# SOME RESULTS ON METRIC TREES

ASUMAN GÜVEN AKSOY

*Department of Mathematics, Claremont McKenna College*
*Claremont, CA 91711*
*E-mail: 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: toikhber@math.uci.edu*

**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 $\pi$ is a contractive retraction from the ambient Banach space $X$ onto $T$ (such a $\pi$ 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, $\epsilon$-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 [52] 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 [49], and in computer science involve string matching [5]. Universal properties of "$\ell_1$ 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 [31].

Since metric trees are described by three different names and several definitions, we start with some definition we will use. The first three are classical, and can be found, for instance, in [10] or [25].

DEFINITION 1.1. Let $x, y \in M$, where $(M, d)$ is a metric space. A *geodesic segment from x to y* (or a *metric segment*, denoted by $[x, y]$) is the image of an isometric embedding $\alpha : [a, b] \to M$ such that $\alpha(a) = x$ and $\alpha(b) = y$. A metric space is called *geodesic* if any two points can be connected by a metric segment.

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

DEFINITION 1.3. (Metric convexity) Suppose $T$ is a metric tree. A subset $S \subset T$ is called *metrically 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 con$S$.

DEFINITION 1.4. Suppose $T$ is a metric tree. Following [2], we call a $t \in T$ a *leaf* (or a *final point*) of $T$ if $t \in [x, y]$ (with $x, y \in T$) implies $t \in \{x, y\}$.

Below are some examples of metric trees.

EXAMPLE 1.5. (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 [19].

EXAMPLE 1.6. (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 1.7. (Finitely generated trees) Suppose $\mathcal{T}$ is a weighted graph-theoretical tree, whose sets of vertices and edges are denoted by $\mathcal{V}$ and $\mathcal{E}$, respectively. Let $\mathbf{d}_e$ denote the length (weight) of the edge $e$. We construct the metric tree $\tilde{\mathcal{T}}$, generated by $\mathcal{T}$, as a union of *elementary segments* $[v_1, v_2]$, where $v_1, v_2 \in \mathcal{V}$ are adjacent. In this case, $[v_1, v_2]$ is identified with the edge $e$, connecting $v_1$ and $v_2$, and $d(v_1, v_2) = \mathbf{d}_e$. The definition of

the distance $d$ is then extended to $\tilde{\mathcal{T}}^2$ in the obvious way. It is easy to see that all the tree axioms are satisfied.

EXAMPLE 1.8. (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

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

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

We can consider this tree as a finitely generated tree (see Example 1.7), arising from a "spider-like" graph, with vertices $v_0, v_1 \ldots, v_n$, and edges of length 1 connecting $v_0$ with $v_1, \ldots, v_n$.

The simplest spider – that with three legs – is called a *tripod* (this terminology comes from [50]).

EXAMPLE 1.9. (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 $\mathbb{H}_{\mathcal{U}}^n$ 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 [23].

We refer the reader to [10] for the properties of metric segments and to [2], [3] and [25] 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], [25, Section 3.3.1]) Let $x, y, z \in T$ ($T$ is a complete metric tree). There exists (necessarily unique) $w \in T$ such that

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

Consequently,

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

3. (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 leaves of $T$.

We also need to mention several properties of metric spaces.

DEFINITION 1.10. A metric space $(X, d)$ is said to be 0-*hyperbolic* (or to satisfy *the four-point inequality*) [19, 25] if, for any $x_1, x_2, x_3, x_4$ in $X$,

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

$X$ is said to satisfy *Reshetnyak's inequality* [50] if, for any $x_1, x_2, x_3, x_4 \in X$,

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

It was proven in [18] (see also [25, Chapter 3]) that any 0-hyperbolic metric space embeds isometrically into a metric tree. Moreover, a metric space $M$ is a metric tree if and only if it is 0-hyperbolic and geodesic. We see below (Lemma 5.1) that the four-point inequality implies Reshetnyak's inequality. The converse is not true, as an example of a Hilbert space shows.

Further properties of metric spaces are encoded in the definition below.

DEFINITION 1.11. A geodesic metric space $X$ is called a *CAT(0) space* or a global metric space of non-positive curvature (NPC), (*global NPC space*) if for every three points $x_0, x_1, x_2 \in X$, the *CN Inequality* holds:

$$d(x_0, y)^2 \leqslant \frac{d(x_0, x_1)^2}{2} + \frac{d(x_0, x_2)^2}{2} - \frac{d(x_1, x_2)^2}{4}$$

whenever $y$ is the midpoint of a metric segment connecting $x_1$ and $x_2$.

For information on these spaces, the reader is referred to [13], [50], or [43]. In [47], it was shown that a geodesic space is a CAT(0) space if and only if it satisfies Reshetnyak's inequality. The class of CAT(0) spaces includes metric trees (see Lemma 5.1), as well as Hilbert spaces and hyperbolic spaces [27].

Generalizing the classical Banach space notion of uniform convexity, we follow [27] in defining the modulus of convexity for geodesic metric spaces.

DEFINITION 1.12. Suppose $(M, d)$ is a geodesic metric space. For numbers $R > 0$, $\epsilon \in [0, 2R]$, and $a \in M$, let $\boldsymbol{\sigma}_M(a, R, \epsilon) = \inf\{1 - d(m, a)/R\}$, where $m$ is the midpoint of a metric segment connecting $x_1$ and $x_2$, and the infimum runs over all pairs $(x_1, x_2)$ with $\max\{d(a, x_1), d(a, x_2)\} \leqslant R$, and $d(x_1, x_2) \geqslant R\epsilon$. Define the *modulus of convexity* of $M$ by setting $\boldsymbol{\sigma}_M(R, \epsilon) = \inf_{a \in M} \boldsymbol{\sigma}_M(a, R, \epsilon)$.

