta_n(A) = \alpha(A)/2$, where*

$$\alpha(A) := \inf\{b > 0 : A \subset \bigcup_{j=1}^k T_j \textit{ for some } T_j \subset A \textit{ diam}(T_j) \leqslant b\}$$

is the set measure of noncompactness of A.

Proof. Follows from the Theorem 6.2 that $\alpha(A) = 2\beta(A)$. ■

The relationship between geometric quantities, such as n-widths and entropy, and analytic entities of bounded linear maps, such as eigenvalues and essential spectrum, has been studied by various authors (see [15] and references therein).

As, in general, metric spaces need not be represented as centrally symmetric subsets of an ambient Banach space (such as L_∞), we need a modified notion of Kolmogorov numbers. Emulating Definition 6.8, we introduce:

Definition 6.12. The *n-th affine Kolmogorov diameter of A in X* is defined as:

$$\delta_n^{(a)}(A, X) = \delta_n^{(a)}(A)$$
$$:= \inf\left\{\sup_{a\in A} d(a, M) \mid M \text{ is an affine subspace of } X,\, \dim M \leqslant n\right\}.$$

Clearly, the sequence $(\delta_n^{(a)}(A))_{n=1}^\infty$ is non-increasing, and

$$\delta_n(A) \geqslant \delta_n^{(a)}(A) \geqslant \delta_{n+1}(\mathrm{conv}\,(A \cup (-A))).$$

To establish the last inequality, suppose $\delta_n^{(a)}(A) < c$. Then there exists an affine subspace M, of dimension not exceeding n, such that for any $a \in A$ there exists $m \in M$ satisfying $\|a-m\| < c$. Any element of the closed convex hull of $A\cup(-A)$ can be approximated arbitrarily well by $x = \sum_{i=1}^N \alpha_i a_i$, with $\sum_i |\alpha_i| \leqslant 1$, and $a_i \in A$. For each i, find $m_i \in M$ such that $\|a_i - m_i\| < c$. Note that $M' = \mathrm{conv}\,(M \cup (-M))$ is a linear subspace of dimension not exceeding $n + 1$, with $m = \sum_i \alpha_i m_i \in M'$. Thus, $\|x - m\| < c$.

Furthermore, if A is centrally symmetric about 0 in X, then $\delta_n(A) = \delta_n^{(a)}(A)$. Indeed, fix $\epsilon > 0$, and find an affine subspace $M \subset X$ of dimension less than n, such that for any $a \in A$ there exists $m \in M$ with the property that $\|a - m\| < \delta_n^{(a)}(A) + \epsilon$. By symmetry, for such an a we can also find $m_- \in -M$ satisfying $\|a-m_-\| < \delta_n^{(a)}(A)+\epsilon$. Note that $m' = (m+m_-)/2 \in M' = M + (-M)$, and the latter is a linear subspace of X, of the same dimension as M. By the triangle inequality, $\|a - m'\| < \delta_n^{(a)}(A) + \epsilon$. Thus, $\delta_n(A) \leqslant \delta_n^{(a)}(A) + \epsilon$. As $\epsilon > 0$ is arbitrary, we are done.

Clearly, for a metric space A, the behavior of the sequence $(\delta_n^{(a)}(A))_{n=1}^\infty$ may depend on the choice of the embedding into the ambient Banach space X (see Remark 6.15 for an example). For a subset A of a complete metric space X, we set $d_n(A) = \delta_n^{(a)}(U(A))$, where $U : T \to \ell_\infty(\mathcal{L})$ is the universal embedding. By Theorem 3.2, $\delta_n^{(a)}(J(A)) \leqslant d_n(A)$ for any isometric embedding $J : X \to Z$ (Z is a 1-injective Banach space).

To estimate $d_n(T)$, recall, from Definition 6.1, that $\mathcal{M}_\epsilon(T)$ denotes the cardinality of the largest ϵ-separated subset of T. Furthermore, $\mathcal{K}_\epsilon^M(T)$

stands for the cardinality of a minimal ϵ-net in the ambient space M, covering T. As complete metric trees are injective, the choice of M is irrelevant; we can take $M = T$, and use the notation $\mathcal{K}_\epsilon(T)$.

