Spaces of algebraic measure trees
and triangulations of the circle

Wolfgang Löhr$^{1,2}$     Anita Winter$^1$

November 29, 2018

Abstract

In this paper we investigate algebraic trees which can be considered as (continuum) metric trees in which the metric distances are ignored and in which therefore the focus lies on the tree structure. We give an axiomatic definition of such trees, which we call algebraic trees, using a branch point map and show that any order separable algebraic tree can be represented by a metric tree. We further consider algebraic measure trees which are algebraic trees additionally equipped with a sampling (probability) measure. This measure gives rise to the branch point distribution which turns out to be the length measure of an intrinsic choice of such a metric tree representation.

We will provide a notion of convergence of algebraic measure trees which resembles the idea of the Gromov-weak topology which itself is defined through weak convergence of sample distance matrices.

Binary algebraic (measure) trees are of particular interest because they often arise in praxis, and also due to their close connection to triangulations of the circle. We use this connection to show that in the subspace of binary algebraic measure trees, weak convergence of sample shapes, sample subtree masses and sample distance matrices are all equivalent and define a compact, metrizable topology. Furthermore, the coding by triangulations is a continuous, surjective operation in this topology.

Contents

1 Introduction 2

2 Algebraic trees 6

2.1 The branch point map . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7

2.2 Morphisms of algebraic trees . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

2.3 Algebraic trees as topological spaces . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11

2.4 Metric tree representations of algebraic trees . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14

2.5 Tree homomorphisms versus homeomorphisms . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20

3 The space of algebraic measure trees 21

4 Triangulations of the circle 28

4.1 The space of sub-triangulations of the circle . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28

4.2 Coding binary measure trees with (sub-)triangulations of the circle . . . . . . . . . . . . . . . . . . . . 31

5 The subspace of binary algebraic measure trees 34

5.1 Convergence in distribution of sampled tree shapes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35

5.2 Convergence in distribution of sampled subtree masses . . . . . . . . . . . . . . . . . . . . . . . . . . . 37

5.3 Equivalence and compactness of topologies . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 41

6 Examples 42

A Appendix 44

1University of Duisburg-Essen, Mathematics, 45117 Essen, Germany
2Corresponding author, e-Mail: wolfgang.loehr@uni-due.de
3Research partially supported by DFG-RTG 2131
1 Introduction

In the recent years the construction and investigation of scaling limits of tree-valued Markov chains became more and more of interest. This started with the continuum analogs of the Aldous-Broder-algorithm for sampling a uniform spanning tree from the complete graph ([EPW06]) and of the tree-valued subtree-prune and regraft Markov chain used in the reconstruction of phylogenetic trees ([EW06]). It continued with the construction of evolving genealogies of infinite size populations in population genetics ([GPW13, DG12, Pio10, GSW16]) and in population dynamics ([Glo12, KW]). Moreover, continuum analogues of pruning procedures were constructed ([ADV10, AD12, LVW15, HWb, HWa]).

In order to provide a unified set-up which includes graph-theoretical trees on the one hand and continuum trees on the other, it became by now a classic approach to encode trees as metric (measure) spaces or bi-measure $\mathbb{R}$-trees and to equip the space of all metric (measure) trees with the Gromov-Hausdorff ([Gro99]), Gromov-weak ([Fuk87, Fuk87, GPW09, Löh13]) or Gromov-Hausdorff-weak ([Vil09, ADH13, ALW17]) topology and the space of all bi-measure $\mathbb{R}$-trees with the leaf-sampling weak-vague topology ([LVW15]). All these approaches have in common that they rely on encoding trees as metric spaces.

With the present paper we want to present a set-up which rather than focusing on pairwise metric distances brings the attention to the tree structure. We thereby want to respond to the observation that sometimes the intrinsic graph distance does not seem to be the right notion of distance. It might, for example, behave too wildly to allow the suitably rescaled family of tree-valued Markov chains to be tight, and consequently let the metric approach fail.

The goal of this paper is to overcome the metric issue by focusing on the tree structure only. We will introduce algebraic measure trees with an axiomatic approach, but it turns out that they can (under a separability constraint) be viewed as metric trees where one has “forgotten” the metric. To this end, we introduce an intrinsic metric which comes from the branch point distribution (Definition 3.5).

The starting point is the notion of an $\mathbb{R}$-tree ([Tit77, DMT96, Chi01, Eva08]). There are many equivalent definitions. We use the following:

**Definition 1.1 ($\mathbb{R}$-trees).** A metric space $(T, r)$ is an $\mathbb{R}$-tree iff it satisfies the following:

(RT1) $(T, r)$ satisfies the so-called 4-point condition, i.e., for all $x_1, x_2, x_3, x_4 \in T$,

$$r(x_1, x_2) + r(x_3, x_4) \leq \max \left\{ r(x_1, x_3) + r(x_2, x_4), r(x_1, x_4) + r(x_2, x_3) \right\}. \quad (1.1)$$

(RT2) $(T, r)$ is a connected metric space.

Notice that any metric space $(T, r)$ satisfying (RT1) and (RT2) admits a branch point map $c: T^3 \to T$, i.e., for all $x_1, x_2, x_3 \in T$ there exists a unique point $c(x_1, x_2, x_3) \in T$ such that

$$\{ c(x_1, x_2, x_3) \} = [x_1, x_2] \cap [x_1, x_3] \cap [x_2, x_3]. \quad (1.2)$$

where here for $x, y \in T$

$$[x, y] := \{ z \in T : r(x, z) + r(z, y) = r(x, y) \}, \quad (1.3)$$
1 INTRODUCTION

Figure 1: The only possible tree shape spanned by 4-points separates the four points into two pairs. Here \( r(x_1, x_2) + r(x_1, x_4) < \max\{r(x_1, x_3) + r(x_2, x_4), r(x_1, x_4) + r(x_2, x_3)\} \), while any other permutation yields equality. Here \( c_1 = c(x_1, x_2, x_3) = c(x_1, x_2, x_4) \) and \( c_2 = c(x_1, x_3, x_4) = c(x_2, x_3, x_4) \).

Figure 2: The graph shown here is not a tree, but the vertices satisfy the 4-point condition with respect to the graph-distance. Condition (MT2) fails.

or equivalently,

\[
[x, y] := \{ z \in T : c(x, y, z) = z \}.
\]  

(1.4)

While condition (RT1) is crucial for trees as it reflects the fact that there is only one possible shape for the subtree spanned by four points (as shown in Figure 1), the assumption of connectedness can be relaxed. In [ALW17], the notion of a metric tree was introduced as a metric space \((T, r)\) which can be embedded isometrically into an \(\mathbb{R}\)-tree such that it contains all branch points \(c(x_1, x_2, x_3), x_1, x_2, x_3 \in T\), as defined by (1.2). To exclude non-tree graphs satisfying the 4-point condition (see, for example, Figure 2), we have to require the property of containing the branch points explicitly.

**Definition 1.2** (Metric trees). A metric space \((T, r)\) is a metric tree if the following two conditions hold:

- (MT1) \((T, r)\) satisfies the 4-point condition (1.1).
- (MT2) \((T, r)\) admits all branch points, i.e., for all \(x_1, x_2, x_3 \in T\) there exists a (necessarily unique) \(c(x_1, x_2, x_3) \in T\) such that

\[
r(x_i, c(x_1, x_2, x_3)) + r(c(x_1, x_2, x_3), x_j) = r(x_i, x_j), \quad \forall i \neq j \in \{1, 2, 3\}.
\]  

(1.5)

Our main goal is to forget the metric while keeping the tree structure encoded by the branch point map. To axiomatize the latter, notice that for metric trees the branch point map satisfies the following obvious properties:

- (BPM1) The map \(c: T^3 \to T\) is symmetric.
- (BPM2) The map \(c: T^3 \to T\) satisfies the 2-point condition, i.e., for all \(x, y \in T\)

\[
c(x, y, y) = y.
\]  

(1.6)

- (BPM3) The map \(c: T^3 \to T\) satisfies the 3-point condition, i.e., for all \(x, y, z \in T\)

\[
c(x, y, c(x, y, z)) = c(x, y, z).
\]  

(1.7)
The map \( c: T^3 \to T \) satisfies the 4-point condition, i.e., for all \( x_1, x_2, x_3, x_4 \in T \),
\[
c(x_1, x_2, x_3) \in \{c(x_1, x_2, x_4), c(x_1, x_3, x_4), c(x_2, x_3, x_4)\}.
\] (1.8)

We therefore define the following:

**Definition 1.3 (Algebraic tree).** An algebraic tree \((T, c)\) consists of a set \( T \neq \emptyset \) and a branch point map \( c: T^3 \to T \) satisfying (BPM1)–(BPM4).

We want to consider algebraic trees as topological spaces where the topology is generated as follows: For each point \( x \in T \), we define an equivalence relation \( \sim_x \) on \( T \setminus \{x\} \) such that for all \( y, z \in T \setminus \{x\} \), \( y \sim_x z \) iff \( c(x, y, z) \neq x \). For \( y \in T \setminus \{x\} \), we denote by
\[
S_x(y) := \{z \in T : z \sim_x y\}
\] (1.9)
the equivalence class w.r.t. \( x \in T \) which contains \( y \). Clearly, any equivalence class \( S_x(y) \) corresponds to a subtree rooted at \( x \) (but not containing) \( x \) in the embedding \( \mathbb{R} \)-tree. We consider the topology generated by sets of the form \( \{1.9\} \) with \( x \neq y \) and denote by \( B(T, c) \) the corresponding Borel \( \sigma \)-algebra.

Our first main result (Theorem 1) relates metric trees with algebraic trees. On the one hand, if \( (T, r) \) is a metric tree, then it is clear that \( T \) together with the map \( c \) from (MT2) yields an algebraic tree. On the other hand, we shall show that every order separable algebraic tree, i.e., a separable tree with at most countably many edges, is induced by a metric tree in this way. More concretely, we will show that if \( \nu \) is a measure on \( B(T, c) \) which is finite and non-zero on non-degenerate intervals, i.e., on sets of the form
\[
[x, y] := \{z \in T : c(x, y, z) = z\}
\] (1.10)
for \( x, y \in T \), \( x \neq y \), then a metric representation of \((T, c)\) is given by
\[
r_\nu(x, y) := \nu([x, y]) - \frac{1}{2}\nu(\{x\}) - \frac{1}{2}\nu(\{y\}).
\] (1.11)

We next want to equip an algebraic tree \((T, c)\) with a sampling (probability) measure \( \mu \) on \( B(T, c) \). An algebraic measure tree \((T, c, \mu)\) consists of an algebraic tree \((T, c)\) and a probability measure \( \mu \) on \( B(T, c) \). Two algebraic measure trees \((T, c, \mu)\) and \((T', c', \mu')\) are equivalent if there are \( A \subseteq T \), \( A' \subseteq T' \) and a bijection \( \phi: A \to A' \) such that the following holds.

- \( \mu(A) = \mu'(A') = 1 \), \( c(A^3) \subseteq A \) and \( c'((A')^3) \subseteq A' \).
- \( \phi \) is measure preserving, and \( c'((\phi(x), \phi(y), \phi(z))) = \phi(c(x, y, z)) \) for all \( x, y, z \in T \).

We denote by \( \mathbb{T} \) the space of all equivalence classes of order separable algebraic measure trees. We shall equip \( \mathbb{T} \) with a notion of convergence based on the Gromov-weak topology. For that purpose, we introduce a particular metric representation of an algebraic measure tree. As metric representations are far from being unique, we will consider the intrinsic metric \( r_\nu \) which comes from the branch point distribution, i.e., the image measure \( \nu := c_*\mu^{\otimes 3} \) of \( \mu^{\otimes 3} \) under the branch point map \( c \). We declare that
\[
(T_n, c_n, \mu_n) \xrightarrow{n \to \infty} (T, c, \mu) \quad \text{iff} \quad (T, r_{c_n}^{(\nu)} \mu_n^{\otimes 3}, \mu_n) \to (T, r_{c_*\mu^{\otimes 3}}, \mu) \quad \text{Gromov-weakly},
\] (1.12)
and refer to this convergence as branch point distribution distance Gromov-weak convergence, or shortly, \( \text{bpdd-Gromov-weak convergence} \).

A particular subclass of interest is the space of binary algebraic measure trees. Similar to encoding compact \( \mathbb{R} \)-trees by a continuous excursion on the unit interval, binary algebraic trees can be encoded by \textit{triangulations of the circle} (see Figure 3). Such an encoding was introduced originally by David Aldous in \[\text{Ald94a, Ald94b}\], and there has since then been an increasing amount of research in the random tree community using this approach (e.g. \[\text{CLG11, BS15, CK15}\]). Also more general angulations and dissections have been considered which allow for encoding not necessarily binary trees (\[\text{CHK15}\]).

Aldous’s originally defines a triangulation of the circle as a closed subset of the disc the complement of which is a disjoint union of open triangles with vertices on the circle (see \[\text{Ald94b, Definition 1}\]). We modify Aldous’s definition in two respects. First, we add a condition which excludes non-tree graphs (such a condition is missing in Aldous’s definition) and under which triangulations of the circle are precisely the Hausdorff-metric limits of triangulations of \( n \)-gons. Second, we extend the definitions to so-called sub-triangulation of the circle (triangulations of a subset of the circle) which allow for encoding not only the algebraic tree but the measure on it in such a way that it is allowed to have point masses on leaves. In fact, any triangulation of the whole circle encodes a binary tree with a non-atomic measure which is relevant in the case of Aldous’s CRT. We then show that sub-triangulations of the circle can indeed be used to encode binary algebraic measure trees with point-masses restricted to the leaves. Furthermore, we show that – similar to the case of coding compact \( \mathbb{R} \)-trees by continuous excursions – the coding map that associates to a sub-triangulation of the circle the corresponding algebraic measure tree is surjective and \textit{continuous} when the set of sub-triangulations is equipped with the Hausdorff metric topology and the set of binary algebraic measure trees with bpdd-Gromov-weak topology (Theorem 2).

We also analyze the space of binary algebraic measure trees with point-masses restricted to the leaves in more detail. Our third main result (Theorem 3) states that this space is a compact, metrizable space under the bpdd-Gromov-weak topology. We also give two more notions on convergence. One is based on weak convergence of the \textit{tree-shapes} spanned by a finite sample. The other on weak convergence of the \textit{tensor of subtree-masses} read off the algebraic measure subtree spanned by a finite sample. It turns out that all three notions of convergence are equivalent on this subspace.
It is obvious that our results can be extended easily to trees with a bound other than three on the degree. However, in order to consider unbounded (or infinite) degrees additional care has to be taken. The reasons is that such subspaces of algebraic measure trees can not any longer expected to be compact.

Close relatives of algebraic measure trees have been recently studied in Forman [For]. The equivalence of bpdd-Gromov-weak topology and weak convergence of sample tree-shapes is related to the space of didentritic systems introduced recently by Evans, Grüber, and Wakolbinger in [EGW17]. Didentritic systems can be considered as ordered binary algebraic measure trees, and the space of didentritic systems is equipped with a kind of sample shape convergence.

Outline. The rest of the paper is organized as follows. In Section 2 we introduce our concept of algebraic trees by formalising the branch point map as a tertiary operation on the tree. We show that under a separability constraint algebraic trees can be seen as subtrees of metric trees, where the metric structure has been “forgotten” (Theorem 1).