The CN inequality implies that, for any CAT(0) space $M$,

$$\boldsymbol{\sigma}_M(R, \epsilon) \geqslant \boldsymbol{\sigma}_H(R, \epsilon) = \boldsymbol{\sigma}_H(\epsilon) = 1 - \sqrt{1 - \epsilon^2/4}$$

(here, $H$ is the Hilbert space of dimension greater than 1). In Lemma 5.2, we obtain a sharper estimate on the moduli of convexity of metric trees.

## 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 Panitch-pakdi [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 $X$ is a 1 absolute Lipschitz retract; that is, 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 $\ell_2$ fails to be hyperconvex, the spaces $\ell_\infty$ and $L_\infty$ are hyperconvex. The connection between hyperconvex metric spaces and metric trees is given in the following theorem:

THEOREM 2.3 ([32], [1]). *A complete metric tree $T$ is hyperconvex. Conversely, any hyperconvex space with unique metric segments is a complete metric tree.*

## 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 graphs (including trees) into Banach spaces. In particular, the connections between Lipschitz embeddability of trees and superreflexivity were investigated in [11], and more recently, in [6] (they are discussed in more detail in Section 3.3). The distortion necessary to embed a metric tree into a uniformly convex Banach space can be found in e.g. [39] (by [36], this problem is equivalent to computing the distortion of embedding the corresponding finite tree).

**3.1.** *Embeddings into $L_\infty$.* First we consider two embeddings into $L_\infty$, with different properties.

THEOREM 3.1 (From [33], 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 $\ell^\infty(X)$.*

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 $L_\infty$ 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 $\lambda$-injective Banach space, there exists a $\lambda$-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 $\lambda$. 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 linear on each interval $\left[\frac{1}{n+1}, \frac{2n+1}{2n(n+1)}\right]$ and $\left[\frac{2n+1}{2n(n+1)}, \frac{1}{n}\right]$. 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 an "injective" counterpart of the "projective" universal embedding $U$. More precisely, suppose $T$ is the "tripod" tree, described in Example 1.8. 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 a space $L_1(\mu)$, with the measure $\mu$ on $T$ defined below (we follow the construction from [26]). For any two points $x, y \in T$, denote by $\phi_{xy}$ the isometry from $[0, d(x, y)]$ to $[x, y]$, mapping 0 to $x$. A set $S \subset T$ is said to be measurable if $\phi_{xy}^{-1}(S)$ is a Lebesgue measurable subset of $[0, d(x, y)]$ for any $x, y \in T$. By the transitivity of betweenness (Section 1) the intersection of two metric segments is either empty, a singleton, or a metric segment, hence any metric segment is measurable.

For an interval $[x, y] \subset T$ and measurable $S \subset T$, we define $\mu_{[x,y]}(S) = \lambda(\phi_{xy}^{-1}(S))$, where $\lambda$ is the Lebesgue measure. Denote now by $\mathcal{F}$ the set of all finite unions $F = \cup_{k=1}^n [x_k, y_k]$ of disjoint unions of metric segments. For a measurable $S \subset T$, and $F$ as above, set $\mu_F(S) = \sum_{k=1}^n \mu_{[x_k, y_k]}(S)$. Finally, let $\mu_T(S) = \mu(S) = \sup_{F \in \mathcal{F}} \mu_F(S)$. It is easy to see that $\mu$ is indeed a measure, vanishing on countable sets, such that $\mu([x, y]) = d(x, y)$ for any $x, y \in T$.

The "semicanonical" embedding $U = U_{x_0}$ of $T$ into $L_1(\mu)$ ($x_0$ is a point in $T$), is defined by $U(x) = \chi_{[x_0, x]}$. To verify that $U$ is isometric, note that, for any $x, y \in T$, there exists a unique $z \in [x_0, x]$ s.t. $[x, y] = [x, z] \cup [z, y]$, and $[z, y] \cap [x_0, x] = \{z\}$. Then $U(x) - U(y) = \chi_{[x, z)} - \chi_{[y, z)}$, hence

$$\|U(x) - U(y)\| = \|\chi_{[x, z)}\| + \|\chi_{[y, z)}\| = d(x, z) + d(y, z) = d(x, y).$$

An embedding of a finitely generated tree into $\ell_1^N$ is described into [25, Section 2.5].

**3.3.** *A characterization of superreflexivity.* Recall that a Banach space $X$ is called *superreflexive* if all its ultrapowers are reflexive, or equivalently, any Banach space which can be finitely represented in $X$ must be reflexive. The reader is referred to [7] for many properties and characterizations of superreflexive spaces.

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.*

This theorem should be compared with the characterizations of superreflexive Banach spaces due to F. Baudier and J. Bourgain [6, 11]. Their results concern the *binary tree of height $n$* $\mathcal{T}_n = \cup_{j=0}^n \{-1, 1\}^j$ ($n \geqslant 0$), and the *infinite binary tree* $\mathcal{T}_\infty = \cup_{j=0}^\infty \{-1, 1\}^j$. We can view these objects as graphs, where the only edges are those connecting $([\alpha])$ with $([\alpha], \pm 1)$ ($\alpha \in \{-1, 1\}^j$). The graph structure induces the *hyperbolic distance* $d$, defined as follows. For $\alpha = (\alpha_1, \ldots, \alpha_k) \in \{-1, 1\}^k$ and $\beta = (\beta_1, \ldots, \beta_\ell) \in \{-1, 1\}^\ell$, denote by $s = s(\alpha, \beta)$ the smallest integer $j$ for which $\alpha_{j+1} \neq \beta_{j+1}$ (if $\alpha_1 \neq \beta_1$, or if either $k$ or $\ell$ equals 0, set $s = 0$). Let $d(\alpha, \beta) = k + \ell - 2s$.

The *Lipschitz constant* of an embedding $f : A \to B$ between metric spaces is defined as

$$\mathrm{L}(f) = \sup_{x \neq y} \frac{d_B(f(x), f(y))}{d_A(x, y)} \sup_{x \neq y} \frac{d_A(x, y)}{d_B(f(x), f(y))}$$

(here $d_A$ and $d_B$ are the distances in the spaces $A$ and $B$, respectively). We say that $A$ has a *Lipschitz embedding* into $B$ if there exists an embedding $f : A \to B$ with finite Lipschitz constant. A family $(A_n)$ is said to have a *uniform Lipschitz embedding* into $B$ if there exist embeddings $f_n : A_n \to B$, with $\sup_n \mathrm{L}(f_n) < \infty$.