Proposition 6.13. *Suppose T is a complete metric tree. Let*

$$c_1 = \inf\{\epsilon > 0 \mid \mathcal{K}_\epsilon(T) \leqslant n\}, \quad \textit{and } c_2 = \sup\{\epsilon > 0 \mid \mathcal{M}_\epsilon(T) \geqslant n+2\}.$$

(1) *Suppose X is a Banach space, and $U : T \to X$ is a 1-Lipschitz map. Then $\delta_n^{(a)}(U(T)), X) \leqslant c_1$. Thus, $d_n(T) \leqslant c_1$.*
(2) *$d_n(T) \geqslant c_2/2$.*

Proof. (1) Fix $c > c_1$, and suppose $t_1, \ldots, t_K$ is a c-net in T. Let E be the affine span of $U(t_1), \ldots, U(t_K)$. Then $\dim E < K$. Furthermore, for any $t \in T$,

$$d(U(t), E) \leqslant \min_i d(U(t), U(t_i)) = \min_i d(t, t_i) \leqslant c,$$

which shows that $d_n(T) \leqslant c$. Taking the infimum, we conclude $d_n(T) \leqslant c_1$.

(2) Let $M = \mathcal{M}_\epsilon(T)$. For $c < c_2$, let $t_1, \ldots, t_M$ be a c-separated subset of T. $d(t_0, t_i) > c/2$ for any i. Indeed, by relabeling if necessary, we can assume that $d(t_1, t_2) \leqslant d(t_i, t_j)$ whenever i and j are different. Let t_0 be the midpoint of $[t_1, t_2]$. We claim that $d(t_0, t_i) > c/2$ for any i. The inequality clearly holds for $i \in \{1, 2\}$. If $i > 2$ and $d(t_0, t_i) \leqslant c/2$, then

$$d(t_1, t_i) \leqslant d(t_1, t_0) + d(t_0, t_i) \leqslant \frac{d(t_1, t_2)}{2} + \frac{c}{2} < d(t_1, t_2),$$

a contradiction.

Let $\mathcal{L}$ be the set of 1-Lipschitz functions from T to $\mathbb{R}$, taking t_0 to 0. For any $\sigma = (\sigma_1, \ldots, \sigma_M) \in \{-1, 1\}^M$, define $g_\sigma : \{t_0, t_1, \ldots, t_M\} \to \{-c/2, c/2\} \subset \mathbb{R}$ by setting $g_\sigma(t_0) = 0$, and $g_\sigma(t_i) = \sigma_i c/2$ for $i \geqslant 1$. Clearly g_σ is 1-Lipschitz. By the injectivity of $\mathbb{R}$, it has an extension $h_\sigma : T \to \mathbb{R}$, belonging to $\mathcal{L}$.

Let $L = 2^M$, and identify $\{-1, 1\}^M$ with $\{1, \ldots, L\}$. Define the map $A : \ell_\infty(\mathcal{L}) \to \ell_\infty^L : (a_h)_{h \in \mathcal{L}} \to (a_{h_\sigma})_{\sigma=1}^L$ ($\mathcal{L}$ is the set of bounded 1-Lipschitz functions from T to $\mathbb{R}$, taking t_0 to 0). Clearly, this is a linear contraction, hence $d_n(T) \geqslant \delta_n^{(a)}(A \circ U(T))$. Moreover, $U(t_i) = (h(t_i))_{h \in \mathcal{L}}$, hence $e_i := A \circ U(t_i) = (h_\sigma(t_i))_\sigma$. For any real numbers $\alpha_1, \ldots, \alpha_M$,

$$\|\sum_i \alpha_i e_i\| = \max_\sigma |\sum_i \alpha_i h_\sigma(t_i)| = \frac{c}{2} \sum_i |\alpha_i|.$$

Therefore, $e_1, \ldots, e_M$ are linearly independent. Moreover, for $x \in \mathrm{span}[\pm e_1, \ldots, \pm e_M]$, $\|x\| \leqslant c/2$ if and only if $x \in S$, where $S = \mathrm{conv}(\pm e_1, \ldots, \pm e_M)$. In other words, S is the ball of a M-dimensional subspace of ℓ_∞^L, of radius $c/2$. By Lemma 2.c.8 of [36], $\delta_k(S) = c/2$ if $k < M$. Therefore,

$$d_n(T) \geqslant \delta_n^{(a)}(A \circ U(T)) \geqslant \delta_n^{(a)}(\{e_1, \ldots, e_M\}) \geqslant \delta_{n+1}(S) = \frac{c}{2}$$

whenever $n \leqslant M - 2$. ∎

As an application, we consider the sequence $(d_n(T))$ for finitely generated trees. Recall that such a tree T can be written as a finite union of elementary segments (that is, metric segments whose endpoints are either branching points or leaves, and whose interiors are devoid of branching points. Denote the sum of the length of these segments by $|T|$).

Corollary 6.14. *For a finitely generated tree T, $\delta_n^{(a)}(T) \sim |T|/n$, for any $n \geqslant N = N(T)$. More precisely, there exist universal constants c_1 and c_2 such that, for any finitely generated tree T, the inequality $c_1|T|/n \leqslant \delta_n^{(a)}(T) \leqslant c_2|T|/n$ holds for any $n \geqslant N = N(T)$.*

In fact, one can see that $N(T)$ depends on the minimum of lengths of the elementary segments comprising T.