In Section 3 we introduce the space of (equivalence classes of) order separable algebraic measure trees, and equip it with the Gromov-weak topology with respect to the metric associated with the branch point distribution.

In Section 4 we give a definition and characterisation of triangulations of the circle. We also formalize the notion of the algebraic (measure) tree associated with a given triangulation of the circle. This correspondence has often been pointed out in the literature but has never been made precise (except for discrete, graph-theoretic trees where it is more or less obvious). We show that the resulting coding map (which associates a triangulation of the circle with a tree) is well-defined and surjective onto the space of binary algebraic measure trees with non-atomic measure (Theorem 2).

In Section 5 we restrict ourselves to the subspace of binary, order separable, algebraic measure trees, and introduce two other, natural notions of convergence. We will use the construction of the coding map from Section 4 to show that on the subspace of binary algebraic measure trees, all three notions of convergence define the same topology (Theorem 3). This topology turns our space of binary algebraic measure trees into a compact, metrizable space. Furthermore, we show that the coding map is continuous if the space of triangulations is equipped with the Hausdorff metric topology, and the space of trees with the bpdd-Gromov-weak topology.

In Section 6 we consider the example of the continuum limit of so-called $\beta$-splitting trees introduced in [Ald96a].

2 Algebraic trees

In this section we introduce algebraic trees. In Subsection 2.1 we formalize the “tree-structure” common to both graph-theoretic trees and metric trees by a function that maps every triplet of points in the tree to the corresponding branch point. We show that the set of defining properties is rich enough to obtain known concepts such as leaves, branch points, degree, edges, intervals, subtrees spanned by a set, discrete and continuum trees, etc. In Subsection 2.2 we introduce the notion of structure preserving morphisms. In Subsection 2.3 we equip algebraic trees with a canonical Hausdorff topology. We also characterize compactness and a concept we call order separability, which is closely related to second countability of the topology. Finally, in Subsection 2.4 we show that any order separable algebraic tree
is induced by a metric tree (which is not true without order separability), and establish the condition under which this metric tree can be chosen to be a compact $\mathbb{R}$-tree.

### 2.1 The branch point map

In this subsection we introduce algebraic trees. Recall from Definition 2.2 the definition of a metric tree, and the properties (BPM1)–(BPM4) of the map which sends a triplet of 3 points in a metric tree to its branch point.

**Definition 2.1** (algebraic trees). An algebraic tree $(T, c)$ consists of a set $T \neq \emptyset$ and a branch point map $c: T^3 \to T$ satisfying (BPM1)–(BPM4).

The following useful property reflects the fact that any four points in an algebraic tree can be associated with a shape as illustrated in Figure 1 above.

**Lemma 2.2** (a consequence of (BPM4)). Let $(T, c)$ be an algebraic tree. Then for all $x_1, x_2, x_3, x_4 \in T$ the following hold:

(i) If $c(x_1, x_2, x_3) = c(x_1, x_2, x_4)$, then $c(x_1, x_3, x_4) = c(x_2, x_3, x_4)$.

(ii) If $c(x_1, x_2, x_3) = c(x_1, x_2, x_4)$, then $c(x_1, x_2, x_3) = c(x_1, x_2, c(x_1, x_3, x_4))$.

**Proof.** Let $x_1, x_2, x_3, x_4 \in T$ with $c_1 := c(x_1, x_2, x_3) = c(x_1, x_2, x_4)$, and $c_2 := c(x_1, x_3, x_4)$.

(i) Condition (BPM4) implies that

$$c_2 \in \{c_1 = c(x_1, x_3, x_2), c(x_2, x_3, x_4), c_1 = c(x_1, x_2, x_4)\}. \tag{2.1}$$

Thus $c_1 = c_2$, or $c_2 = c(x_2, x_3, x_4)$. The second case is the claim. In the first case, we apply Condition (BPM4) once more to find that

$$c(x_2, x_3, x_4) \in \{c_1 = c(x_1, x_2, x_3), c_2 = c(x_1, x_3, x_4), c_1 = c(x_1, x_2, x_4)\} = \{c_1, c_2\} = \{c_2\}, \tag{2.2}$$

so that the claim also holds in this case.

(ii) Condition (BPM3) implies that

$$c(x_1, x_3, c_2) = c(x_1, x_3, c(x_1, x_3, x_4)) = c(x_1, x_3, x_4) = c_2, \tag{2.3}$$

and similarly also $c(x_2, x_3, c_2) = c(x_2, x_3, x_4) = c_2$. Now part (i) with $x_4$ replaced by $c_2$ yields $c(x_1, x_2, x_3) = c(x_1, x_2, c_2)$ as claimed.

We have seen that the four axiomatizing properties of the branch point map are necessary. In many respects they are also sufficient to capture the tree structure. For example, in analogy to (LE3) we can define for each $x, y \in T$ the *interval* $[x, y]$ by

$$[x, y] := \{w \in T : c(x, y, w) = w\}. \tag{2.4}$$

We also use the notation $(x, y) := [x, y] \setminus \{x, y\}$, and similarly $(x, y), [x, y]$. The following properties of intervals are known to hold in $\mathbb{R}$-trees (compare, e.g., to [Chi01, Chapter 2] or [Eva08, Chapter 3]):

**Lemma 2.3** (properties of intervals). Let $(T, c)$ be an algebraic tree. Then the following hold:
(i) If \( x, v, w, z \in T \) are such that \( w \in [x, z] \) and \( v \in [x, w] \), then \( v \in [x, z] \).

(ii) If \( x, y, z \in T \), then
\[
[x, y] \cap [y, z] = [c(x, y, z), y].
\] (2.5)
In particular,
\[
[x, c(x, y, z)] \cap [c(x, y, z), z] = \{c(x, y, z)\}. \tag{2.6}
\]
(iii) If \( x, y, z \in T \), then
\[
[x, y] \cup [y, z] = [x, z] \cup (c(x, y, z), y].
\] (2.7)
In particular,
\[
[x, y] \cup [y, z] = [x, z] \iff y \in [x, z]. \tag{2.8}
\]
(iv) For all \( x, y, z \in T \),
\[
[x, y] \cap [y, z] \cap [z, x] = \{c(x, y, z)\}. \tag{2.9}
\]

Proof. (i) Let \( x, v, w, z \in T \) with \( w = c(x, w, z) \) and \( v = c(x, v, w) \). Then by Condition (BPM4),
\[
c(x, v, z) \in \{c(x, v, w), w = c(x, w, z), c(v, w, z)\}. \tag{2.10}
\]
We discuss the three cases separately. If \( c(x, v, z) = c(x, v, w) \), then \( c(v, w, z) = c(x, w, z) = w \) by Lemma 2.2(i). It then follows that \( c(x, v, z) = c(x, v, c(x, w, z)) = c(x, v, w) = v \) by Lemma 2.2(ii), which gives the claim in this case.

If \( c(x, v, z) = w \) then \( v = c(v, w, x) = c(v, w, z) \) by Lemma 2.2(i). It then follows that \( c(x, v, z) = c(x, z, c(z, w, v)) = c(x, v, v) = v \) by Lemma 2.2(ii), which gives the claim in this case.

If \( c(x, v, z) = c(v, w, z) \) then \( v = c(x, v, w) = c(x, w, z) = w \) by Lemma 2.2(i). Thus \( v = w \in [x, z] \), and the claim holds also in this case.

(ii) Let \( x, y, z \in T \), and \( v \in [x, y] \cap [y, z] \). That is, \( v = c(x, v, y) = c(y, v, z) \). It follows from Lemma 2.2(i) that \( c(x, z, v) = c(x, z, y) \), and then from Lemma 2.2(ii) together with Condition (BPM2) that
\[
v = c(x, v, y) = c(v, y, c(y, x, z)). \tag{2.11}
\]
Equivalently, \( v \in [c(x, y, z), y] \). This proves the inclusion \( [x, y] \cap [y, z] \subseteq [c(x, y, z), y] \). The other inclusion follows from (i).

Notice that (2.6) follows from (2.5) with the special choice \( y = c(x, y, z) \).

(iii) Notice first that it follows immediately from (i) that the union on the right hand side is disjoint. We claim that
\[
[x, z] \subseteq [x, y] \cup [y, z]. \tag{2.12}
\]
Indeed, let \( v \in [x, z] \), i.e. \( c(x, z, v) = v \). Then by (BPM4) applied to \( (v, x, y, z) \),
\[
v = c(x, z, v) \in \{c(x, y, v), c(x, y, z), c(y, z, v)\}, \tag{2.13}
\]
which implies that \( v \in [x, y] \) (if \( v = c(x, y, v) \)) or \( v \in [x, z] \cap [x, y] \) (if \( v = c(x, y, z) \)) or \( v \in [y, z] \) (if \( v = c(y, z, v) \)). Second, we claim that for all \( x, y, z \in T \),
\[
[x, z] \cup [c(x, y, z), y] \subseteq [x, y] \cup [y, z]. \tag{2.14}
\]
To see this, recall from (ii) that \( [c(x, y, z), y] = [x, y] \cap [z, y] \subseteq [x, y] \cap [z, y] \). As \( [x, c(x, y, z)] \subseteq [x, y] \) by (i), we have \( [x, y] \subseteq [x, c(x, y, z)] \cup [c(x, y, z), y] \subseteq [x, y] \cup (c(x, y, z), y] \). The corresponding inclusion for \( [y, z] \) is shown in the same way, and we have proven Equation (2.11).
(iv) This follows immediately from (ii). \(\square\)

We say that \(\{x, y\} \subseteq T\) with \(x \neq y\) is an edge of \((T, c)\) if and only if there is “nothing in between”, i.e. \([x, y] = \{x, y\}\), and denote by

\[
\text{edge}(T, c) := \{\{x, y\} \subseteq T : x \neq y, [x, y] = \{x, y\}\}
\]  

(2.15)

the set of edges. The following example explains that there is no need to distinguish between finite algebraic trees and graph-theoretical trees, and the definitions of edges are consistent.

**Example 2.4** (finite algebraic trees correspond to graph-theoretical trees). Finite algebraic trees are in one to one correspondence with finite (undirected) graph-theoretical trees. Let \((T, E)\) be a graph-theoretic tree with vertex set \(T\) and edge set \(E\). Then \((T, E)\) corresponds to the algebraic tree \((T, c_E)\) with \(c_E(u, v, w)\) defined as the unique vertex that is on the (graph-theoretic) path between any two of \(u, v, w\). Conversely, if \((T, c)\) is an algebraic tree with \(T\) finite, then \((T, c)\) corresponds to the graph-theoretic tree \((T, E_c)\) with \(E_c := \text{edge}(T, c)\). Obviously, \(c_{E_c} = c\).

For a graph-theoretic tree \((T, E)\), we can allow the vertex set \(T\) to be countably infinite, and still obtain a corresponding algebraic tree as in the previous example. Note, however, that countable algebraic trees do not necessarily correspond to graph-theoretical trees. Indeed, it is possible that \(T\) is countably infinite and \(\text{edge}(T, c) = \emptyset\). This can be seen by taking \(T = \mathbb{Q}\) in the following example, which shows that every totally ordered space naturally corresponds to an algebraic tree.

**Example 2.5** (totally ordered spaces as algebraic trees). For a totally ordered space \((T, \leq)\), define \(c_\leq(x, y, z) := y\) whenever \(x \leq y \leq z\), \((x, y, z) \in T\). Then it is trivial to check that \((T, c_\leq)\) is an algebraic tree and the interval \([x, y]\) coincides with the order interval \([z \in T : x \leq z \leq y]\). \(\diamondsuit\)

Conversely, given an algebraic tree \((T, c)\) and any fixed point \(\rho\) (often referred to as root), we can define a partial order \(\leq_\rho\) by letting for \(x, y \in T\),

\[
x \leq_\rho y \quad \text{iff} \quad x \in [\rho, y].
\]  

(2.16)

**Lemma 2.6** (algebraic trees as semi-lattices). Let \((T, c)\) be an algebraic tree, and \(\rho, x, y \in T\). Then \((T, \leq_\rho)\) is a partially ordered set, and a meet semi-lattice with infimum

\[
x \wedge y = c(\rho, x, y).
\]  

(2.17)

Furthermore, \(\leq_\rho\) is a total order on \([\rho, x]\) for all \(x \in T\).

**Proof.** Let \(x, y \in T\) with \(x \leq_\rho y\) and \(y \leq_\rho x\). That is, \(x = c(\rho, x, y)\) and \(y = c(\rho, y, x)\) which implies that \(x = y\), and proves that \(\leq_\rho\) is antisymmetric. As \(x = c(\rho, x, x)\), \(x \leq_\rho x\) which proves that \(\leq_\rho\) is reflexive. Finally, to show transitivity, let \(x, y, z \in T\) with \(x \leq_\rho y\) and \(y \leq_\rho z\). That is \(x \in [\rho, y]\) and \(y \in [\rho, z]\), which implies that \(x \in [\rho, z]\) by Lemma 2.3(i). Equivalently, \(x \leq_\rho z\) which proves the transience, and thus that \(\leq_\rho\) is a partial order.

For the infimum, notice that \(v \leq_\rho x\) and \(v \leq_\rho y\) if and only if \(v \in [\rho, x] \cap [\rho, y]\), or equivalently by Lemma 2.3(ii), \(v \in [\rho, c(\rho, x, y)]\). As for all \(v \in [\rho, c(\rho, x, y)]\) we have \(v \leq c(\rho, x, y)\), the claim (2.17) follows.
Fix $x \in T$. For totality on $[\rho, x]$, let $v, w \in [\rho, x]$, i.e., $v = c(\rho, v, x)$ and $w = c(\rho, w, x)$. Applying Condition (BPM4) to $\{\rho, v, w, x\}$ we find that one of the following three cases must occur: $c(\rho, v, w) = c(\rho, v, x)$ (which implies that $v = c(\rho, v, w)$, or equivalently, $v \leq \rho w$), $c(\rho, w, v) = c(\rho, w, x)$ (which implies that $w = c(\rho, w, v)$, or equivalently, $w \leq \rho v$), or $c(\rho, x, v) = c(\rho, x, w)$ (which implies that $w = v$).

**Corollary 2.7.** Let $(T, c)$ be an algebraic tree, and $\rho, x, y \in T$. If $v \in [x, y]$, then $v \geq c(x, v, y)$.

**Proof.** Let $\rho, x, y \in T$ and $v \in [x, y]$. That is, $v = c(x, v, y)$. We need to show that $c(\rho, v, c(\rho, x, y)) = c(\rho, x, y)$.

By Condition (BPM4) applied to $\{x, y, \rho, v\}$ we have one of the following three cases: $c(x, y, \rho) = c(x, y, v)$ (in which case $c(\rho, x, y) = c(\rho, v, y)$) or $c(\rho, y, x) = c(\rho, v, y)$ (in which case $c(x, v, \rho) = c(x, v, y) = v$ by Lemma 2.2(i) and thus $v \in [\rho, x]$; the claim then follows since this implies that $v \in [\rho, x] \cap [x, y] = [c(\rho, x, y), y]$ by Lemma 2.3(ii)), or $c(\rho, x, y) = c(\rho, x, v)$ (in which we conclude similar to the second case that $v \in [c(\rho, x, y), x]$).