J. Bourgain [11] proved that a Banach space $X$ is not superreflexive if and only if the family $(\mathcal{T}_n)$ has uniform Lipschitz embedding into $X$. Recently, F. Baudier [6] established that $\mathcal{T}_\infty$ Lipschitz embeds into any non-superreflexive space. Together with Theorem 3.3, these results yield:

THEOREM 3.4. *For a Banach space $X$, the following are equivalent:*

1. *$X$ is not superreflexive.*
2. *There exists a Lipschitz embedding of $\mathcal{T}$ into $X$.*
3. *There exist Lipschitz embeddings $f_n : \mathcal{T} \to X$ $(n \in \mathbb{N})$, with $\sup_n \mathrm{L}(f_n) < \infty$.*
4. *Any finitely generated metric tree embeds isometrically into an isomorphic copy of $X$.*

*Remark* 3.5. It is easy to note that a strictly convex space cannot contain a tripod, described in Example 1.8 (a Banach space $X$ is called strictly convex if the equality $2(\|x\|^2 + \|y\|^2) = \|x - y\|^2$ implies $x = -y$). If $X$ is separable, we can find an injection $T : X \to \ell_2$, and equip $X$ with the equivalent strictly convex norm $\||x\|| = (\|x\|^2 + \|Tx\|^2)^{1/2}$. Thus, there exist non-superreflexive Banach spaces which do not contain the tripod isometrically. This shows that renorming is essential in Theorems 3.3 and 3.4(4).

For the proof of Theorem 3.3, we need a simple

LEMMA 3.6. *Suppose a Banach space $X$ is not superreflexive, and $c \in (0, 1)$. Then for every $n \in \mathbb{N}$ there exists a family $(x_i)_{i=1}^n$ in the unit ball of $X$, such that*

1. *If $(a_i)$ is a sequence of scalars, changing signs at most once, then $\|\sum_{i=1}^n a_i x_i\| \geqslant 2c \sum_{i=1}^n |a_i|$.*
2. *For every $i$, $\mathrm{dist}(x_i, \mathrm{span}[x_j : j \neq i]) > c$.*

*Proof.* Fix $\lambda \in (c, 1)$. By a Ramsey-style result from [28], there exists $m = m(n, c, \lambda) \in \mathbb{N}$ with the following property: if $(y_i)_{i=1}^m$ is a subset of the unit ball of a Banach space $X$, and $\|y_i - y_j\| \geqslant 2\lambda$ whenever $i \neq j$, then there exist $1 \leqslant s_1 < s_2 < \ldots < s_n \leqslant m$, such that $dist(y_{s_i}, \mathrm{span}[y_{s_j} : j \in \{1, \ldots, n\} \backslash \{i\}]) > c$ for any $i \in \{1, \ldots, n\}$.

Now suppose $X$ is not superreflexive. By [48] (see also [7, Part 4]), there exist $y_1, \ldots, y_m$ in the unit ball of $X$, such that, for every $1 \leqslant k \leqslant m$, $\|y_1 + \ldots + y_k - y_{k+1} - \ldots - y_m\| > m + \lambda - 1$. We claim that $\|\sum_{j=1}^m \alpha_j y_j\| \geqslant \lambda \sum_{j=1}^m |\alpha_j|$ if the sequence $(\alpha_j)$ changes sign at most once. Indeed, suppose $\alpha_j \geqslant 0$ for $j \leqslant k$, and $\alpha_j \leqslant 0$ for $j \geqslant k + 1$. By scaling, we can assume $\sum_j |\alpha_j| = 1$. Then $0 \leqslant \alpha_j \leqslant 1$ for $1 \leqslant j \leqslant k$, and

$0 \geqslant \alpha_j \geqslant -1$ for $k+1 \leqslant j \leqslant m$. By the triangle inequality,

$$\|\sum_{j=1}^{m} \alpha_j y_j\| \geqslant \|y_1 + \ldots + y_k - y_{k+1} - \ldots - y_m\| - \sum_{j=1}^{m}(1 - |\alpha_j|) > (m + \lambda - 1) - m + 1 = \lambda.$$

By our choice of $m$, we can find the vectors $x_k = y_{s_k}$ with the required properties. ■

*Proof of Theorem 3.3.* A weighted tree graph $\mathcal{T} = (\mathcal{V}, \mathcal{E})$ ($\mathcal{V}$ and $\mathcal{E}$ denote the sets of vertices and edges, respectively) gives rise to the metric tree $T$, as in Example 1.7. Select $v_{\emptyset} \in \mathcal{V}$, and call it the *root*. Enumerate the immediate descendants of $v_{\emptyset}$ (that is, the vertices connected to $v_{\emptyset}$ by edges) by $v_1, \ldots, v_{n_{\emptyset}}$. For $1 \leqslant i \leqslant n_{\emptyset}$, let $a_i = d(v_{\emptyset}, v_i)$. For each $i$, enumerate its own immediate descendants $v_{11}, \ldots, v_{1n_1}$, and set $a_{ij} = d(v_i, v_{ij})$. Proceeding further in the same manner, we write $\mathcal{V}$ as the collection of points $v_S$, for a finite collection $\mathcal{S}$ of finite strings $S$. Then $v_{S'}$ is a descendant of $v_S$ if and only if $S' = S \smile j$, for some $j \in [1, n_S]$. Set $a_{S'} = d(v_S, v_{S'})$, where $S$ is the unique immediate predecessor of $S'$.

For $S_1, S_2 \in \mathcal{S}$, write $S_1 \prec S_2$ if $v_{S_1}$ is a predecessor of $v_{S_2}$, 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 is no such $k$, set $S_1 \wedge S_2 = \emptyset$. Then $v_{S_1 \wedge S_2}$ is the largest common predecessor of $v_{S_1}$ and $v_{S_2}$.