Remark 6.15. For a general metric tree T, we have no good estimates for $\inf_{A,X} \delta_n^{(a)}(A(T)), X)$, where the infimum runs over all isometric embeddings A of T into a Banach space X. As we are interested in the infimum, we can assume that $X = \ell_\infty(I)$, for some index set I. For certain trees T and isometric embeddings A, $\delta_n^{(a)}(A(T)), X)$ can be much smaller than $d_n(T)$. For instance, pick $N \in \mathbb{N}$, and let $L = 2^N$. Consider a "spider" T with L limbs of length 1. More precisely, T consists of the "root" o, and the pairs (i, t), with $1 \leqslant i \leqslant L$, and $0 < t \leqslant 1$. For convenience, we identify o with $(i, 0)$. The metric on T is described in Example 1.5. We can embed T into ℓ_∞^N isometrically: let $e_1, \ldots, e_L$ be an enumeration of the vertices of the unit ball of ℓ_∞^N. Then the map $A : T \to \ell_\infty^N$, taking (i, t) to te_i, is an isometric embedding. Therefore, $\delta_n^{(a)}(A(T)) = 0$ for $n \geqslant N$. On the other hand, T contains a 2-separated set of cardinality $L = 2^N$ (the endpoints of the limbs of T). By Theorem 6.13, $d_n(T) \geqslant 1$ for $n \leqslant L - 2$.

References

1. A. G. Aksoy, B. Maurizi, *Metric trees, hyperconvex hulls, and extensions*, Turkish Math. J. 32 (2008), 219–234.
2. A. G. Aksoy, M. S. Borman, and A. L. Westfahl, *Compactness and measures of noncompactness in metric trees*, in: Proc. of the Inter. Symp. on Banach and Funct. Sp. II, M. Kato and L. Maligranda (eds.), Yokohama Publishers, Yokohama, 2008, 1–16.
3. A. G. Aksoy and M. A. Khamsi, *A selection theorem in metric trees*, Proc. of Amer. Math. Soc. 134 (2006), 2957–2966.
4. N. Aronszajn, P. Panitchpakdi, *Extensions of uniformly continuous transformations and hyperconvex metric spaces*, Pacific J. Math. 6 (1956), 405–439.
5. I. Bartolini, P. Ciaccia, and M. Patella, *String matching with metric trees using approximate distance*, SPIR, Lecture Notes in Computer Science, Springer-Verlag, 2476 (2002), 271–283.
6. F. Baudier, *Metrical characterization of super-reflexivity and linear type of Banach spaces*, Arch. Math. 89 (2007), 419–429.
7. B. Beauzamy, *Introduction to Banach spaces and their geometry*, North-Holland, Amsterdam, 1985.