The partial orders $\leq \rho$ allow us to define a notion of completeness of algebraic trees.

**Definition 2.8** (directed order completeness). Let $(T, c)$ be an algebraic tree. We call $(T, c)$ **(directed) order complete** if for all $\rho \in T$ the supremum of every totally ordered, non-empty subset exists in the partially ordered set $(T, \leq \rho)$.

Obviously, in an order complete algebraic tree, infima of totally ordered sets exists, because they are either $\rho$ if the set is empty or a non-empty supremum w.r.t. a different root. This notion of completeness allows us to define the analogs of complete $\mathbb{R}$-trees.

**Definition 2.9** (algebraic continuum tree). We call an algebraic tree $(T, c)$ **algebraic continuum tree** if the following two conditions hold:

1. **(ACT1)** $(T, c)$ is order complete.
2. **(ACT2)** $\text{edge}(T, c) = \emptyset$.

### 2.2 Morphisms of algebraic trees

Like any decent algebraic structure (or in fact mathematical structure), algebraic trees come with a notion of structure-preserving morphisms.

**Definition 2.10** (morphisms). Let $(T, c)$ and $(\hat{T}, \hat{c})$ be algebraic trees. A map $f : T \to \hat{T}$ is called a **tree homomorphism** (from $T$ into $\hat{T}$) if for all $x, y, z \in T$,

$$f(c(x, y, z)) = \hat{c}(f(x), f(y), f(z)).$$

We refer to a bijective tree homomorphism as **tree isomorphism**.

As we have seen that the tree structure can be expressed also in terms of intervals or partial orders rather than the branch point map, and we obtain the following equivalences.

**Lemma 2.11** (equivalent definitions). Let $(T, c)$ and $(\hat{T}, \hat{c})$ be algebraic trees, and $f : T \to \hat{T}$. Then the following are equivalent:
1. \( f \) is a tree homomorphism.

2. For all \( \rho \in T \), \( f \) is an order preserving map from \((T, \leq_{\rho})\) to \((\hat{T}, \leq_{f(\rho)})\).

3. For all \( x, y \in T \), \( f([x, y]) \subseteq [f(x), f(y)] \).

Proof. 

\( 1 \Rightarrow 2 \). Let \( x, y, \rho \in T \) with \( x \leq_{\rho} y \). Then \( x = c(\rho, x, y) \) and thus \( f(x) = \hat{c}(f(\rho), f(x), f(y)) \). Therefore \( f(x) \leq_{f(\rho)} f(y) \).

\( 2 \Rightarrow 3 \). Let \( x, y, z \in T \) with \( z \in [x, y] \). Then \( z \leq_{x} y \) and thus \( f(z) \leq_{f(x)} f(y) \), i.e. \( f(z) \in [f(x), f(y)] \).

\( 3 \Rightarrow 1 \). Let \( x, y, z \in T \). Then \( \{c(x, y, z)\} = [x, y] \cap [x, z] \cap [y, z] \). Hence

\[
\{f(c(x, y, z))\} \subseteq \{f(x), f(y)\} \cap \{f(y), f(z)\} \cap \{f(x), f(z)\} = \{\hat{c}(f(x), f(y), f(z))\}. \tag{2.19}
\]

Therefore, \( f(c(x, y, z)) = \hat{c}(f(x), f(y), f(z)) \).

The image of an algebraic tree under a homomorphism is a subtree in the following sense.

**Definition 2.12** (subtree). Let \((T, c)\) be an algebraic tree, and \( \emptyset \neq A \subseteq T \). \( A \) is called a **subtree (of \((T, c))\)** if

\[
c(A^3) \subseteq A. \tag{2.20}
\]

We refer to \( c(A^3) \) as the **algebraic subtree generated by \( A \)**.

Obviously, a subtree \( A \) of \((T, c)\), implicitly equipped with the restriction of \( c \) to \( A^3 \), is an algebraic tree in its own right. Furthermore, the following lemma is easy to check.

**Lemma 2.13** (tree homomorphisms). Let \((T, c)\) and \((\hat{T}, \hat{c})\) be two algebraic trees, and \( f : T \to \hat{T} \) a homomorphism. Then the image \( f(T) \) is a subtree of \( \hat{T} \). If \( f \) is injective, \( f^{-1} \) is a tree homomorphism from \( f(A) \) into \( T \).

In particular, if \((\hat{T}, \hat{c})\) is another algebraic tree, and \( g \) is a homomorphism from \((\hat{T}, \hat{c})\) to \((T, c)\), then \( g \circ f \) is a homomorphism from \((\hat{T}, \hat{c})\) to \((T, c_\hat{T})\).

### 2.3 Algebraic trees as topological spaces

In contrast to metric trees, there is a priori no topology defined on a given algebraic tree. In this section, we therefore equip algebraic trees with a canonical topology.

For each \( x \in T \), we introduce a (component) relation \( \sim_x \) by letting \( y \sim_x z \) if and only if \( x \notin [y, z] \), where \( y, z \in T \). Let for each \( y \in T \setminus \{x\} \)

\[
S_x(y) = S_x^{(T, c)}(y) := \{ z \in T \setminus \{x\} : z \sim_x y \} \tag{2.21}
\]

be the equivalence class of \( T \setminus \{x\} \) containing \( y \), and note that \( S_x(y) \) is a subtree for all \( x, y \in T \), and \( S_x(y) = S_x(z) \) whenever \( z \in S_x(y) \). We refer to \( S_x(y) \) as the **component** of \( T \setminus x \) containing \( y \). Now and in the following, we equip \((T, c)\) with the topology

\[
\tau := \tau(\{S_x(y) : x, y \in T, x \neq y\}) \tag{2.22}
\]

generated by the set of components, i.e. with the coarsest topology such that all components are open sets. We call \( \tau \) the **component topology** of \((T, c)\).
Example 2.14 (on totally ordered trees, \( \tau \) is the order topology). If \((T, \leq)\) is a totally ordered space, and \((T, c)\) the corresponding algebraic tree as in Example 2.15, then \( \tau \) coincides with the order topology (i.e. the one generated by sets of the form \( \{ y \in T : y > x \} \) and \( \{ y \in T : y < x \} \) for \( x \in T \)).

Example 2.15 (intervals are closed sets). Let \((T, c)\) be an algebraic tree, and \( x, y \in T \). Then

\[
T \setminus [x, y] = \bigcup \{ S_u(v) : u \in [x, y], v \in T, S_u(v) \cap [x, y] = \emptyset \} \in \tau. \tag{2.23}
\]

This means that \([x, y]\) is closed in the component topology \( \tau \).

Lemma 2.16. Let \((T, c)\) be an algebraic tree. Then \( c \) is continuous w.r.t. the component topology \( \tau \).

Proof. By definition of \( \tau \), it is sufficient to show that for any \( x, y \in T, x \neq y \), the set \( c^{-1}(S_x(y)) \) is open in \( T^3 \). By definition of \( S_x(y) \) and the property \( c(u, v, w) \in [u, v] \cap [v, w] \cap [w, u] \) shown in Lemma 2.13, \( c(u, v, w) \in S_x(y) \) if and only if (at least) two of \( u, v, w \) are in \( S_x(y) \). Because \( S_x(y) \) is open, the same is true for \( \{(u, v, w) \in T^3 : u, v, w \in S_x(y)\} \) in the product topology. Hence \( c^{-1}(S_x(y)) \) is a union of open set and thus open.

Next, we show that \( \tau \) is a Hausdorff topology and characterize compactness of algebraic trees in this topology.

Lemma 2.17 (\( \tau \) is Hausdorff). Let \((T, c)\) be an algebraic tree. Then the component topology \( \tau \) defined in (2.22) is a Hausdorff topology on \( T \).

Proof. To show that \((T, \tau)\) is Hausdorff, let \( x, y \in T \) be distinct. If \( S_y(x) \cap S_x(y) = \emptyset \), then \( S_y(x) \) and \( S_x(y) \) are clearly disjoint neighbourhoods of \( x \) and \( y \), respectively. Assume that this is not the case, and choose \( z \in S_x(y) \cap S_y(x) \). Then \( \rho := c(x, y, z) \notin \{x, y\} \). Furthermore, \( c(x, x, y) = c(x, y, z) = \rho \), and hence \( x \not\sim \rho \ y \). Thus \( S_x(x) \) and \( S_y(y) \) are disjoint neighbourhoods of \( x \) and \( y \), respectively. Hence \( \tau \) is Hausdorff.

Proposition 2.18 (characterizing compactness). Let \((T, c)\) be an algebraic tree with component topology \( \tau \). Then \((T, \tau)\) is compact if and only if \((T, c)\) is directed order complete.

Proof. “only if”. Assume first that \((T, c)\) is not order complete. Then we can choose \( \rho \in T \) and \( \emptyset \neq A \subseteq T \) such that \( A \) is totally ordered w.r.t. \( \leq \rho \) but does not have a supremum in \((T, \leq \rho)\). For \( x, y \in T \), let \( U_x := \{ z \in T : z \not\sim \rho x \} \) and \( V_y := \{ z \in T : z \not\sim \rho y \} \). Then \( U_x \) and \( V_y \) are open sets. We claim that \( U := \{ U_x : x \in A \} \cup \{ V_y : y \geq A \} \) is an open cover of \( T \). Indeed, if \( z \geq \rho A \), then, because \( A \) has no supremum, there is \( y \in T \) with \( A \not\leq \rho y \leq z \), hence \( z \in V_y \in U \). Otherwise, if \( z \not\sim \rho A \), there is \( z \in A \) with \( z \in U_x \in U \). Thus \( U \) is a cover of \( T \).

\( U \) has no finite sub-cover, because if \( U' = \{ U_{x_1}, ..., U_{x_n}, V_{y_1}, ..., V_{y_m} \} \) were such a finite sub-cover, then \( \{U_{x_1}, ..., U_{x_n}\} \) would cover \( A \). This, however, would imply that \( \max\{x_1, ..., x_n\} \) would be a supremum of \( A \), contradicting our assumption. Hence \((T, \tau)\) is not compact.

“if”. Assume that \((T, c)\) is order complete. Consider a cover \( U \) of \( T \) with components, i.e. \( U \subseteq \{ S_y(x) : x, y \in T, x \neq y \} \). By the Alexander subbase theorem, for compactness of \( \tau \), it is sufficient to show that \( U \) has a finite sub-cover.

To this end, fix an element \( \rho \in T \) and consider the set \( U_\rho := \{ U \in U : \rho \in U \} \neq \emptyset \). By Hausdorff’s maximal chain theorem (or Zorn’s lemma), there is a maximal chain \( I \) in the
partially ordered set \((U_\rho, \subseteq)\). For every \(U \in I\), we have \(\rho \in U\), and thus there is \(x_U \in T\) such that \(U = S_{x_U}(\rho)\). We claim that \(U \subseteq V\) implies \(x_U \leq_\rho x_V\). Indeed, \(x_V \notin V\) and hence \(x_V \notin U\) which is equivalent to \(x_V \geq_\rho x_U\). Therefore, \(z := \sup\{x_U : U \in I\}\) exists in \((T, \leq_\rho)\) by directed order completeness of \(T\). Because \(U\) is a cover, there is \(V \in U\) with \(z \in V\), hence \(V = S_y(z)\) for some \(y \in T\). Because \(V \notin I\) and \(I\) is a maximal chain, \(y \not\geq_\rho z\). Hence there is \(U \in I\) with \(y \not\leq_\rho x_U := x\). We claim that \(T = S_y(z) \cup S_x(\rho)\). Indeed, let \(w \in T \setminus S_x(\rho)\). Then \(w \geq_\rho x\). Using \(z \geq_\rho x\) and \(c(w, z, y) \in [w, z]\), we obtain \(c(w, z, y) \geq_\rho x\), and hence \(c(w, z, y) \neq y\). Thus \(w \in S_y(z)\) as claimed, and \(\{S_y(z), S_x(\rho)\}\) is the desired sub-cover.

It turns out that the following separability condition, which we will call order separability, is crucial for us.

**Proposition 2.19** (order separability). Let \((T, c)\) be an algebraic tree with component topology \(\tau\). Then the following are equivalent:

1. There exists a countable set \(D\) such that for all \(x, y \in T\) with \(x \neq y\),
   \[D \cap \{x, y\} \neq \emptyset.\]  \tag{2.24}
2. The topological space \((T, \tau)\) is second countable (i.e. \(\tau\) has a countable base), and \(\text{edge}(T, c)\) is countable.
3. The topological space \((T, \tau)\) is separable, and \(\text{edge}(T, c)\) is countable.

**Proof.** \[\implies \] Assume that \(\text{edge}(T, c)\) is countable, and that \((T, \tau)\) is separable. Then there exists a countable, dense subset \(\tilde{D} \subseteq T\). We claim that
\[D := c(\tilde{D}^3) \cup \{z \in T : \exists x \in T\text{ such that }\{x, z\} \in \text{edge}(T, c)\}\]  \tag{2.25}
satisfies \((2.24)\). Indeed, \(D\) is countable by assumption. Moreover, let \(x, y \in T\). Then two cases are possible: either \(S_x(y) \cap S_y(x) = \emptyset\). In this case, \(\{x, y\} \in \text{edge}(T, c)\), which implies that \([x, y] \cap D \neq \emptyset\). Or \(S_x(y) \cap S_y(x) \neq \emptyset\). In this case, as \(S_x(y) \cap S_y(x)\) is open by definition of \(\tau\), there is \(z \in D \cap S_x(y) \cap S_y(x)\). Let \(v := c(x, y, z)\). Then \(v \in (x, y)\), and either \(v = z \in D\), or the three components \(S_v(x), S_v(y), S_v(z)\) are distinct. In the second case, we can choose \(x' \in D \cap S_v(x)\) and \(y' \in D \cap S_v(y)\) to see that \(v = c(x', y', z) \in D\). In any case, \(v \in [x, y] \cap D\).

\[\implies \] Let \(D\) be a countable set satisfying \((2.24)\). Then for all \(\{x, y\} \in \text{edge}(T, c)\),
\[D \cap \{x, y\} = \{x\}\]. This implies that \(\text{edge}(T, c)\) is countable. We consider the countable set \(U = \{S_v(u) : u, v \in D\} \subseteq \tau\) and claim that it is a subbase for \(\tau\) (i.e. generates \(\tau\)). To this end, let \(x, y \in T\). We show that \(U := S_x(y)\) is a union of sets from \(U\), i.e. for every \(z \in U\) we construct \(V \in U\) with \(z \in V \subseteq U\). By assumption on \(D\), there is \(v \in D \cap [x, z]\) and \(u \in D \cap [v, z]\). Let \(V := S_v(u) \in U\). Because \(c(u, v, z) = u \neq v\), we have \(z \in V\). Let \(w \in T \setminus U\). Because \(w \in U\), we have \(U = S_w(u)\) and therefore \(x \in [u, w]\). Similarly, \(x \in [v, w]\). In particular, by Lemma 2.2, \(c(u, v, w) = c(u, v, c(u, x, w)) = c(u, v, x) = v\), and thus \(w \notin V\).

\[\implies \] Trivial, because every second countable topological space is separable. \qed

**Definition 2.20** (order separability). We call an algebraic tree \((T, c)\) order separable if the equivalent conditions of Proposition 2.19 are satisfied. We call a set \(D \subseteq T\) order dense if it satisfies \((2.24)\).
Example 2.21 (uncountable star tree). This example shows that in (2.24) we can not replace

\([x, y]\) by \([x, y]^{\prime}\). Let \(T := \{0\} \cup [1, 2]\) with \(c(x, y, z) := 0\) whenever \(x, y, z \in T\) are distinct. Then if \(D \subseteq T\) is such that \(D \cap [x, 0] \neq \emptyset\) for all \(x \in [1, 2]\) then \([1, 2] \subseteq D\), and thus \(D\) is uncountable and \((T, c)\) not order separable. On the other hand, \(D := \{0\}\) has the property that \(D \cap [x, y] \neq \emptyset\) for all \(x, y \in T\) with \(x \neq y\).