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

$$d(v_{S_1}, v_{S_2}) = d(v_{S_1}, v_S) + d(v_S, v_{S_2}) = \sum_{m=k_0+1}^{k_1} a_{(i_{11} \ldots a_{m1})} + \sum_{m=k_0+1}^{k_2} a_{(i_{12} \ldots a_{m2})}. \quad (3.1)$$

The main step is to renorm $X$ (making it into $Y$) in such a way that there exists an isometry $J_{\mathcal{V}} : \mathcal{V} \to Y$. We then extend it to $J : T \to Y$ so that $J|_{\mathcal{V}} = J_{\mathcal{V}}$. For $t \in T \backslash \mathcal{E}$, there exist unique $v_1, v_2 \in \mathcal{V}$ such that $t$ belongs to the elementary segment $[v_1, v_2]$. Let $\lambda = d(v_1, t)/d(v_1, v_2)$. Define $J(t) = \lambda J_{\mathcal{V}}(v_1) + (1 - \lambda) J_{\mathcal{V}}(v_2)$. Clearly, $J$ is an isometry on any elementary segment. By the description of metric segments given in [10, Lemma 15.1], $J$ is an isometry on $T$.

To construct $J_{\mathcal{V}} : \mathcal{V} \to Y$, denote the cardinality of $\mathcal{V}$ by $N$. By Lemma 3.6, there exists, for every $M \in \mathbb{N}$, a family $(x_{iM})_{i=1}^{N} \subset B(0; 1)$ such that

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

for any sequence $(\alpha_i)$ with at most one change of signs, and $\|\sum_{i=1}^{N} \alpha_i x_{iM}\| \geqslant (1 + 2^{-M})^{-1} \max_i |\alpha_i|$ for any sequence $(\alpha_i)$.

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 $S_\alpha = S \smile j_{1\alpha} \smile \ldots \smile j_{m_\alpha \alpha}$ ($\alpha = 1, 2$). 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 $J_M : \mathcal{V} \to X$ by setting, for $S = (i_1, \ldots i_k)$,

$$J_M(v_S) = \sum_{j=1}^{k} d(v_{i_1 \ldots v_{j-1}}, v_{i_1 \ldots v_j}) x_{\phi(i_1 \ldots i_j) M},$$

and $J_M(v_\emptyset) = 0$, By (3.1) and (3.2),

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

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 $(\alpha_i)_{i=1}^N$. For any such sequence, $\max_i |\alpha_i| \leqslant \Phi(\alpha_1, \ldots, \alpha_N) \leqslant \sum_i |\alpha_i|$. Thus, we can define a normed space $Z$ by setting $\|\sum_i \alpha_i e_i\| = \Phi(\alpha_1, \ldots, \alpha_N)$, where $(e_i)_{i=1}^N$ is the canonical basis for $\mathbb{R}^N$. Denote the span of $(x_{iM})_{i=1}^N$ in $X$ by $Z_M$, and define the map $U_M : Z_M \to Z : x_{iM} \mapsto e_i$. Find $M$ so large that $c = \max\{\|U_M\|, \|U_M^{-1}\|\} < \sqrt{1 + \epsilon}$. Consider $J = U_M \circ J_M : \mathcal{V} \to Z$. As $\|\sum_i \alpha_i e_i\| = \sum_i |\alpha_i|$ if the sequence $(\alpha_i)$ changes sign no more than once, the map $J$ is an isometry.

To renorm $X$, embed $Z$ isometrically into $\ell_\infty$. Then there exists $\tilde{U} : X \to \ell_\infty$, with $\tilde{U}|_{Z_M} = U_M$, and $\|\tilde{U}\| \leqslant c$. For $x \in X$, define $\|x\|_Y = \max\{c^{-1}\|x\|, \|\tilde{U}x\|\}$. Clearly, $c^{-1}\|x\| \leqslant \|x\|_Y \leqslant c\|x\|$. Moreover, for $x \in Z_M$, $\|\tilde{U}x\| = \|U_M x\| \geqslant c^{-1}\|x\|$, hence $\|x\|_Y = \|U_M x\|$. In other words, $Y$ contains $Z$ isometrically.

Consider $J = U_M \circ J_M : \mathcal{V} \to Z$. As $\|\sum_i \alpha_i e_i\| = \sum_i |\alpha_i|$ if the sequence $(\alpha_i)$ changes sign no more than once, the map $J$ is an isometry. Therefore, the map $J_M : \mathcal{V} \to Y$ is an isometry. ∎

**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 [50]. 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 $\pi$ 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

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

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 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)

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

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 $(1) \Rightarrow (2)$ 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 $\phi : T \rightarrow \mathbb{R}$, defined by

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

Then

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

$$\geqslant |\sum_k \alpha_k(\|x_k - x\| - \|x - x\|)| = \sum_k \alpha_k \|x_k - x\| \geqslant \|x_0 - x\|.$$

∎

THEOREM 4.4. *Suppose $T$ is a compact metric tree, embedded into $L_1(\mu_T)$ 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.* As in Theorem 4.3, we only need to establish (2) $\Rightarrow$ (1). Suppose $x$ is a leaf of the tree $T$. By Proposition 4.1, we have to show that, if $x_0 \in T$ is such that $d(x_0, x) \leqslant \sum_k \alpha_k d(x_k, x)$, then $\|x_0 - x\| \leqslant \|\tilde{x}^{(\alpha)} - x\|$. The semicanonical embedding of $T$ into $L_1$ identifies $t \in T$ with $\chi_{[x_0, t]}$ (by translation, we can identify $x_0$ with 0). For each $k$, find $u_k \in T$ satisfying $[x_0, x_k] \cap [x_0, x] = [x_0, u_k]$. Then $y_k = x_k - u_k = \chi_{[x_k, u_k]}$ and $z_k = x - u_k = \chi_{[x, u_k]}$ have disjoint support, hence

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

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 $x \in T$, $\|\sum_k \alpha_k (x_k - x)\| = \sum_k \alpha_k \|x_k - x\|$. Thus, it suffices to establish that, at any point $y \in T$, the signs of $(x_k - x)(y) = \chi_{[x_0, x_k]}(y) - \chi_{[x_0, x]}(y)$ are independent of $k$. If $y \notin [x_0, x]$, then $\chi_{[x_0, x_k]}(y) - \chi_{[x_0, x]}(y) \geqslant 0$ for any $k$. On the other hand, if $y \notin [x_0, x]$, then $\chi_{[x_0, x_k]}(y) - \chi_{[x_0, x]}(y) \leqslant 0$ for any $k$. $\blacksquare$

*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 [50]. 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 $x_1, \dots, x_n$ (see Definition 1.3 for the definition of metric convexity).