8. Y. Benyamini and Y. Lindenstrauss, *Geometric nonlinear functional analysis I*, Amer. Math. Soc., Providence, RI, 2000.
9. M. Bestvina, *R-trees in topology, geometry, and group theory*, in: Handbook of geometric topology, R. Daverman and R. Sher (eds.), 55–91, North-Holland, Amsterdam, 2002.
10. L. M. Blumenthal, *Theory and applications of distance geometry*, Oxford University Press, London, 1953.
11. J. Bourgain, *The metrical interpretation of superreflexivity in Banach spaces*, Israel J. Math. 56 (1986), 222–230.
12. J. Bourgain, V. Milman, and H. Wolfson, *On type of metric spaces*, Trans. Amer. Math. Soc. 294 (1986), 295–317.
13. M. Bridson and A. Haefliger, *Metric spaces of nonpositive curvature*, Grundlehren der Mathematischen Wissenschaften, vol. 319, Springer-Verlag, Berlin, 1999.
14. B. Carl, I. Kyrezi, and A. Pajor, *Metric entropy of convex hulls in Banach spaces*, J. London Math. Soc. (2) 60 (1999), 871–896.
15. B. Carl and I. Stephani, *Entropy, compactness, and the approximation of operators*, Cambridge Univ. Press, 1990.
16. V. V. Chepyzhov and M. I. Vishik, *Attractors for equations of mathematical physics*, Amer. Math. Soc. Colloquium Publications 49 (2002), 163–197.
17. G. Darbo, *Punti uniti in trasformazioni a codominio non compatto*, Rend. Sem. Mat. Univ. Padova 24 (1955), 84–92.
18. D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory 52 (2006), 1289–1306.
19. D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, *Data compression and harmonic analysis*, IEEE Trans. Inform. Theory 44 (1998), 2435–2476.
20. I. Doust and A. Weston, *Enhanced negative type for finite metric trees*, J. Funct. Anal. 254 (2008), 2336–2364.
21. A. Dress, V. Moulton and W. Terhalle, *T-theory: An overview* Europ. J. Combinatorics 17 (1996), 161–175.
22. A. Dyubina and I. Polterovich, *Explicit construction of universal $\mathbb{R}$-trees and asymptotic geometry of hyperbolic spaces*, Bull. London Math. Soc. 33 (2001), 727–734.
23. D. E. Edmunds and H. Triebel, *Function spaces, entropy numbers and differential operators*, Cambridge University Press, Cambridge, 1996.
24. S. Evans, *Probability and real trees*, Springer, Berlin, 2008.
25. G. Godefroy, *Renormings of Banach spaces*, in: Handbook of the geometry of Banach spaces, Vol. I, W. Johnson and Y. Lindenstrauss (eds.), North-Holland, Amsterdam, 2001, 781–835.
26. S. Heinrich, *Ultraproducts in Banach space theory*, J. Reine Angew. Math. 313 (1980), 72–104.
27. P. Hjorth, P. Lisonek, S. Markvorsen, and C. Thomassen, *Finite metric spaces of strictly negative type*, Linear Algebra Appl. 270 (1998), 255–273.
28. W. Johnson and Y. Lindenstrauss, *Basic concepts in the geometry of Banach spaces*, in: Handbook of the geometry of Banach spaces, Vol. I, W. Johnson and Y. Lindenstrauss (eds.), North-Holland, Amsterdam, 2001, 1–84.
29. W. Johnson, Y. Lindenstrauss, D. Preiss, and G. Schechtman, *Lipschitz quotients from metric trees and from Banach spaces containing l_1*, J. Funct. Anal. 194 (2002), 332–346.
30. W. A. Kirk, *Hyperconvexity of R-trees*, Fund. Math. 156 (1998), 67–72.
31. W. A. Kirk and B. Sims (editors), *Handbook of metric fixed point theory*, Kluwer Acad. Publ., Dordrecht, 2001.
32. A. N. Kolmogorov and V. M. Tihomirov, *ε-entropy and ε-capacity of sets in function spaces*, Uspehi Mat. Nauk 14 (1959), 3–86.
33. K. Kuratowski, *Sur les espaces complets*, Fund. Math. 15 (1930), 301–309.

34. J. Lee, A. Naor, and Y. Peres, *Trees and Markov convexity*, Geom. Funct. Anal. 18 (2009), 1609–1659.
35. C. Lennard, A. Tonge, and A. Weston, *Generalized roundness and negative type*, Michigan Math. J. 44 (1997), 37–45.
36. J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces I*, Springer-Verlag, Berlin, 1977.
37. J. Matousek, *On embedding trees into uniformly convex Banach spaces*, Israel J. Math. 114 (1997), 221–237.
38. M. Mendel and A. Naor, *Metric cotype*, Ann. of Math. (2) 168 (2008), 247–298.
39. A. Naor, Y. Peres, O. Schramm, and S. Sheffield, *Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces*, Duke Math. J. 134 (2006), 165–197.
40. A. Pinkus, *n-widths in approximation theory*, A Series of Modern Surveys in Mathematics 7, Springer-Verlag, Berlin, 1985.
41. G. Pisier, *Probabilistic methods in the geometry of Banach spaces*, in: Probability and Analysis (Varenna, 1985), G. Letta and M. Pratelli (eds.), Lecture Notes in Math. 1206, Springer, New York, 1986, 167–241.
42. E. C. Posner and E. R. Rodemich, *ϵ-entropy and data compression*, Annals of Mathematical Statistics 42 (1971), 2079–2125.
43. C. Semple and M. Steel, *Phylogenetics*, Oxford Lecture Series in Mathematics and its Applications, 24, 2003.
44. K. T. Sturm, *Probability measures on metric spaces of nonpositive curvature*, in: Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), P. Auscher, T. Coulhon and A. Grigor'yan (eds.), Contemp. Math., 338, Amer. Math. Soc., Providence, RI, 2003, 357–390.
45. S. Szarek, *On the best constants in the Khinchin inequality*, Studia Math. J. 58 (1976), 197–208.
46. J. Tits, *A Theorem of Lie-Kolchin for trees*, in: Contributions to algebra: a collection of papers dedicated to Ellis Kolchin, H. Bass, P. Cassidy, and J. Kovacic (eds.), Academic Press, New York, 1977.

Asuman Güven Aksoy, Department of Mathematics, Claremont McKenna College Claremont, CA 91711
E-mail address: aaksoy@cmc.edu

Timur Oikhberg, Department of Mathematics, University of California-Irvine, Irvine, CA, 92697, *and* Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801
E-mail address: toikhber@math.uci.edu