An order complete, order separable algebraic tree is, in its component topology \(\tau\), a compact, second countable Hausdorff space by Propositions 2.18 and 2.19. In particular, it is metrizable. In fact, order separability already implies metrizability, as we will see in Subsection 2.4. The following example shows that (topological) separability of \((T, \tau)\) alone, without requiring the number of edges to be countable, is neither sufficient for order separability nor for metrizability of \((T, \tau)\).

Example 2.22 (a continuum ladder). Let \(T = [0, 1] \times \{0, 1\}\) with the lexicographic order \(\leq\) on \(T\), and define the canonical branch point map \(c_{\leq}\) as in Example 2.5. Then \(\text{edge}(T, c_{\leq}) = \{(x, 0), (x, 1) : x \in [0, 1]\}\) is uncountable, and hence \((T, c_{\leq})\) is not order separable. Because \((\mathbb{Q} \cap [0, 1]) \times \{0, 1\}\) is a countable dense set, \((T, \tau)\) is (topologically) separable. The topological subspace \([0, 1] \times \{1\}\) is the Sorgenfrey line, which is known to be non-metrizable (see [SS78, Counterexample 51]). Thus also \((T, \tau)\) cannot be metrizable.

Definition 2.23 (Borel \(\sigma\)-algebra \(B(T, c)\)). Let \((T, c)\) be an algebraic tree. We denote the Borel \(\sigma\)-algebra of the component topology \(\tau\) by \(B(T, c)\) and call it Borel \(\sigma\)-algebra of \((T, c)\).

In general, \(B(T, c)\) is not generated by the set of components. Order separability, however, is sufficient to ensure this property because it implies second countability of the component topology.

Corollary 2.24 (Borel \(\sigma\)-algebra generated by components). Let \((T, c)\) be an order separable algebraic tree, and \(D \subseteq T\) an order dense set. Then its Borel \(\sigma\)-algebra is generated by the set of components indexed by \(D\), i.e.

\[
B(T, c) = \sigma\left(\{S_{D}(y) : x, y \in D, x \neq y\}\right). \tag{2.26}
\]

Proof. Define \(U := \{S_D(y) : x, y \in D, x \neq y\}\). By Proposition 2.19, \((T, \tau)\) is second countable. Hence \(B(T, c)\) is generated by any subbase of \(\tau\). If \(D\) is order dense, \(U\) is such a subbase as shown in the proof of Proposition 2.19.

2.4 Metric tree representations of algebraic trees

In this subsection, we discuss the connection of metric trees with algebraic trees. Let \((T, r)\) be a metric tree (recall from Definition 1.2). Then by (MT2), there exists to any three points \(x_1, x_2, x_3 \in T\) a unique branch point \(c_{(T, r)}(x_1, x_2, x_3)\) satisfying (1.5). We refer to \((T, c_{(T, r)})\) as the algebraic tree induced by \((T, r)\), and to \((T, r)\) as a metric representation of \((T, c_{(T, r)})\).

Lemma 2.25 (the algebraic tree induced by a metric tree). Let \((T, r)\) be a metric tree, and \(c_{(T, r)}\) the map which sends any three distinct points to their branch point. Then the following hold:

(i) \((T, c_{(T, r)})\) is an algebraic tree.

(ii) \((T, c_{(T, r)})\) is order separable if and only if \((T, r)\) is separable.
(iii) \((T, c_{(T,r)})\) is directed order complete if and only if \((T, r)\) is bounded and complete. In particular, it is an algebraic continuum tree if and only if \((T, r)\) is a bounded, complete \(\mathbb{R}\)-tree.

**Proof.** (i) It can be easily checked that \((T, c_{(T,r)})\) is an algebraic tree.

(ii) Let \((T, r)\) be separable. Then \(\text{edge}(T, c_{(T,r)})\) is countable. The topology induced by \(r\) is obviously stronger than the topology \(\tau\) introduced in (2.22), hence \(\tau\) is separable and therefore the algebraic tree \((T, c_{(T,r)})\) is order separable. Conversely, if \((T, c_{(T,r)})\) is order separable, then any countable set \(D\) satisfying (2.24) is also dense in \((T, r)\).

(iii) Clearly, \((T, c_{(T,r)})\) admits suprema along any linearly ordered set with respect to some root if and only if \((T, r)\) is bounded and complete. The “in particular” follows because a complete metric tree \((T, r)\) is an \(\mathbb{R}\)-tree if and only if \(\text{edge}(T, c_{(T,r)}) = \emptyset\) \([ALW17, \text{Remark 1.2}]\).

Our first main result states that under the assumption of order separability any algebraic tree can be embedded by an injective homomorphism into a compact \(\mathbb{R}\)-tree and hence is isomorphic to (the algebraic tree induced by) a totally bounded metric tree.

**Theorem 1** (characterisation of order separable algebraic trees). Let \(T\) be a set, \(c: T^3 \to T\).

(i) \((T, c)\) is an order separable algebraic continuum tree if and only if there exists a metric \(r\) on \(T\) such that \((T, r)\) is a compact \(\mathbb{R}\)-tree with

\[ c = c_{(T,r)}. \]  

(ii) \((T, c)\) is an order separable algebraic tree if and only if there is an order separable algebraic continuum tree \((\overline{T}, \overline{c})\) such that \((T, c)\) is a subtree of \((\overline{T}, \overline{c})\). In particular, every order separable algebraic tree is induced by a totally bounded metric tree.

The separability hypothesis in Theorem 1 is crucial and cannot be dropped. In Example 2.22, we have already seen an algebraic tree where the component topology \(\tau\) is not metrizable. Moreover, in this example, \(\tau\) coincides with the order topology which is also the case for the metric topology of any metric tree without branch points. Thus the algebraic tree cannot be induced by a metric tree. The following example shows that also algebraic continuum trees need not be induced by metric trees.

**Example 2.26** (algebraic continuum tree that is not induced by a metric tree). Let \(T = [0, 1] \times [0, 1]\) with lexicographic order, and \((T, c)\) the corresponding algebraic tree as in Example 2.5. It is easy to check that \((T, c)\) is an algebraic continuum tree. It cannot be induced by a metric tree because in its order topology \(\tau\), it is connected but not path-wise connected. These two properties are equivalent for metric trees (see \[Eva08, \text{Theorem 2.20}]\).

In order to prove Theorem 1, given an algebraic tree \((T, c)\), we need to provide a metric \(r\) such that (2.27) holds. For that purpose, we consider for any measure \(\nu\) on \((T, \mathcal{B}(T,c))\) such that \(\nu\) is finite on every interval, the following pseudometric,

\[ r_\nu(x, y) := \nu([x, y]) - \frac{1}{2} \nu(\{x\}) - \frac{1}{2} \nu(\{y\}), \quad x, y \in T. \]  

(2.28)
Lemma 2.27 \((r_\nu)\) is a pseudometric. Let \((T, c)\) be an algebraic tree, and \(\nu\) a measure on \((T, c)\) with \(\nu([x, y]) < \infty\) for all \(x, y \in T\). Then \(r_\nu\) is a pseudometric on \(T\).

Proof. By Lemma 2.25 for all \(x, y, z \in T\),
\[
\nu([x, y]) + \nu([y, z]) = \nu([x, y] \cup [y, z]) + \nu([x, y] \cap [y, z]) \\
= \nu([x, z]) + \nu([c(x, y, z), y]) + \nu([c(x, y, z), y]).
\] (2.29)

Hence
\[
r_\nu(x, y) + \frac{1}{2} \nu(x) + \frac{1}{2} \nu(y) = r_\nu(x, z) + \frac{1}{2} \nu(x) + \frac{1}{2} \nu(z)
\]
\[
= r_\nu(x, z) + \frac{1}{2} \nu(x) + \frac{1}{2} \nu(z) + 2r_\nu(c(x, y, z), y) + \nu(c(x, y, z)) + \nu(y) - \nu(c(x, y, z)),
\] (2.30)
or equivalently,
\[
r_\nu(x, y) + r_\nu(y, z) = r_\nu(x, z) + 2r_\nu(c(x, y, z), y).
\] (2.31)

This implies that \(r_\nu\) satisfies the triangle inequality. 

We denote the quotient metric space by \((T_\nu, r_\nu)\), i.e. \(T_\nu\) is the set of equivalence classes of points in \(T\) with \(r_\nu\)-distance zero, and the quotient metric on \(T_\nu\) is again denoted by \(r_\nu\). Furthermore, let \(\pi_\nu: T \to T_\nu\) be the canonical projection.

Lemma 2.28 \((T_\nu, r_\nu)\) is a metric tree. Let \((T, c)\) be an algebraic tree, and \(\nu\) a measure on \((T, c)\) with \(\nu([x, y]) < \infty\) for all \(x, y \in T\). Then the quotient space \((T_\nu, r_\nu)\) is a metric tree, and the canonical projection \(\pi_\nu\) is a tree homomorphism.

Proof. Let \(x_1, \ldots, x_4 \in T\). By Condition (BPM4), we can assume w.l.o.g. that \(c(x_1, x_2, x_3) = c(x_2, x_3, x_4)\). Then by Lemma 2.2(ii), \(c(x_1, x_2, x_3) \in [x_1, x_2] \cap [x_1, x_3] \cap [x_2, x_3] \cap [x_1, x_4] \cap [x_2, x_4]\), and (2.31) yields that for \(i, j \in \{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\),
\[
r_\nu(x_i, x_j) = r_\nu(x_i, c(x_1, x_2, x_3)) + r_\nu(c(x_1, x_2, x_3), x_j).
\] (2.32)

Therefore,
\[
r_\nu(x_1, x_3) + r_\nu(x_2, x_4) \\
= r_\nu(x_1, c(x_1, x_2, x_3)) + r_\nu(c(x_1, x_2, x_3), x_3) + r_\nu(c(x_1, x_2, x_3), x_4) \\
= r_\nu(x_1, x_2) + r_\nu(c(x_1, x_2, x_3), x_3) + r_\nu(c(x_1, x_2, x_3), x_4) \\
\geq r_\nu(x_1, x_2) + r_\nu(x_3, x_4),
\] (2.33)

and analogously,
\[
r_\nu(x_1, x_4) + r_\nu(x_2, x_3) = r_\nu(x_3, c(x_1, x_2, x_3)) + r_\nu(c(x_1, x_2, x_3), x_4) + r_\nu(x_1, x_2) \\
\geq r_\nu(x_1, x_2) + r_\nu(x_3, x_4).
\] (2.34)

This means that the four point Condition (MT\(\square\)) is satisfied. Moreover, (2.32) implies Condition (MT\(\spadesuit\)) with branch point \(\pi_\nu(c(x_1, x_2, x_3))\). In particular, \(\pi_\nu\) is a tree homomorphism. 

Remark 2.29. Lemma 2.28 also explains why we had defined \(r_\nu\) as in (2.28) and not just as \(r_\nu' := \nu([x, y])\) for \(x \neq y\). Namely, in the latter case we would still have (MT\(\square\)), but (MT\(\spadesuit\)) might fail. Take, for example, \(T := \{1, 2, 3\}, c(1, 2, 3) = 2, \) and \(\nu = \delta_2\). In this case, \(r_\nu'\) is the discrete metric on \(T\), thus \(2\) does not lie on the interval \([1, 3]\) anymore. 

\(\diamond\)
Let \((T, c)\) be an algebraic tree. For all \(v \in T\), define the degree of \(v\) in \((T, c)\) by
\[
\text{deg}(v) := \text{deg}_{(T, c)}(v) := \#\{S_v(y) : y \in T\}.\tag{2.35}
\]
We say that \(v \in T\) is a leaf if \(\text{deg}_{(T, c)}(v) = 1\), and a branch point if \(\text{deg}_{(T, c)}(v) \geq 3\). Notice that
\[
\text{If}(T, c) := \{u \in T : c(u, v, w) \neq u \forall v, w \in T \setminus \{u\}\}
\]
equals the set of leaves of \(T\), and
\[
\text{br}(T, c) := \{u \in T : c(x, v, w) = u \text{ for some } x, v, w \in T \setminus \{u\}\}
\]
the set of branch points. Moreover, note that any \(\nu\)-mass on \(\text{If}(T, c)\) that is not atomic does not contribute to \(r_\nu\).

**Proposition 2.30** (metric representations of algebraic trees). Let \((T, c)\) be an algebraic tree, \(\nu\) a measure on \((T, \mathcal{B}(T, c))\) with \(\nu([x, y]) < \infty\) for all \(x, y \in T\), and \(r_\nu\) defined by (1.11). Then the following hold:

(i) If \((T, c)\) is order separable and \(\nu\) has at most countably many atoms, then \((T_\nu, r_\nu)\) is separable.

(ii) If \(\# T > 1\), \((T, c)\) is order complete, and \([x, y]\) is order separable for every \(x, y \in T\), then \((T_\nu, r_\nu)\) is connected if and only if \(\nu\) is non-atomic. In this case, \((T_\nu, r_\nu)\) is a complete \(\mathbb{R}\)-tree.

**Proof.** Throughout the proof denote by \(\pi_\nu : T \to T_\nu\) the canonical projection.

(i) It is easy to see that if a set \(A \subseteq T\) satisfies (2.24) and contains all atoms of \(\nu\), then \(\pi_\nu(A)\) is dense in \((T_\nu, r_\nu)\). Therefore, by Proposition 2.19 order separability of \((T, c)\) implies separability of \((T_\nu, r_\nu)\).

(ii) For all \(x, y \in T\) with \(x \neq y\), \(r_\nu(x, y) \geq \frac{1}{2}\nu\{x\}\). Hence \((T_\nu, r_\nu)\) cannot be connected if \(\nu\) has atoms. Conversely, assume that \(\nu\) is non-atomic. For \(x, z \in T\), consider \([|x, z|], \leq_{\nu}\), which is a totally ordered space according to Lemma 2.6, and define \(y := \sup\{v \in [x, z] : 2\nu([x, v]) \leq \nu([x, z])\}\). The supremum exists due to order completeness of \((T, c)\). Because of the order separability of \([x, z]\) and the non-atomicity of \(\nu\), we obtain \(2\nu([x, y]) = \nu([x, z]) = 2\nu([y, z])\) and therefore \(2r_\nu(x, y) = r_\nu(x, z) = 2r_\nu(y, z)\). From this equality, connectedness follows once we have shown completeness, and every connected metric tree is an \(\mathbb{R}\)-tree.