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 1.8). If $T$ is embedded into $\ell_\infty(T)$ (respectively $L_1$) 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, metrically 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 [50],

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

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

In certain cases, when the structure of $x_1, \ldots, x_n$ 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 1.8). 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,

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

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

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

and

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

$$= 1 + \frac{1 - c}{3} = 1 + \beta.$$

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 $S = (i_0, [0, \beta]) \cup \left(\cup_{i \neq i_0} (i, [0, \alpha])\right)$ 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

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

with $a \in [0, 1]$ to be determined later (here, $\delta_x$ 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

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

hence

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

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 $S = (1, [0, 1/3]) \cup \left(\cup_{i \neq 1} (i, [0, \alpha])\right)$, 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 $yu \leqslant 0$ $\mu$-a.e. (we view $y$ and $u$ as functions on $\Omega$). As $x_1, \ldots, x_n \in S$, we also have $\tilde{x}^{(\alpha)} y \leqslant 0$ $\mu$-a.e.. Therefore, $\|y - \tilde{x}^{(\alpha)}\| \geqslant \|y\| = d(x', y)$, which is what we need. ■

EXAMPLE 4.10. Proposition 4.9 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 $\ell_\infty^3$ 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 $\ell_\infty^3$). 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 $\ell_\infty^3$, $\mathbf{P}(\tilde{x}_0) = \{(4, 1/3)\}$. 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, cotype, and convexity of Metric Trees.** In this section we consider properties of metric spaces, such as the four-point inequality and Reshetnyak's inequality (see

Definition 1.10), type, and cotype. The notion of metric type was introduced in [12] (see also [45]). More recently, metric cotype was defined in [40].

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

*Proof.* Suppose the elements $x_1, x_2, x_3, x_4$ of a metric space $(X, d)$ 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,

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

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. ∎

Therefore, any metric tree is a CAT(0) space. Below we show that metric trees are "more convex" (that is, their moduli of convexity are larger) than those of "generic" CAT(0) spaces.

LEMMA 5.2. *Suppose $T$ is a complete metric tree. Then, for any $R > 0$ and $\epsilon \in [0, 2R]$, $\sigma_M(R, \epsilon) \geqslant 1 - \epsilon/2$.*

*Proof.* Suppose $a, x_1, x_2 \in T$ are such that $\max\{d(a, x_1), d(a, x_2)\} \leqslant R$, and $d(x_1, x_2) \geqslant R\epsilon$. We have to show that $d(a, m) \leqslant R - R\epsilon/2$, where $m$ is the midpoint of $[x_1, x_2]$. Find $y \in [x_1, x_2]$ such that $[a, x_1] = [a, y] \cup [y, x_1]$, and $[a, x_2] = [a, y] \cup [y, x_2]$. Relabeling if necessary, we can assume that $y \in [x_2, m]$. Then $m \in [y, x_1]$, hence $R \geqslant d(a, x_1) = d(a, m) + d(m, x_1) \geqslant d(a, m) + R\epsilon/2$. Thus, $d(a, m) \leqslant R - R\epsilon/2$. ∎

DEFINITION 5.3. Suppose $1 \leqslant p \leqslant 2$, and $K > 0$. A metric space $(X, d)$ is said to have *metric type $p$* (or *BMW type $p$*, after Bourgain, Milman, and Wolfson, who introduced this notion in [12]) with constant $K$ if, for any $n \in \mathbb{N}$, and any function $f : \{-1, 1\}^n \to X$, we have

$$\sum_{\epsilon \in \{-1,1\}^n} d(f(\epsilon), f(-\epsilon))^2 \leqslant K n^{1/p-1/2} \sum_{\epsilon \in \{-1,1\}^n} \sum_{i=1}^n d(f(\epsilon), f(\epsilon^{[i]}))^2, \qquad (5.1)$$

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 (5.1) 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 [40].

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.* Part (1) was proved in [42]. For Part (2), 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 [51] with the basic properties of cotype, described in e.g. [30]). Therefore, by Theorem 1.2 of [40], $L_1$ has metric cotype 2, with the constant $90\sqrt{2}$. We have seen that any finitely generated metric tree embeds isometrically into $\ell_1^N$, 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. ∎

We next tackle the negative type of metric trees. Recall that a metric space $X$ has *negative type $p$ ($p > 0$)* if, for any $x_1, \ldots, x_n \in X$, the $n \times n$ matrix $(d(x_i, x_j)^p)$ is conditionally negative definite. Recall 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. [22, 37]. 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 [29] 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 $t_1, \ldots, t_n$. The corresponding $(n+1) \times (n+1)$ distance matrix is

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

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 $\xi$ 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 [41].

## 6. Entropy Quantities and Other Measures of Compactness.

**6.1.** *$\epsilon$-entropy and related quantities.* Kolmogorov introduced the notion of $\epsilon$-entropy as a measure of the massiveness of sets [34]. This notion has been useful in function spaces (see [24]), 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 [21], [46], [20]). 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 the $\epsilon$-entropy of a bounded subset $A$ of a tree $T$ is equal to the $\epsilon$-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. Suppose $A$ is a subset of a metric space $M$.