Recall from Lemma 2.28 that \((T_\nu, r_\nu)\) is a metric tree. The same holds for its metric completion \(\overline{T}_\nu\). Assume for a contradiction that there is a sequence \((x_n)_{n \in \mathbb{N}}\) in \(T_\nu\) converging to some \(x \in \overline{T}_\nu \setminus T_\nu\). Then \(x\) cannot be a branch point and one of the at most two components of \(\overline{T}_\nu \setminus \{x\}\) contains infinitely many \(x_n\). Thus we may assume w.l.o.g. that \(x \in \text{If}(\overline{T}_\nu)\). Define \(y_n := c_{\overline{T}_\nu}(x_1, x_n, x)\). Then \(y_n \to x\) and, for large enough \(m\), we have \(y_n = c_{\overline{T}_\nu}(x_1, x_m, x_m)\).

Hence \(y_n \in T_\nu\) for all \(n \in \mathbb{N}\) and we may choose representatives \(x'_n \in \pi_\nu^{-1}(y_n)\) such that \(x'_n = c(x, x'_m, x'_m)\) for \(\rho := x'_n\) and all sufficiently large \(m\). By Lemma 2.6 \(\{x'_n : n \in \mathbb{N}\}\) is totally ordered w.r.t. \(\leq_\rho\), and hence \(x' := \sup\{x'_n : n \in \mathbb{N}\} \in T\) exists by order completeness. Obviously, \(\pi_\nu(x') = x\) and \(x \in T_\nu\).

\(\square\)

In order to prove Theorem 1.11 using Proposition 2.30 we need a non-atomic probability measure \(\nu\) (to ensure connectedness of \((T_\nu, r_\nu)\)) charging all intervals (so that \(\pi_\nu\) is injective). Such a measure always exists in the case of order separable algebraic continuum trees.
Lemma 2.31. Let \((T, c)\) be an order separable algebraic continuum tree with \(\#T > 1\). Then there exists a non-atomic probability measure \(\nu\) on \((T, \mathcal{B}(T, c))\) with \(\nu(\mathbb{I}(T, c)) = 0\) and
\[
\nu([x, y]) > 0 \quad \forall x, y \in T, \ x \neq y.
\] (2.38)

Proof. Fix \(\rho \in T\). Then, for every \(x \in T \setminus \{\rho\}\), the interval \([\rho, x]\) is a separable linear continuum in the sense of order theory, i.e., a totally ordered space (proven in Lemma 2.6 without jumps (what we call here edges) or gaps (what follows from directed order completeness)). Due to Cantor’s order characterisation of \(\mathbb{R}\) (e.g. [Das14, Theorem 560]), this means that \([\rho, x]\) is order isomorphic to the unit interval. Obviously, every order isomorphism is measurable and bijective, and the image of Lebesgue measure on the unit interval is a non-atomic probability measure \(\nu_x\) on \([\rho, x]\). Then \(\sum_{n \in \mathbb{N}} 2^{-n}\nu x_n\), where \(\{x_n : n \in \mathbb{N}\}\) satisfies (2.24), is a non-atomic probability satisfying (2.38) and \(\nu(\mathbb{I}(T, c)) = 0\).

Any separable \(\mathbb{R}\)-tree \((T, r)\) comes with an intrinsic measure, called length measure, that generalizes the Lebesgue measure on \(\mathbb{R}\). More generally, if \((T, r)\) is a complete, separable metric tree and \(\rho \in T\) a fixed root, the length measure \(\lambda = \lambda^{(T, r, \rho)}\) is uniquely defined by the two properties \(\lambda([\rho, x]) = r(\rho, x)\) for all \(x \in T\), and \(\lambda(\mathbb{I}(\rho, T)) = 0\), where \(\mathbb{I}(\rho, T)\) is the set of non-isolated leaves (see [ALW17 Section 2.1]). Note that the total mass \(\lambda(T)\) (the “total length” of the metric tree) does not depend on the choice of \(\rho\).

Proposition 2.32 (total length of \((T_\nu, r_\nu)\)). Let \((T, c)\) be an order separable, order complete algebraic tree, \(\nu\) a measure on \((T, \mathcal{B}(T, c))\) with \(\nu([x, y]) < \infty\) for all \(x, y \in T\) and such that \(\nu|_{\mathbb{I}(T, c)}\) is purely atomic, and \(r_\nu\) be defined by (1.11). Then the following hold:

(i) The total length of the metric tree \((T_\nu, r_\nu)\) is given by
\[
\lambda(T_\nu) = \frac{1}{2} \int_T \deg(T, c) \, d\nu.
\] (2.39)

(ii) \(\int_T \deg(T, c) \, d\nu = \int_{T_\nu} \deg(T_\nu, r_\nu) \circ \pi_\nu \, d\nu\).

Proof. (i) Let \(D := \{v_n : n \in \mathbb{N}\}\) be a subset of \((T, c)\) which contains the atoms of \(\nu\) and satisfies (2.24), and \(\pi_\nu : T \to T_\nu\) be the canonical projection. We use \(\rho := \pi_\nu(v_1)\) as the root of \((T_\nu, r_\nu)\). Then
\[
T \setminus \mathbb{I}(T, c) \subseteq [D] = \bigcup_{n \in \mathbb{N}} [v_1, \ldots, v_n],
\] (2.40)
where \([A] := \bigcup_{x, y \in A} [x, y]\). Hence \(\nu(T \setminus [D]) = 0\), and
\[
\lambda(T_\nu, r_\nu) = \lim_{n \to \infty} \lambda(T_\nu, r_\nu)(\pi_\nu([v_1, \ldots, v_n])).
\] (2.41)

Abbreviate \(T_n := [v_1, \ldots, v_n]\) and \(\ell_n := \lambda(T_\nu, r_\nu)(\pi_\nu([v_1, \ldots, v_n]))\). If \(v_{n+1} \in T_n\), then \(T_{n+1} = T_n\) and \(\lambda(T_\nu, r_\nu)(\pi_\nu(T_{n+1})) = \lambda(T_\nu, r_\nu)(\pi_\nu(T_n))\). Otherwise, there exists a unique \(u_n \in T\) with \(T_{n+1} = T_n \cup (u_n, v_{n+1}]\), and thus
\[
\ell_{n+1} = \ell_n + r_\nu(u_n, v_{n+1}) = \ell_n + \nu((u_n, v_{n+1}]) - \frac{1}{2} \nu\{v_{n+1}\} + \frac{1}{2} \nu\{u_n\}.
\] (2.42)
For $v \in T_n$, let $\deg_n(v)$ be the degree of $v$ in the tree $(T_n, e|_{T_n})$. In the case $v_{n+1} \notin T_n$, we have $\deg_{n+1}(v) = \deg_n(v)$ for $v \in T_n \setminus \{u_n\}$, and $\deg_{n+1}(u_n) = \deg_n(u_n) + 1$. By induction over $n$, we obtain

$$\ell_n = \frac{1}{2} \int_{T_n} \deg_n \, dv$$

(2.43)

Note that $\deg_n(v)$ is monotonically increasing in $n$, and $\deg(v) = \lim_{n \to \infty} \deg_n(v)$ holds for all $v \in [D]$. Thus using the monotone convergence theorem, combining (2.41) and (2.43) yields (2.49).

(ii) If $\deg_{(T,c)}(v) \neq \deg_{(T_r,c)}(\pi(v))$, then either $\pi(S_v(y)) = \{\pi(v)\}$ for some $y \in T$ (and thus $\deg_{(T,c)}(v) > \deg_{(T_r,c)}(\pi(v)))$, or $\pi(v) = \pi(v')$ for some $v' \in Br(T,c)$ (and thus $\deg_{(T,c)}(v) < \deg_{(T_r,c)}(\pi(v))$). In both cases we have that $\nu\{v\} = \nu\{\pi(v)\} = 0$, and thus the claim follows.

Corollary 2.33 (compactness for bounded degree trees). Let $(T, c)$ be an order separable algebraic tree, and $\nu$ a finite measure on $(T, B(T, c))$ with $\nu\{v \in T : \deg(v) > d\} = 0$ for some $d \in \mathbb{N}$. Then the completion of $(T_\nu, r_\nu)$ is compact.

Proof. W.l.o.g. assume that $\nu|_{H(T,c)}$ is non-atomic (if $\nu|_{H(T,c)}$ has a non-atomic part, we can remove it without changing $r_\nu$). Then by Proposition 2.32(i), $(T_\nu, r_\nu)$ has finite total length. As complete metric trees with finite total length are necessarily compact, the statement follows.

We are now in a position to prove Theorem 1.

Proof of Theorem 1. (i) “$\Rightarrow$” Since every compact metric space is bounded, complete and separable, this step follows from Lemma 2.25.

“$\Leftarrow$” Let $(T, c)$ be an order separable algebraic continuum tree. To avoid trivialities, assume that $T$ contains more than two points. By Lemma 2.31 we can choose a non-atomic probability measure $\nu$ on $(T, B(T, c))$ satisfying (2.38). Define $r_\nu$ by (1.11). Then the equivalence classes in $T_\nu$ are singletons by (2.38), and we may identify $T_\nu$ with $T$.

By Proposition 2.30 $(T, r_\nu)$ is a complete $\mathbb{R}$-tree and the identity is a tree homomorphism by Lemma 2.28. Thus $c$ is induced by $r_\nu$. Moreover, $\nu(\text{br}(T,c)) = 0$ because $\text{br}(T,c)$ is countable and $\nu$ is non-atomic. We can therefore conclude with Corollary 2.33 that $(T, r_\nu)$ is also compact.

(ii) “$\Rightarrow$” This is obvious because every order separable algebraic continuum tree is induced by a separable $\mathbb{R}$-tree according to part (i), and subspaces of separable metric spaces are separable.

“$\Leftarrow$” Let $(T, c)$ be an order separable algebraic tree and $D \subseteq T$ a countable set satisfying (2.24). Let $\nu$ be any probability measure on $D$ with $\nu\{x\} > 0$ for all $x \in D$, and $r_\nu$ defined by (1.11). The equivalence classes in $T_\nu$ are singletons, and we may again identify $T_\nu$ with $T$. By Proposition 2.32 $(T, r_\nu)$ is a metric tree with (2.27). As $(T, c)$ is order separable, $(T, r_\nu)$ is separable by Proposition 2.30(i). Moreover, the diameter of $(T, r_\nu)$ is bounded by 1. Hence, by [Eva08, Theorem 3.38] there is a bounded, separable $\mathbb{R}$-tree $(\tilde{T}, \tilde{r})$ such that $T \subseteq \tilde{T}$ and $r_\nu$ is the restriction of $\tilde{r}$ to $T$. By Lemma 2.25 this $\mathbb{R}$-tree induces an algebraic continuum tree $(\tilde{T}, \tilde{c})$, and $T$ is a subtree of $\tilde{T}$.

“in particular”. According to part (i), there is a metric $\tilde{r}$ on $\tilde{T}$ such that $(\tilde{T}, \tilde{r})$ is a compact $\mathbb{R}$-tree inducing $(\tilde{T}, \tilde{c})$. Let $r$ be the restriction of $\tilde{r}$ to $T$. Then $(T, r)$ is a totally bounded metric tree inducing $(T, c)$.
2.5 Tree homomorphisms versus homeomorphisms

Since order separable algebraic continuum trees are $\mathbb{R}$-trees where we have “forgotten” the metric, the question arises how homeomorphisms of $\mathbb{R}$-trees relate to tree homomorphism of the corresponding algebraic trees. A first observation is that homeomorphisms are necessarily tree homomorphisms. This statement relies on connectedness of the $\mathbb{R}$-trees and we cannot replace “$\mathbb{R}$-tree” by “metric tree”: every bijection between finite metric trees is obviously a homeomorphism because the topologies are discrete, but not necessarily a tree homomorphism.

**Lemma 2.34** (homeomorphisms are tree isomorphisms). Let $(T, r)$, $(\hat{T}, \hat{r})$ be $\mathbb{R}$-trees, and $f: T \to \hat{T}$ a homeomorphism. Then $f$ is a tree homomorphism.

**Proof.** The branch point map can be expressed in terms of intervals by (1.2). In an $\mathbb{R}$-tree $(T, r)$, the interval $[x, y]$, $x, y \in T$, is the unique simple path from $x$ to $y$, which is a purely topological notion, and hence preserved by homeomorphisms. $\Box$

**Example 2.35** (tree isomorphisms need not be homeomorphisms). In Lemma 2.34, the converse is not true: bijective tree homomorphisms need not be homeomorphisms, even if the trees are order separable. To see this, let $r, \hat{r}$ the metrics on $\mathbb{N}$ defined by $r(n, m) = \frac{1}{n} + \frac{1}{m}$, $\hat{r}(n, m) = 2$ for distinct $n, m \in \mathbb{N}$. Let $T$ and $\hat{T}$ be the $\mathbb{R}$-trees generated by $(\mathbb{N}, r)$ and $(\mathbb{N}, \hat{r})$, respectively. Then both $\hat{T}$ and $T$ are the countable star with set $\mathbb{N}$ of leaves. In $T$, the distance from the branch point to leaf $n$ is $\frac{1}{n}$, while it is 1 in $\hat{T}$. Hence $T$ is compact while $\hat{T}$ is not. The identity on $\mathbb{N}$ can be extended to a bijective tree homomorphism $f: T \to \hat{T}$ which cannot be continuous. $\diamond$

Example 2.35 shows that there are non-homeomorphic (topologically non-equivalent) $\mathbb{R}$-trees inducing isomorphic (equivalent) algebraic continuum trees. This can only happen if at least one of the trees is non-compact.

**Proposition 2.36** (tree isomorphisms of compact $\mathbb{R}$-trees are homeomorphisms). Let $T, \hat{T}$ be $\mathbb{R}$-trees, and $f: T \to \hat{T}$.

(i) If $\hat{T}$ is compact, $f(T)$ is connected, and $f$ a tree homomorphism, then $f$ is continuous.

(ii) If both $T$ and $\hat{T}$ are compact and $f$ is bijective, then $f$ is a homeomorphism if and only if it is a tree homomorphism.

**Proof.** (ii) is obvious from (i) and Lemma 2.34.

Assume $f$ is a tree homomorphism, $f(T)$ is connected, and $\hat{T}$ is compact. Choose a root $\rho \in T$. Let $v_n \to v$ be a convergent sequence in $T$, and $w \in \hat{T}$ an accumulation point of $f(v_n)$. Then there is a subsequence $(n_k)_{k \in \mathbb{N}}$ with $f(v_{n_k}) \to w$. We have

$$v = \sup_{k \in \mathbb{N}, i > k} \inf_{k \in \mathbb{N}, i > k} v_{n_i} \quad \text{and} \quad w = \sup_{k \in \mathbb{N}, i > k} \inf_{k \in \mathbb{N}, i > k} f(v_{n_i}),$$

(2.44)

where sup and inf are w.r.t. the partial orders $\leq_{\rho}$ and $\leq_{f(\rho)}$ in the first and second equality, respectively. In the following, we show $w = f(v)$. Because $f$ is order preserving for these partial orders due to Lemma 2.11 we obtain $w \leq_{f(\rho)} f(v)$. Assume for a contradiction $w \neq f(v)$. Because $f(T)$ is connected, there is $y \in \hat{T}$ with $w <_{f(\rho)} y <_{f(\rho)} f(v)$ and $x \in T$
with \( y = f(x) \). For \( u := c(\rho, x, v) \), we have \( f(u) = \hat{c}(f(\rho), y, f(v)) = y \), \( u \leq_\rho v \), and \( u \neq v \). Therefore, \( u \leq_\rho v_n \), for all sufficiently large \( i \), and thus \( y = f(u) \leq f(\rho) \) for those \( i \). Now \((2.44) \ implies \ y \leq f(\rho) \) \( w \) in contradiction to the choice of \( y \), finishing the proof of \( w = f(v) \).