- $A$ is *centered* if for all $U \subset A$ such that $\mathrm{diam}\,(U) = 2r$, there exists $a \in A$ such that $U \subset B_c(a; r)$. By $B_c(a; r)$ we mean the closed ball of radius $r$ centered at $a$.
- $\{m_i\}_{i \in I} \subset M$ is an *$\epsilon$-net for $A$ in $M$* if

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

- $\{U_\alpha\}_{\alpha \in I}$, where $U_\alpha \subset M$, is an *$\epsilon$-cover for $A$* if $\mathrm{diam}\,(U_\alpha) \leqslant 2\epsilon$ and

$$A \subset \bigcup_{\alpha \in I} U_\alpha.$$

- $U \subset A$ is a *$\epsilon$-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 $\mathcal{N}_\epsilon(A)$ ($\mathcal{K}_\epsilon^M(A)$) be the cardinality of a minimal $\epsilon$-cover of $A$ (rep. minimal $\epsilon$-net for $A$ in $M$). Define $\mathcal{M}_\epsilon(A)$ as the maximal cardinality of an $\epsilon$-separated subset of $A$.

Note that, if $A$ is a complete metric tree, then it is injective, hence $\mathcal{K}_\epsilon^M(A) = \mathcal{K}_\epsilon^A(A)$ for any ambient space $M$.

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

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

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. Alternatively, one can prove Theorem 6.2 by recalling, from Theorem 2.3, that any complete metric tree is hyperconvex. By the definition of hyperconvexity (Definition 2.1), any hyperconvex set is centered. 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 $\mathcal{K}_\epsilon^T(A) = \mathcal{M}_\epsilon(A)$.*

*Proof.* Any complete metric tree is centered. Thus, any $\epsilon$-net for $A$ is equivalent to an $\epsilon$-cover. ∎

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 6.6. *Suppose $S$ is a compact subset of a complete metric tree $T$ and $\epsilon_1, \epsilon_2$ are positive numbers. Then*

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

*Proof.* For the sake of brevity, set $N = N_{\epsilon_1}(S)$, $D = \mathrm{diam}\,(S)$, and $S' = \mathrm{con}(S)$. Convexity of the norm (see [50]) implies that the diameter of $S'$ equals $D$. By Theorem 6.2, there exists $x_0 \in S'$ such that for any $y \in S'$, $d(x_0, y) \leqslant \mathrm{diam}\,(S')/2 = D/2$, 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 [50], 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 "flatness" of compact subsets of linear spaces. Since then, Kolmogorov diameters have been widely used in approximation theory (see [44] 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 [35] 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 show the connections between them.

DEFINITION 6.7. Given a subset $A$ of a normed linear space $X$ and $n \geqslant 0$, define the *n-th Kolmogorov diameter (n-width) of $A$ in $X$* as:

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

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

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

Observe that the sequences $\{\delta_n(A)\}_{n=1}^\infty$ and $\{\delta_n^{(a)}(A)\}_{n=1}^\infty$ are non-increasing, and

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

Indeed, the left hand side inequality is obvious. To establish the right hand side, 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 $x \in A \cup (-A)$ can be expressed as $x = \sum_{i=1}^{N} \alpha_i a_i$, with $\sum_i |\alpha_i| \leqslant 1$, and $a_i \in A$. For each $i$, find $m_i \in M$ such that $\|a_i - m_i\| < c$. Then $M' = \operatorname{conv}(M \cup (-M))$ is a linear subspace of dimension not exceeding $n + 1$, $m = \sum_i \alpha_i m_i \in M'$, and $\|x - m\| < c$.

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 not exceeding $n$, 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.

DEFINITION 6.8. Suppose $A$ is a subset of the metric space $M$. Define the *ball (Hausdorff) measure of non-compactness* and the *set measure of non-compactness* as

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

and

$$\alpha(A) := \inf \left\{ a > 0 \mid A \subset \bigcup_{j=1}^{k} A_j \text{ for some } A_j \subset A \text{ with } \operatorname{diam}(A_j) \leqslant a \right\},$$

respectively.

Note that $\beta(A, M)$ is an "extrinsic" measure of non-compactness, and may depend on the ambient space $M$. On the other hand, $\alpha(A)$ is intrinsic, and is independent of $M$. It is easy to observe that $\beta(A, M) \leqslant \alpha(A) \leqslant 2\beta(A, M)$.

Connections between entropy of linear maps, their Kolmogorov numbers (and other $s$-numbers), and their analytic properties, such as eigenvalues and essential spectrum, have been studied extensively (see [15] and references therein). Below we present some results illuminating the connections between Kolmogorov numbers and entropy properties of metric spaces.

THEOREM 6.9. *Suppose $A$ is a bounded subset of a Banach space $X$. Then*

$$\lim_{n \to \infty} \delta_n(A, X) = \beta(A, X) = \lim_{n \to \infty} \delta_n^{(a)}(A, X).$$

*Proof.* We only show the equality involving $\delta_n(A, X) = \delta_n(A, X)$, as the one with $\delta_n^{(a)}(A, X)$ is handled in a similar manner. By the boundedness of $A$, $\{\delta_n(A, X)\}_{n=1}^{\infty}$ forms a non-increasing sequence of nonnegative numbers, hence $\lim_{n \to \infty} \delta_n(A, X)$ exists.

(1) $\beta(A, X) \leqslant \lim_{n \to \infty} \delta_n(A, X)$. Pick $c > b > \lim_n \delta_n(A, X)$, and show $\beta(A, X) \leqslant c$. Thus, there exists $n$ such that $\delta_n(A, X) < b$. This means there there exists an $n$-dimensional subspace $E$ of $X$ such that $\sup_{a \in A} d(a, E) < b$. Let $Q = \{e \in E : \|e\| \leqslant b + \sup_{a \in A} \|a\|\}$. By compactness, $Q$ contains a finite $(c - b)$-net $\{q_n\}$. Then $A \subset \cup_n B(q_n, c)$, hence $\beta(A, X) \leqslant c$.

(2) $\beta(A, X) \geqslant \lim_{n \to \infty} \delta_n(A, X)$. Let $b > \beta(A, X)$, and show that $\delta_n(A, X) \leqslant b$. Find a finite $b$-net $\{x_j\}_{j=1}^n \subset X$ for $A$. Then the dimension of $E = \text{span}(\{x_j\}_{j=1}^n)$ does not exceed $n$. For any $a \in A$, $\text{dist}(a, E) \leqslant \min_j \|a - x_j\| < b$, hence $\delta_n(A, X) \leqslant b$. ∎

COROLLARY 6.10. *Suppose a complete metric tree $T$ is embedded isometrically into a Banach space $X$, and $A$ is a bounded subset of $T$. Then*

$$\lim_{n \to \infty} \delta_n(A, X) = \lim_{n \to \infty} \delta_n^{(a)}(A, X) = \frac{\alpha(A)}{2}.$$

*Proof.* Clearly, $\alpha(A) \leqslant 2\beta(A, X)$. By Theorem 6.2, $\alpha(A) = 2\beta(A, T) \geqslant 2\beta(A, X)$. Thus, $\alpha(A) = 2\beta(A, X)$. An application of Theorem 6.9 completes the proof. ∎

Next we consider affine Kolmogorov diameters of $V(S)$, where $V$ is an embedding of a metric space $S$ into a Banach space $X$. It is well known (see e.g. [15, 44], or Remark 6.13 below) that Kolmogorov diameters may depend heavily on the ambient space $X$. If $X$ is a subspace of $Y$, then $\delta_n(V(S), Y) \leqslant \delta_n(V(S), X)$. Furthermore, if $X$ is contained in a $\lambda$-injective space $Z$, then $\delta_n(V(S), Z) \leqslant \lambda \delta_n(V(S), Y)$. Similar inequalities hold for $\delta_n^{(a)}$.

Suppose now that $V$ is an isometric embedding of $S$ into a 1-injective Banach space $X$. Then, by Theorem 3.2, $\delta_n^{(a)}(V(S), X) \leqslant d_n(S)$, where $d_n(S) = \delta_n^{(a)}(U(S), \ell_\infty(\mathcal{L}))$ (the universal embedding $U : S \to \ell_\infty(\mathcal{L})$ was defined in Section 3.1).

PROPOSITION 6.11. *Suppose $S$ is a metric space. Let*

$$c_1 = \inf\{\epsilon > 0 \mid \mathcal{K}_\epsilon^S(S) \leqslant n\}, \quad and \ c_2 = \sup\{\epsilon > 0 \mid \mathcal{M}_\epsilon(S) \geqslant n + 1\}.$$

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

*Proof.* (1) Fix $c > c_1$, and suppose $s_1, \ldots, s_n$ is a $c$-net in $S$. Let $E$ be the affine span of $V(s_1), \ldots, V(s_n)$. Then $\dim E \leqslant n$. Furthermore, for any $s \in S$,

$$d(V(s), E) \leqslant \min_i d(V(s), V(s_i)) \leqslant \min_i d(s, s_i) \leqslant c,$$

which shows that $\delta_n^{(a)}(V(S)), X) \leqslant c$. As $c > c_1$ is arbitrary, $\delta_n^{(a)}(V(S)), X) \leqslant c_1$.

(2) Let $M = n + 1$. For $c < c_2$, let $s_1, \ldots, s_M$ be a $c$-separated subset of $S$. Then there exists $s_0 \in T$ such that $d(s_0, s_i) > c/2$ for any $i$. Indeed, by relabeling if necessary, we can assume that $d(s_1, s_2) \leqslant d(s_i, s_j)$ whenever $i$ and $j$ are different. Let $s_0$ be the midpoint of $[s_1, s_2]$. We claim that $d(s_0, s_i) > c/2$ for any $i$. The inequality clearly holds for $i \in \{1, 2\}$. If $i > 2$ and $d(s_0, s_i) \leqslant c/2$, then

$$d(s_1, s_i) \leqslant d(s_1, s_0) + d(s_0, s_i) \leqslant \frac{d(s_1, s_2)}{2} + \frac{c}{2} < d(s_1, s_2),$$

a contradiction.

Let $\mathcal{L}$ be the set of 1-Lipschitz functions from $T$ to $\mathbb{R}$, taking $s_0$ to 0. For any $\sigma = (\sigma_1, \ldots, \sigma_M) \in \{-1, 1\}^M$, define $g_\sigma : \{s_0, s_1, \ldots, s_M\} \to \{-c/2, c/2\} \subset \mathbb{R}$ by setting $g_\sigma(s_0) = 0$, and $g_\sigma(s_i) = \sigma_i c/2$ for $i \geqslant 1$. Clearly $g_\sigma$ is 1-Lipschitz. By the injectivity of $\mathbb{R}$, it has an extension $h_\sigma : S \to \mathbb{R}$, belonging to $\mathcal{L}$.

Let $L = 2^M$, and identify $\{-1, 1\}^M$ with $\{1, \ldots, L\}$. Define the map $A : \ell_\infty(\mathcal{L}) \to \ell_\infty^L :$ $(a_h)_{h \in \mathcal{L}} \to (a_{h_\sigma})_{\sigma=1}^L$. Clearly, this is a linear contraction, hence $d_n(S) \geqslant \delta_n^{(a)}(A \circ U(S))$.