Compactness of \( \hat{T} \) and uniqueness of accumulation points implies \( f(v_n) \to f(v) \), and \( f \) is continuous.

In view of Theorem 1, Proposition 2.36 implies that order separable algebraic continuum trees are in one-to-one correspondence with homeomorphism classes of compact \( \mathbb{R} \)-trees. Furthermore, the unique metric topology induced by the compact \( \mathbb{R} \)-tree coincides with the topology \( \tau \) introduced in Subsection 2.3. But be aware that there may be other, non-homeomorphic, non-compact \( \mathbb{R} \)-trees inducing the same order separable algebraic continuum tree, as shown in Example 2.35.

**Corollary 2.37** (uniqueness of inducing \( \mathbb{R} \)-tree). Every order separable algebraic continuum tree is induced by a compact \( \mathbb{R} \)-tree that is unique up to homeomorphism, and the unique induced topology coincides with the component topology \( \tau \) defined in \((2.22)\).

**Proof.** That an order separable algebraic continuum tree is induced by a compact \( \mathbb{R} \)-tree is Theorem 1. Any two such compact \( \mathbb{R} \)-trees are isomorphic as algebraic trees, hence homeomorphic by Proposition 2.36. The topology \( \tau \) is a Hausdorff topology and clearly weaker than the topology induced by the \( \mathbb{R} \)-tree. Hence, by compactness of the \( \mathbb{R} \)-tree, the two topologies coincide.

### 3 The space of algebraic measure trees

In this section, we define algebraic measure trees, and equip the space of (equivalence classes) of algebraic measure trees with a notion of convergence. In what follows, the order separability of the underlying algebraic tree is crucial. Therefore, we include it already in the following definition of algebraic measure trees.

**Definition 3.1** (algebraic measure trees). An algebraic measure tree \((T, c, \mu)\) is an order separable algebraic tree \((T, c)\) together with a probability measure \(\mu\) on \(B(T, c)\).

**Definition 3.2** (equivalence of algebraic measure trees). (i) We call two algebraic measure trees \((T_i, c_i, \mu_i), i = 1, 2, \) equivalent if there exist subtrees \(A_i\) of \(T_i\) with \(\mu_i(A_i) = 1\), and a measure preserving tree isomorphism \(f\) from \(A_1\) onto \(A_2\). In this case, we call \(f\) isomorphism of the algebraic measure trees.

(ii) A metric measure tree \((T, r, \mu)\) is called a metric representation of the algebraic measure tree \((T', c', \mu')\) if its induces algebraic measure tree \((T, c(T'), \mu)\) is equivalent to \((T', c', \mu')\).

In the following, we denote for an algebraic measure tree \(x := (T, c, \mu)\) by \(\text{supp}(x)\) the algebraic subtree generated by the support of \(\mu\), i.e.

\[
\text{supp}(x) := c(\text{supp}(\mu)^3),
\]

and by

\[
\text{br}(x) := \text{br}(T, c) \cap \text{supp}(x)
\]
the set of branch points of $\chi$. It is easy to check that an isomorphism $f$ from $\chi = (T, c, \mu)$ to $\chi' = (T', c', \mu')$ induces a bijection between $\text{br}(\chi)$ and $\text{br}(\chi')$ (although it need neither be defined nor injective on all of $\text{supp}(\chi)$). Also note that $\chi$ is equivalent to $\text{supp}(\chi)$ equipped with the appropriate restrictions of $c$ and $\mu$.

**Remark 3.3** (a note on our definition of equivalence). Every algebraic measure tree is equivalent to an algebraic continuum measure tree, and has a metric representation with a compact $\mathbb{R}$-tree by Theorem 1. For the definition of equivalence of algebraic measure trees it is important that we do not require the whole trees to be isomorphic (see Example 3.11 below). On the other hand, it is also important that the isomorphism is injective on a subtree (as opposed to only a subset) of full measure, because otherwise it would not be an equivalence relation and every tree with $n$ leaves and uniform distribution on them would be equivalent to the $n$-star.

**Example 3.4** (the linear non-atomic measure tree). There is only one equivalence class of linearly ordered algebraic measure trees with non-atomic measure. Indeed, let $(T, c, \mu)$ be an algebraic measure tree with $\text{br}(T, c) = \emptyset = \text{at}(\mu)$. Then, by Theorem 1 there is a tree isomorphism from $T$ into $[0, 1]$ and we may assume $T \subseteq [0, 1]$ to begin with. Let $F_{\mu} : [0, 1] \to [0, 1]$ be the distribution function of $\mu$. Then $F_{\mu}$ is continuous and maps $\mu$ to Lebesgue-measure $\lambda_{[0,1]}$. Let $A := \{x \in \text{supp}(\mu) : \text{there is no } y_n \in [0, 1] \setminus \text{supp}(\mu) : y_n < x, y_n \to x\}$ be the support of $\mu$ with left boundary points removed. Then $F_{\mu}$ restricted to $A$ is bijective and hence a measure preserving tree isomorphism onto $[0, 1]$ (with Lebesgue measure and canonical branch point map). Thus $(T, c, \mu)$ is equivalent to $[0, 1]$.

Let

$$T := \{\text{equivalence classes of algebraic measure trees}\}. \quad (3.3)$$

Next, we equip $T$ with a topology. We shall base this notion of convergence on the fact that algebraic measure trees allow for metric representations (see Theorem 1), and require convergence in Gromov-weak topology of particular representations. To this end, let

$$\mathbb{H} := \{\text{equivalence classes of (separable) metric measure trees}\}, \quad (3.4)$$

where we consider two metric measure trees $(T, r, \mu)$ and $(T', r', \mu')$ as equivalent if there exists a measure preserving isometry between the metric completions of $\text{supp}(\mu)$ and $\text{supp}(\mu')$.

In order to get a useful topology on $T$, we cannot take arbitrary (optimal) metric representations. Instead, given an algebraic measure tree $(T, c, \mu)$, we use the metric $r_{\mu}$ defined in $(2.25)$ for the branch point distribution $\nu$, namely the distribution of the random branch point obtained by sampling three points with the sampling measure $\mu$.

**Definition 3.5** (branch point distribution). The branch point distribution of an algebraic measure tree $(T, c, \mu)$ is the push-forward of $\mu^{\otimes 3}$ under the branch point map,

$$\nu := c_{\ast} \mu^{\otimes 3}. \quad (3.5)$$

Note that the branch point distribution is not necessarily supported by $\text{br}(T, c)$. For instance, every atom of $\mu$ is also an atom of $\nu$. If $(T, c, \mu)$ and $(T', c', \mu')$ are equivalent algebraic measure trees with branch point distributions $\nu$ and $\nu'$, respectively, then the isomorphism is also an isometry w.r.t. $r_{\nu}$ and $r_{\nu'}$. Therefore, the following embedding map is well-defined
Definition 3.6 (embedding \( \iota \)). Define the map \( \iota : T \to \mathbb{H} \) by
\[
\iota(T, c, \mu) := (T_\nu, r_\nu, \mu_\nu),
\]
where \( \nu = c_\mu \otimes 3 \) is the branch point distribution of \((T, c, \mu)\), and \( \mu_\nu \) is the image of \( \mu \) under the canonical projection \( \pi_\nu \).

The topology we use on \( T \) is the Gromov-weak topology w.r.t. the branch point distribution distance. That is, it is the topology induced by \( \iota \), i.e. the weakest (coarsest) topology on \( T \) such that \( \iota \) is continuous.

Definition 3.7 (bpdd-Gromov-weak topology). Let \( \mathbb{H} \) be equipped with the Gromov-topology. We call the topology induced on \( T \) by the embedding \( \iota \) branch point distribution distance Gromov-weak topology (bpdd-Gromov-weak topology).

The following reconstruction theorem is crucial for the usefulness of bpdd-Gromov-weak convergence.

Proposition 3.8 (\( \iota \) is injective). The map \( \iota : T \to \mathbb{H} \) is injective, and \( \iota(x) \) is a metric representation of \( x \in T \).

Proof. If we show that \( \iota(x) \) is a metric representation of \( x = (T, c, \mu) \in T_2 \), it is obvious that \( \iota \) is injective, because equivalence of metric measure spaces implies equivalence of the corresponding algebraic measure trees by Lemma 2.34.

Choosing an appropriate representative, we can assume that \( \nu\{v\} > 0 \) for all \( v \in \text{br}(T, c) \). The canonical projection \( \pi_\nu : T \to T_\nu \) is a tree homomorphism by Lemma 2.28. To show equivalence of \((T, c, \mu)\) and \((T_\nu, c(T_\nu, r_\nu), \mu_\nu)\), we have to show that \( \pi_\nu \) is injective on a subtree \( A \subseteq T \) with \( \mu(A) = 1 \). Let \( N := \{ v \in T : \pi_\nu(v) \neq \{v\} \} \). Then \( \mu(\pi_\nu(v)) = 0 \) for all \( v \in N \), and \( w \in \pi_\nu(v) \) implies \( [v, w] \subseteq \pi_\nu(v) \) because \( \pi_\nu \) is a tree homomorphism. Because there are at most countably many non-degenerate, disjoint closed intervals in \( T \) due to order separability, this implies that \( \pi_\nu(N) \) is countable, and thus \( \mu(N) = 0 \). Define \( A = T \setminus N \). Then \( \mu(A) = 1 \), and \( \pi_\nu \) is injective on \( T \setminus N \). To see that \( A \) is a subtree, pick \( x, y, z \in A \). If \( v := c(x, y, z) \in \{x, y, z\} \), then \( v \in A \). Otherwise, \( v \in \text{br}(T, c) \), and hence \( \nu\{v\} > 0 \). This implies \( \nu_\nu(v) = \{v\} \), i.e. \( v \in A \). \( \square \)

Corollary 3.9 (metrizability). \( T \) equipped with bpdd-Gromov-weak topology is a separable, metrizable space.

Proof. The Gromov-weak topology on \( \mathbb{H} \) is separable, and metrizable, e.g. by the Gromov-Prohorov metric \( d_\text{GP} \) (see [GPW09]). Because \( \iota \) is injective by Proposition 3.8, \( d_\text{BP}(x, y) := d_\text{GP}(\iota(x), \iota(y)), x, y \in T \), is a metric on \( T \) inducing bpdd-Gromov-weak topology. \( \square \)

Remark 3.10 (distance polynomials). By definition, a sequence \((x_n)_{n \in \mathbb{N}}\) in \( T \) converges to \( x \in T \) bpdd-Gromov-weakly if and only if \( \iota(x_n) \xrightarrow{n \to \infty} \iota(x) \) Gromov-weakly. It has been shown that the Gromov-weak convergence is equivalent to the convergence of the distribution of the distance matrix ([GPW09, Theorem 5]). Therefore, the bpdd-Gromov-weak convergence is equivalent to
\[
\Phi(x_n) \xrightarrow{n \to \infty} \Phi(x)
\]
for all so-called polynomials $\Phi : \mathbb{T} \to \mathbb{R}$, which are test functions of the form

$$
\Phi(T, c, \mu) = \Phi^n(\phi(T, c, \mu) := \int_{T^n} \phi(\{r_{c, i} \in \mathbb{R}^3(x_i, x_j)\}_{1 \leq i, j \leq n}) \mu^\otimes n(dx),
$$

(3.8)

where $n \in \mathbb{N}$ and $\phi \in C_b(\mathbb{R}^{n \times n})$. Note that the set $\Pi_t$ of all polynomials is an algebra, and therefore also convergence determining for $T$-valued random variables (see [Loh13, BK10]).

As pointed out in Remark 3.3, the equivalence class of every algebraic measure tree contains an algebraic continuum measure tree. The following example shows that $\iota$ would not be injective if we had defined it on the set of algebraic continuum measure trees with the stricter notion of equivalence, where the whole algebraic continuum trees have to be measure preserving isomorphic.

**Example 3.11.** For $x \geq 0$, let $T_x$ be the $\mathbb{R}$-tree generated by the interval $I_x = [-x, 1]$ together with additional leaves $\{v_n\}$, $n \in \mathbb{N}$, where $c(0, 1, v_n) = \frac{1}{n}$ and $r(\frac{1}{n}, v_n) = \frac{1}{n}$, i.e. at each point $\frac{1}{n} \in I_x$ there is a branch of length $\frac{1}{n}$ attached. Then $T_x$ is a compact $\mathbb{R}$-tree for every $x \geq 0$, hence induces an algebraic continuum tree by Theorem 1. Let $\mu_x\{-x\} = \frac{1}{2}$, and $\mu_x\{v_n\} = 2^{-n-1}$ for $n \in \mathbb{N}$. Then $\chi_x := (T_x, \mu_x) \in T_2$. Now $\iota(x) = \iota(y)$ for every $x, y \geq 0$, but $T_x$ and $T_0$ are not homeomorphic, hence not isomorphic by Proposition 2.30.

Note that $A_x := \{x\} \cup \{v_n : n \in \mathbb{N}\} \cup \{\frac{1}{n} : n \in \mathbb{N}\}$ is a subtree of $T_x$ with $\mu_x(A_x) = 1$, and $A_x$ is isomorphic (although not homeomorphic) to $A_0$.

In order to construct algebraic measure trees, it is of course not necessary to specify the mass of every Borel subset. To the contrary, we can use the following Carathéodory-type extension result. To this end, recall for $x, y \in T$ with $x \neq y$ from (2.21) the component $S_x(y) = S_x^{(T, c)}(y)$ of $T \setminus \{x\}$ which contains $y$, and define

$$
S_x(x) := \{x\}.
$$

(3.9)

Then $T$ is the disjoint union of the $\deg(x) + 1$ sets in

$$
C_x := \{S_x(y) : y \in T\}.
$$

(3.10)

Note that $C_x = \{S_x(y) : y \in V\}$ for order dense $V \subseteq T$ with $x \in V$. In particular, $C_x$ is countable if $(T, c)$ is order separable. For $y \in T$, $V \subseteq T$, we call a function $f : V \to \mathbb{R}$ order left-continuous on $V$ w.r.t. $\leq_y$ if the following holds. For all $x, x_n \in V$ with $x_1 \leq_y x_2 \leq_y \cdots$ and $x = \sup_{n \in \mathbb{N}} x_n$ w.r.t. $\leq_y$ (in short $x_n \uparrow x$), we have $\lim_{n \to \infty} f(x_n) = f(x)$. Recall the notion of algebraic continuum tree from Definition 2.9.

**Proposition 3.12** (extension to a measure). Let $(T, c)$ be an order separable algebraic continuum tree, and $V \subseteq T$ order dense. Then $\mu_0 : C_V := \bigcup_{x \in V} C_x \to [0, 1]$ has a unique extension to a probability measure on $\mathcal{B}(T, c)$ if it satisfies

1. For all $x \in V$, $\sum_{A \in C_x} \mu_0(A) = 1$
2. For all $x, y \in V$ with $x \neq y$,

$$
\mu_0(S_x(y)) + \mu_0(S_y(x)) \geq 1
$$

(3.11)
3. For every \( y \in V \), the function \( \psi_y : x \mapsto \mu_0(S_x(y)) \) is order left-continuous on \( V \) w.r.t. \( \leq_y \).

Proof. Note that \( \psi_y(x) = \psi_z(x) \) for \( z \in S_x(y) \). We therefore may write \( \psi_A(x) := \psi_y(x) \) for any \( A \subseteq S_x(y) \). Define the \( \cap \)-stable set system

\[
A := \left\{ \bigcap_{k=1}^{n} A_k : n \in \mathbb{N}, A_k \in \mathcal{C}_V \right\}. \tag{3.12}
\]

By Corollary 2.24 \( A \) generates the Borel \( \sigma \)-algebra \( \mathcal{B}(T, c) \). Let \( \emptyset \neq A \in \mathcal{A} \) and \( y \in A \). Because \( (T, c) \) has no edges and is order complete, we have \( A = \bigcap_{x \in \partial A} S_x(y) \), where \( \partial \) denotes the boundary w.r.t. the component topology \( \tau \), which is a finite set in the case of \( A \). Using (3.11), we obtain for \( v \in V \), \( x_0, \ldots, x_n \in V \setminus \{v\} \) such that \( S_v(x_0), \ldots, S_v(x_n) \) are distinct, that

\[
\psi_{x_0}(v) \leq 1 - \sum_{k=1}^{n} \psi_{x_k}(v) \leq 1 - \sum_{k=1}^{n} (1 - \psi_v(x_k)). \tag{3.13}
\]

This implies for \( \emptyset \neq A \in \mathcal{A} \), by induction over \( \# \partial A \), that

\[
\mu(A) := 1 - \sum_{x \in \partial A} (1 - \psi_A(x)) \geq 0, \tag{3.14}
\]

hence \( \mu \) is a non-negative extension of \( \mu_0 \) to \( \mathcal{A} \). We claim that \( \mu \) is super-additive, additive and inner regular for compact sets. From this it follows by standard arguments that it has a unique extension to a measure on the generated \( \sigma \)-algebra \( \sigma(\mathcal{A}) = \mathcal{B}(T, c) \).

Additivity. Let \( n \in \mathbb{N} \setminus \{1\} \), and \( A_1, \ldots, A_n \in \mathcal{A} \setminus \{\emptyset\} \) disjoint with \( A := \bigcup_{k=1}^{n} A_k \in \mathcal{A} \). Define \( D := \bigcup_{k=1}^{n} \partial A_k \). Then \( \partial A \subseteq D \) and there is \( x \in D \setminus \partial A \subseteq A \). Let \( I_x := \{k \in \{1, \ldots, n\} : x \in \partial A_k\} \) and choose \( y_k \in A_k \). Then, because the \( A_k \) are disjoint, the \( S_x(y_k) \), \( k \in I \), are distinct, and because the \( A_k \) cover \( A \), we have \( \{S_x(y_k) : k \in I_x\} = C_x \). In particular, \( \sum_{k \in I_x} \psi_{y_k}(x) = 1 \), and \( B_x := \bigcup_{k \in I_x} A_k \in \mathcal{A} \) with \( \partial B_x = \bigcup_{k \in I_x} \partial A_k \setminus \{x\} \). We obtain

\[
\sum_{k \in I_x} \mu(A_k) = \sum_{k \in I_x} \left( 1 - (1 - \psi_{y_k}(x)) - \sum_{z \in \partial A_k \setminus \{x\}} (1 - \psi_{y_z}(z)) \right) \tag{3.15}
\]

\[
= \sum_{k \in I_x} \psi_{y_k}(x) - \sum_{z \in \partial B_x} (1 - \psi_{z}(x)) = \mu(B_x).
\]

By induction over \( n \), this implies additivity of \( \mu \).

Super-additivity. Let \( A_1, \ldots, A_n \in \mathcal{A} \setminus \{\emptyset\} \) be disjoint and \( \bigcup_{k=1}^{n} A_k \subseteq A \in \mathcal{A} \). The case \( n = 1 \) is trivial, and we proceed by induction over \( n \). Choose \( y \in A_1 \) and let \( D := \partial A_1 \setminus \partial A \).

For \( x \in D \), \( C \in C_x := C_x \setminus S_x(y) \) and \( k \in \{2, \ldots, n\} \), either \( A_k \subseteq C \), or \( A_k \cap C = \emptyset \). Therefore, we have the decomposition \( \{2, \ldots, n\} = \bigcup_{x \in D} \bigcup_{C \in C_x} I_C \) with \( I_C := \{k : A_k \subseteq C\} \). Because \( C \cap A \in \mathcal{A} \), and \( A_k \subseteq C \cap A \) for \( k \in I_C \), we can use the induction hypothesis to obtain

\[
\sum_{k \in I_C} \mu(A_k) \leq \mu(C \cap A) = \psi_C(x) - \sum_{z \in \partial A \cap C} (1 - \psi_A(x)). \tag{3.16}
\]
Therefore,
\[
\mu(A_1) = 1 - \sum_{x \in \partial A_1} (1 - \psi(x)) - \sum_{x \in D} (1 - \psi(x)) \\
= \mu(A) + \sum_{x \in \partial A \setminus \partial A_1} (1 - \psi(x)) - \sum_{x \in D} \sum_{C \in C'_x} \mu(A_k) \\
\leq \mu(A) - \sum_{x \in D} \sum_{C \in C'_x} \sum_{k \in I_C} \mu(A_k) \\
= \mu(A) - \sum_{k=2}^n \mu(A_k).
\] (3.17)

Compact regularity. According to Proposition 2.18, all closed subsets of $T$ are compact. Let $y \in A \in \mathcal{A}$. Because $(T, c)$ is an order separable algebraic continuum tree, and $V$ is order dense, we find for $z \in \partial A$ a sequence $(x_n(z))_{n \in \mathbb{N}}$ in $A \cap V$ with $x_n(z) \uparrow z$ w.r.t. $\leq y$ as $n \to \infty$. Define $A_n := \bigcap_{z \in \partial A} S_{x_n(z)}(y) \in \mathcal{A}$ and $K_n := A_n \cup \partial A_n$. Then $K_n$ is compact, $A_n \subseteq K_n \subseteq A$, and because $\partial A$ is finite, we have $\partial A_n = \{x_n(z) : z \in \partial A\}$ for sufficiently large $n$. Thus, by order left-continuity of $\psi_y$,
\[
\lim_{n \to \infty} \mu(A_n) = 1 - \lim_{n \to \infty} \sum_{z \in \partial A} (1 - \psi(x_n(z))) = 1 - \sum_{z \in \partial A} (1 - \psi(z)) = \mu(A),
\] (3.18)
and $\mu$ is inner compact regular as claimed.

We conclude this section with an extension result, which will be very useful for reading off algebraic measure trees from (sub-)triangulations of the circle in Section 4 below. In Proposition 3.12 we assumed the whole tree to be known, and considered the question of constructing a probability measure on it. Now, we assume that not the whole tree is given a priori, but only the (countably many) branch points. The question is, whether there is an extension of the tree which is rich enough to carry a measure with the specified masses of components.

**Proposition 3.13** (construction of algebraic measure trees). Let $(V, c_V)$ be a countable algebraic tree, and for each $x \in V$, let $A \mapsto \psi_A(x)$ be a probability measure on $C_x$. Define $\psi(x) := \psi_{S_x(y)}(x)$. Assume that for $x, y \in V$ with $x \neq y$,
\[
\psi_x(y) + \psi_y(x) \geq 1.
\] (3.19)
Then there is a unique (up to equivalence) algebraic measure tree $\chi = (T, c, \mu)$ such that
(i) $V \subseteq T$, $\operatorname{br}(T, c) = \operatorname{br}(V, c_V)$,
(ii) $\mu(S_{x}^{(T, c)}(y)) = \psi_y(x)$ for all $x, y \in V$,
(iii) $\operatorname{at}((\mu)) \subseteq V$, where $\operatorname{at}((\mu))$ denotes the set of atoms of $\mu$.

Note that in general we cannot obtain $\operatorname{If}(T, c) \subseteq \operatorname{If}(V, c_V)$. To the contrary, $\operatorname{If}(T, c)$ can be uncountable (for every representative of $\chi$).
has a unique extension to a probability measure \( \tilde{V} \), such that \( \operatorname{br}(\tilde{V}, \tilde{c}) \) is an algebraic tree with \( x \leq y z \leq x \) for all \( n \). Furthermore, we extend \( \psi \) to \( \tilde{V} \) by defining \( \tilde{\psi}(z) \) as \( \tilde{\psi}(z) = 0 \) and \( \tilde{\psi}_x(z) = 1 - \phi_y(x) \). It is easy to check that \( (\tilde{V}, \tilde{c}) \) together with \( \tilde{\psi} \) satisfies the prerequisites of the Proposition, \( \operatorname{br}(\tilde{V}, \tilde{c}) = \operatorname{br}(V, c) \), and \( \{ x \in \tilde{V} : \tilde{\psi}_x(x) > 0 \} = \{ x \in V : \psi_x(x) > 0 \} \).

Now assume w.l.o.g. that \( \psi_y \) is already order left-continuous for all \( y \in V \). Because \( \psi_y \) is countable, it is in particular order separable and according to Theorem 1, there is an order separable algebraic continuum tree \((T, c)\) such that \((T, c)\) is a subtree. We can choose \((T, c)\) such that \( \operatorname{br}(T, c) = \operatorname{br}(V, c) \). Consider the closure \( \overline{V} \) of \( V \) w.r.t. the component topology \( \tau \). For \( x \in \overline{V} \setminus V \), we define

\[
\psi_y(x) := \sup\{ \psi_y(z) : z \in V \cap S_x(y) \}.
\]

Then (3.19) holds for \( x, y \in \overline{V} \), \( x \neq y \), and \( \psi_y \) is order left-continuous. For every \( \{x, y\} \in \text{edge}(\overline{V}, \overline{c}) \), where \( \overline{c} \) is the restriction of \( c \) to \( T \), we fix an order isomorphism \( \varphi_{x, y} : [x, y] \to [0, 1] \), which exists by Cantor’s order characterization of \( \mathbb{R} \) because \( [x, y] \) is a linearly ordered, separable algebraic continuum tree. For every \( z \in T \setminus V \), there exists \( \{ x, y \} \in \text{edge}(\overline{V}, \overline{c}) \), with \( z \in [x, y] \). We define

\[
\psi_y(z) := (1 - \varphi_{x, y}(z))\psi_y(x) + \varphi_{x, y}(z)(1 - \psi_x(y)),
\]

\[
\psi_x(z) := 1 - \psi_y(z)
\]
and \( \psi_z(z) := 0 \). Now we can use Proposition 3.12 to see that

\[
\mu_0(S_x(y)) := \psi_y(x)
\]
has a unique extension to a probability measure \( \mu \) on \( B(T, c) \).

The last step in the construction is to remove point-masses outside \( V \) by expanding them to intervals. To this end, let \( P := \text{at}(\mu) \setminus V \), and \( \overline{T} := (T \setminus P) \cup (P \times [0, 1]) \). Because \( P \subseteq T \setminus V \) contains no branch points, we can extend the restriction of \( c \) to \( T \setminus P \) to a branch point map \( \overline{c} \) on \( \overline{T} \) in a canonical way such that \( [(x, 0), (x, 1)] = \{x\} \times [0, 1] \) for \( x \in P \). Define the Markov kernel \( \kappa \) from \( T \) to \( \overline{T} \) by

\[
\kappa(x) := \begin{cases} 
\delta_x, & x \in (T \setminus P), \\
\delta_x \otimes \lambda_{[0,1]}, & x \in P,
\end{cases}
\]

where \( \delta_x \) is the Dirac measure in \( x \) and \( \lambda_{[0,1]} \) is Lebesgue measure. Let \( \overline{\mu} := \kappa_*(\mu) \) be the push-forward of \( \mu \) under \( \kappa \). Then \((\overline{T}, \overline{c}, \overline{\mu})\) is a separable algebraic measure tree, and by construction \( \operatorname{br}(\overline{T}, \overline{c}) = \operatorname{br}(V, c) \) as well as \( \operatorname{at}(\overline{\mu}) = \operatorname{at}(\mu) \cap V \subseteq V \). Furthermore, for \( x, y \in V \), we have \( \overline{\mu}(S_x(\overline{T}, \overline{c})(y)) = \mu(S_x(T, c)(y)) = \psi_y(x) \) as claimed.

Uniqueness. Follows similarly, where we note that it does not matter how we distribute the mass on an edge of \((\overline{T}, \overline{c})\) in a non-atomic way, because all algebraic measure trees without branch points and non-atomic measure are equivalent by Example 3.3.

\( \square \)
4 Triangulations of the circle

In this section we want to encode binary algebraic measure trees by triangulations of subsets of the circle. This is comparable with the encoding of compact (ordered, rooted) metric (probability) measure trees by excursions over the unit interval, where the height profile encodes the branch point map as well as the metric distances. Moreover, also the measure can be encoded by the excursion by identifying the lengths of sub-excursions with the mass of the corresponding subtrees. Similarly, it turns out we can encode binary algebraic measure trees by what we call sub-triangulations of the circle, and as in the case of coding metric measure trees with excursions, the resulting coding map associating to a sub-triangulation the algebraic measure tree is continuous.

First, in Subsection 4.1, we introduce the space of sub-triangulations of the circle. In Subsection 4.2, we construct the coding map.

4.1 The space of sub-triangulations of the circle

Let \( D \) be a (fixed) closed disc of circumference 1, and \( S := \partial D \) the circle. As usual, for a subset \( A \subseteq D \), we denote by \( \bar{A}, \dot{A}, \partial A \) and conv(\( A \)) the closure, the interior, the boundary and the convex hull of \( A \), respectively. Furthermore, let
\[
\Delta(A) := \{ \text{connected components of } \text{conv}(A) \setminus A \},
\]
and
\[
\circ(A) := \{ \text{connected components of } D \setminus \text{conv}(A) \}.
\]
Then we have the disjoint decomposition
\[
D = A \cup \bigcup \Delta(A) \cup \bigcup \circ(A).
\]

Definition 4.1 ((sub-)-triangulations of the circle). A sub-triangulation of the circle is a closed, non-empty subset \( C \) of \( D \) satisfying the following two conditions:

(Tri1) \( \Delta(C) \) consists of open interiors of triangles.

(Tri2) \( C \) is the union of non-crossing (non-intersecting except at endpoints), possibly degenerate closed straight line segments with endpoints in \( S \).

We denote the set of sub-triangulations of the circle by \( \mathcal{T} \), i.e.
\[
\mathcal{T} := \{ \text{sub-triangulations of the circle} \},
\]
and call \( C \in \mathcal{T} \) triangulation of the circle if and only if \( S \subseteq C \).