Moreover, $U(s_i) = (h(s_i))_{h \in \mathcal{L}}$. For $1 \leqslant i \leqslant M$ consider $e_i = A \circ U(s_i) = (h_\sigma(s_i))_\sigma \in \ell_\infty^L$. For any real numbers $\alpha_1, \ldots, \alpha_M$,

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

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 $x \in C$, where $C = \mathrm{conv}(\pm e_1, \ldots, \pm e_M)$. In other words, $C$ is the ball of a $M$-dimensional subspace of $\ell_\infty^L$, of radius $c/2$. By Lemma 2.c.8 of [38], $\delta_k(C, \ell_\infty^L) = c/2$ if $k < M$. The set $C$ is centrally symmetric, hence $\delta_k(C, \ell_\infty^L) = \delta_k^{(a)}(C, \ell_\infty^L)$. As $n = M - 1$, we conclude

$$d_n(S) \geqslant \delta_n^{(a)}(A \circ U(S)) \geqslant \delta_n^{(a)}(C, \ell_\infty^L) \geqslant \delta_n(C, \ell_\infty^L) = \frac{c}{2}.$$

Since this inequality is valid for any $c < c_2$, we are done. ∎

As an application, we estimate $d_n(T)$ for finitely generated trees. That is, suppose $T$ arises from a weighted graph theoretical tree $\mathcal{T}$ (see Example 1.7). For such a tree, denote by $|T|$ the sum of weights (lengths) of the edges of the original graph.

COROLLARY 6.12. *There exist $0 < c_1 < c_2$ with the property that, for any finitely generated tree $T$, there exists $N = N(T) \in \mathbb{N}$ such that 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 edges of $\mathcal{T}$.

*Remark* 6.13. For a metric tree $T$, we have no good estimates for $\inf_{V,X} \delta_n^{(a)}(V(T), X)$, where the infimum runs over all isometric embeddings $V$ 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 $A$, $\delta_n^{(a)}(V(T), X)$ 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 1.8. We can embed $T$ into $\ell_\infty^N$ isometrically: let $e_1, \ldots, e_L$ be an enumeration of the vertices of the unit ball of $\ell_\infty^N$. Then the map $V : T \to \ell_\infty^N$, taking $(i, t)$ to $te_i$, is an isometric embedding. Therefore, $\delta_n^{(a)}(V(T), \ell_\infty^N) = 0$ 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 6.11, $d_n(T) \geqslant 1$ for $n \leqslant L - 1$.

## 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 noncom-
      pactness 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 and 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 approxi-
      mate 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, Amster-
      dam, 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*, Cam-
      bridge 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]  A. Dress, *Trees, tight extensions of metric spaces, and the cohomological dimension of
      certain groups: a note on combinatorial properties of metric spaces*, Adv. Math. 53 (1984),
      321–402.

[19]  A. Dress, V. Moulton, and W. Terhalle, *T-theory: an overview*, European J. Combin. 17
      (1996), 161–175.

[20]  D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory 52 (2006), 1289–1306.

[21]  D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, *Data compression and
      harmonic analysis*, IEEE Trans. Inform. Theory 44 (1998), 2435–2476.

[22]  I. Doust and A. Weston, *Enhanced negative type for finite metric trees*, J. Funct. Anal.

254 (2008), 2336–2364.

[23] 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.

[24] D. E. Edmunds and H. Triebel, *Function spaces, entropy numbers and differential operators*, Cambridge University Press, Cambridge, 1996.

[25] S. Evans, *Probability and real trees*, Springer, Berlin, 2008.

[26] A. Godard, *Tree metrics and their Lipschitz-free spaces*, preprint, available at http://xxx.lanl.gov/abs/0904.3178.

[27] K. Goebel and S. Reich, *Uniform convexity, hyperbolic geometry, and nonexpansive mappings* Monographs and textbooks in pure and applied mathematics, vol. 83, Marcel Dekker, 1984.

[28] W. Henson, N. Kalton, T. Peck, I. Terescak, and P. Zlatos, *Some Ramsey type theorems for normed and quasinormed spaces*, Studia Math. 124 (1997), 81–100.

[29] P. Hjorth, P. Lisonek, S. Markvorsen, and C. Thomassen, *Finite metric spaces of strictly negative type*, Linear Algebra Appl. 270 (1998), 255–273.

[30] 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.

[31] 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.

[32] W. A. Kirk, *Hyperconvexity of R-trees*, Fund. Math. 156 (1998), 67–72.

[33] W. A. Kirk and B. Sims (editors), *Handbook of metric fixed point theory*, Kluwer Acad. Publ., Dordrecht, 2001.

[34] A. N. Kolmogorov and V. M. Tihomirov, *$\varepsilon$-entropy and $\varepsilon$-capacity of sets in function spaces*, Uspehi Mat. Nauk 14 (1959), 3–86.

[35] K. Kuratowski, *Sur les espaces complets*, Fund. Math. 15 (1930), 301–309.

[36] J. Lee, A. Naor, and Y. Peres, *Trees and Markov convexity*, Geom. Funct. Anal. 18 (2009), 1609–1659.

[37] C. Lennard, A. Tonge, and A. Weston, *Generalized roundness and negative type*, Michigan Math. J. 44 (1997), 37–45.

[38] J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces I*, Springer-Verlag, Berlin, 1977.

[39] J. Matousek, *On embedding trees into uniformly convex Banach spaces*, Israel J. Math. 114 (1999), 221–237.

[40] M. Mendel and A. Naor, *Metric cotype*, Ann. of Math. (2) 168 (2008), 247–298.

[41] 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.

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

[43] S.-i. Ohta and M. Pichot, *A note on Markov type constants*, Arch. Math. 92 (2009), 80–88.

[44] A. Pinkus, *n-widths in approximation theory*, A Series of Modern Surveys in Mathematics 7, Springer-Verlag, Berlin, 1985.

[45] 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.

[46] E. C. Posner and E. R. Rodemich, *$\epsilon$-entropy and data compression*, Annals of Mathematical Statistics 42 (1971), 2079–2125.

[47] T. Sato, *An alternative proof of Berg and Nikolaev's characterization of* CAT(0)*-spaces*

*via quadrilateral inequality*, Arch. Math. 93 (2009), 487–490.

[48]    J. Schaffer and K. Sundaresan, *Reflexivity and the girth of spheres*, Math. Ann. 184 (1969/1970), 163–168.

[49]    C. Semple and M. Steel, *Phylogenetics*, Oxford Lecture Series in Mathematics and its Applications, 24, 2003.

[50]    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.

[51]    S. Szarek, *On the best constants in the Khinchin inequality*, Studia Math. 58 (1976), 197–208.

[52]    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.