In particular, (Tri1) implies that \( \partial \text{conv}(C) \subseteq C \), and we may call \( C \) triangulation of \( \partial \text{conv}(C) \). Given (Tri1), (Tri2) implies that \( \circ(A) \) consists of circular segments with the bounding straight line excluded and the rest of the bounding arc included. We want to point out that our definition of triangulation of the circle differs from the one given by Aldous in [Ald94b, Definition 1]. Namely, Aldous required only Condition (Tri1). For the characterization of triangulations of the circle as limits of triangulations of \( n \)-gons given in Proposition 4.3 below, however, Condition (Tri2) is necessary. See Figure 4 for an example of a triangulation in the sense of Aldous that is excluded by Condition (Tri2), a sub-triangulation of the circle that is no triangulation of the circle, and a triangulation of the circle.
For a metric space \((X,d)\), let
\[
\mathcal{F}(X) := \{ F \subseteq X : F \neq \emptyset, \text{ } F \text{ closed} \},
\] (4.4)
and equip \(\mathcal{F}(X)\) with the *Hausdorff metric topology*. That is, we say that a sequence \((F_n)_{n \in \mathbb{N}}\) converges to \(F\) in \(\mathcal{F}(X)\) if and only if for all \(\varepsilon > 0\) and all large enough \(n \in \mathbb{N}\),
\[
F^n \supseteq F \quad \text{and} \quad F^\varepsilon \supseteq F_n,
\] (4.5)
where for all \(A \in \mathcal{F}(X)\), as usual, \(A^\varepsilon := \{ x \in X : d(x, A) < \varepsilon \}\). It is well-known that if \((X,d)\) is compact, then \(\mathcal{F}(X)\) is a compact metrizable space as well. As sub-triangulations of the circle are elements of \(\mathcal{F}(D)\), we naturally equip \(\mathcal{T}\) with the Hausdorff metric topology. A first observation is that \(\mathcal{T}\) is actually a closed, and therefore compact subspace of \(\mathcal{F}(D)\).

**Lemma 4.2** (compactness of \(\mathcal{T}\)). *Both the space of triangulations of the circle, and the space \(\mathcal{T}\) of sub-triangulations of the circle, are compact metrizable spaces in the Hausdorff metric topology.*

*Proof.* Because \(D\) is compact, \(\mathcal{F}(D)\) is compact as well, and it is sufficient to show that \(\mathcal{T}\) and the set of triangulations of the circle are closed subsets of \(\mathcal{F}(D)\).

Let \(C_n \in \mathcal{T}\) with \(C_n \xrightarrow{n \to \infty} C \in \mathcal{F}(D)\) in the Hausdorff metric topology. (Tri1) is easily seen to be a closed property, thus \(C\) satisfies (Tri1). Let \(L_n\) be a set of non-crossing line segments with endpoints in \(S\) such that \(C_n = \bigcup L_n\). The closure of \(L_n\) in \(\mathcal{F}(D)\) has the same property (it possibly differs from \(L_n\) by a set of degenerated one-point segments contained in non-degenerate segments of \(L_n\)), so we may assume \(L_n\) is closed to begin with, so that \(L_n \in \mathcal{F}(\mathcal{F}(D))\). Because \(\mathcal{F}(\mathcal{F}(D))\) is compact, we may assume, taking a subsequence if necessary, that \(L_n \to L\) for some \(L \in \mathcal{F}(\mathcal{F}(D))\). Obviously, \(L_n\) consists of non-crossing line segments with endpoints in \(S\). Because the union operator \(\bigcup : \mathcal{F}(\mathcal{F}(D)) \to \mathcal{F}(D)\) is continuous, we have \(\bigcup L = C\). In particular, (Tri2) holds for \(C\), and \(C \in \mathcal{T}\). Obviously, also the property that \(S \subseteq C\) is preserved by Hausdorff metric limits, thus the set of triangulations of the circle is closed as well.

We now show two characterizations of sub-triangulations of the circle. Namely, condition (Tri2) can be replaced by existence of “triangles in the middle” which is the major technical ingredient for the construction of the branch point map in the next subsection. Furthermore, they are precisely the limits of finite sub-triangulations, where we consider a sub-triangulation $C$ as finite if $C \cap S$ is a finite set, or equivalently, $C$ consists of finitely many line segments.

**Proposition 4.3** (characterization of (sub-)triangulations). Let $\emptyset \neq C \subseteq \overline{D}$ be closed. Then the following are equivalent.

1. $C$ is a sub-triangulation of the circle.
2. Condition (Tri1) holds, all extreme points of $\text{conv}(C)$ are contained in $S$, and

   (Tri2)' For $x, y, z \in \Delta(C) \cup \partial(C)$ pairwise distinct, there exists a unique $c_{xyz} \in \Delta(C)$ such that $x, y, z$ are subsets of pairwise different connected components of $\overline{D} \setminus \partial c_{xyz}$.

3. There exists a sequence $(C_n)_{n \in \mathbb{N}}$ of finite sub-triangulations of the circle with $C_n \to C$ in the Hausdorff metric topology.

Furthermore, $C$ is a triangulation of the circle if and only if $C_n$ in $\Delta$ can be chosen as a triangulation of a regular $n$-gon inscribed in $S$.

**Remark 4.4** (Condition (Tri2)'). That $x, y, z$ are subsets of different connected components of $\overline{D} \setminus \partial c_{xyz}$ means that either $c_{xyz} \in \{x, y, z\}$ and the two elements of $\{x, y, z\} \setminus \{c_{xyz}\}$ are subsets of different connected components of $\overline{D} \setminus \overline{c_{xyz}}$, or $c_{xyz} \notin \{x, y, z\}$ and $x, y, z$ are subsets of pairwise different connected components of $\overline{D} \setminus \overline{c_{xyz}}$.

**Proof of Proposition 4.3**. $\Rightarrow \Rightarrow$. Because $C$ is the union of line segments with endpoints on $S$, it is obvious that the extreme points of $\text{conv}(C)$ are contained in $S$. We have to show (Tri2)', so let $x, y, z \in \Delta(C) \cup \partial(C)$ be pairwise distinct and note that uniqueness is obvious. If one of the elements of $\{x, y, z\}$, say $x$, is such that the other two are subsets of two different connected components of $\overline{D} \setminus \overline{x}$, then necessarily $x \in \Delta(C)$, and $c_{xyz} := x$ has the desired properties. So assume this is not the case.

Fix a set $L$ of non-crossing, closed lines with endpoints in $S$ such that $C = \bigcup L$. Define

$$L_x := \{ \ell \in L : \ell \text{ separates } x \text{ from } y \cup z \text{ in } \overline{D}, \}$$

(4.6)

note that $L_x \neq \emptyset$ because $y$ and $z$ are in the same connected component of $\overline{D} \setminus \overline{x}$ by assumption, and order $L_x$ by distance from $x$. Similarly, define $L_y$ as set of lines separating $y$ from $x \cup z$ ordered by distance from $y$, and $L_z$ as set of lines separating $z$ from $x \cup y$, ordered by distance from $z$. Define $\ell_x := \sup L_x$, $\ell_y := \sup L_y$, and $\ell_z := \sup L_z$, which exist because $C$ is closed. In particular, they are non-crossing, and because $\text{conv}(C) \setminus C$ may only consist of triangles, they have to be the sides of some $c_{xyz} \in \Delta(C)$ which has the desired properties.

$\Rightarrow \Rightarrow$. Because the extreme points of $\text{conv}(C)$ are on the circle, for every $x \in \partial(C)$, the boundary $\partial_0 x$ in $\overline{D}$ is a single straight line with endpoints in $S$. Let $(V_n)_{n \in \mathbb{N}}$ be an increasing sequence of finite subsets of $\Delta(C) \cup \partial(C)$ such that $c_{xyz} \in V_n$ for pairwise distinct $x, y, z \in V_n$, and $V_n \uparrow \Delta(C) \cup \partial(C)$. Let $A_n := \overline{D} \setminus \bigcup V_n$. Then $A_n \to C$ in the Hausdorff metric topology. Because $c_{xyz} \in V_n$ for distinct $x, y, z \in V_n$, the boundary of each of the finitely many connected components of $A_n \setminus S$ consists of one or two line segments and one or two connected sub-arcs of $S$. Therefore, there is a finite sub-triangulation $C_n \subseteq A_n$ of the circle with Hausdorff distance from $A_n$ less than $e^{-n}$. Thus $C_n \to C$. 

31. Obvious, because $T$ is a closed subset of $F(D)$ by Lemma 4.2.

"Furthermore". If $C_n$ is a triangulation of the $n$-gon, it contains the $n$-gon, and hence any Hausdorff metric limit as $n \to \infty$ contains the circle, and hence is a triangulation of the circle. That triangulations of the circle can be approximated by triangulations of regular $n$-gons is a slight modification of the arguments above. Details are left to the reader.

The most prominent random tree is Aldous’s Brownian CRT, which is the limit of uniform random trees. Similarly, one can define the Brownian triangulation of the circle.

Example 4.5 (Brownian triangulation). The uniform random triangulation of the $n$-gon converges in law with respect to the Hausdorff metric topology to the so-called Brownian triangulation $C_{CRT}$, see [Ald94a, Ald94b, CK14]. A realisation is shown in the right of Figure 4. It has a.s. Hausdorff dimension $\frac{3}{2}$ (see [Ald94a]).

4.2 Coding binary measure trees with (sub-)triangulations of the circle

Given an algebraic tree $(T,c)$, recall the set of leaves $\text{lf}(T,c)$, and the degree $\text{deg}_{(T,c)}(v)$ of $v \in T$ from (2.36) and (2.35), respectively. In this section, we are interested in the following subspace of the space of all binary algebraic measure trees.

Definition 4.6 (our space $T_2$). Let $T_2 \subseteq T$ be the set of (equivalence classes of) algebraic measure trees $(T,c,\mu)$ with $(T,c)$ binary (i.e. $\text{deg}_{(T,c)}(v) \leq 3$ for all $v \in T$) and $\text{at}(\mu) \subseteq \text{lf}(T,c)$.

The space $T_2$ is of particular interest to us, as it is invariant under the dynamics of the Aldous diffusion on cladograms, the construction of which was one of the motivations for studying algebraic measure trees, and because, as we will see, it is precisely the space of algebraic measure trees that can be coded by sub-triangulations of the circle.

To illustrate the construction of the tree coded by a sub-triangulation, we first consider a triangulation $C$ of the regular $n$-gon into necessarily $n - 2$ triangles (see Figure 3). Here, every triangle corresponds to a branch point of the tree, and two branch points are connected by an edge if and only if the triangles share a common edge. We then add a leaf for every edge of the $n$-gon and obtain a graph-theoretic binary tree with $n$ leaves and $n - 2$ internal vertices. Recall from Example 2.4 that the finite graph-theoretic tree corresponds to a unique algebraic tree. We finally assign to each leaf mass $n - 1$ (which corresponds to the length of the arcs of the circle connecting two endpoints of edges of the $n$-gon if we inscribe it in a circle of unit length), and obtain an algebraic measure tree.

The main result of this section is that there is a natural, surjective coding map from $T$ onto $T_2$, which is also continuous. To state that formally, we need further notation. Given a sub-triangulation $C \subseteq D$, recall $\Delta(C)$ and $\cap(C)$ from (4.2) and (4.1), respectively. For $x \in \Delta(C) \cup \cap(C)$, and $y \subseteq D$ connected and disjoint from $\partial_D x$, where $\partial_D$ denotes the boundary in the space $D$, let

$$\text{comp}_x(y) := \text{the connected component of } D \setminus \partial_D x \text{ which contains } y.$$  (4.7)

For $x \in \Delta(C)$, let $p_i(x), i = 1, 2, 3$, be the mid-points of the three arcs of $S \setminus \partial x$, and define

$$\Box(C) := \{ \{p_i(x)\} : x \in \Delta(C), i \in \{1, 2, 3\}, \text{comp}_x(\{p_i(x)\}) \subseteq C\};$$  (4.8)

as well as $\text{comp}_p(p) := \{p\}$ for $p \in \Box(C)$ (see Figure 5). Recall the definition of components $S_v(w)$ in an algebraic tree from (2.24).
Proposition 3.13 yields existence and uniqueness of the desired algebraic tree.

arcs add up to the total length of $\lambda$ of Proposition 3.13. Indeed, $V_T$ that 


corresponds to the set $\Delta(C)$, and $\deg(x) = 1$ for all $x \in \partial(C)$.

Figure 5: Triangulation $C$ with $\#\Delta(C) = 1$, $\partial(C) = \emptyset$, and $\#\Box(C) = 3$. The coded tree consists of three line segments with non-atomic measure of $\frac{1}{3}$ each, glued together at one branch point.

Lemma 4.7 (induced branch point map). For $C \in T$, let $V_C := \Delta(C) \cup \partial(C) \cup \Box(C)$. If $V_C \neq \emptyset$, then there is a unique branch point map $c_V : V_C^3 \to V_C$ such that $(V_C, c_V)$ is an algebraic tree with $S_{(V_C, c_V)}(y) = \{v \in V_C : \operatorname{comp}_C(y) = \operatorname{comp}_V(v)\}$ for $x, y \in V_C$. Furthermore, $\deg(x) = 3$ for all $x \in \Delta(C)$, and $\deg(x) = 1$ for $x \in \partial(C) \cup \Box(C)$.

Proof. Recall from Proposition 3.13 that for a sub-triangulation $C$ of the circle and pairwise distinct $x, y, z \in \Delta(C) \cup \partial(C)$, there is a triangle $c_{xyz} \in \Delta(C)$ “in the middle”. It is straightforward to see that this defines a branch point map and can naturally be extended to $V_C^3$.

The following theorem states that all sub-triangulations $C$ of the circle can be associated with an element in $T_2$ for which $\Delta(C)$ corresponds to the set of branch points, $\partial(C)$ corresponds to the set

$\operatorname{lf}_\mu(x) := \{x \in \operatorname{lf}(T, c) : \mu(\{x\}) > 0\}$

of leaves which carry an atom, and $\operatorname{comp}_C(w)$ corresponds to the component $S_C(w)$.

Theorem 2 (algebraic measure tree associated to a sub-triangulation). (i) For every sub-triangulation $C \subseteq \mathbb{D}$ of the circle, there is a unique (up to equivalence) algebraic measure tree $\chi_C = (T_C, c_C, \mu_C) \in T_2$ with the following properties:

(CM1) $V_C \subseteq T_C$, $\operatorname{br}(\chi_C) = \Delta(C)$, and $c_C$ is an extension of $c_V$, where $(V_C, c_V)$ is defined in Lemma 3.7.

(CM2) $\mu_C(S_{(T_C, c_C)}(y)) = \lambda_S(S \cap \operatorname{comp}_x(y))$ for all $x, y \in V_C$, where $\lambda_S$ denotes the Lebesgue measure on $S$.

(CM3) $\operatorname{at}(\mu_C) = \partial(C)$.

(ii) The coding map $\tau : T \to T_2$, $C \mapsto \chi_C$ is surjective and continuous, where $T$ is equipped with the Hausdorff metric topology and $T_2$ with the bpd-Gromov-weak topology.

Proof. (i) Let $C$ be a sub-triangulation of the circle. If $C = \mathbb{D}$, then $\Delta(C) = \partial(C) = \emptyset$, which requires by (CM1) that $\operatorname{br}(\chi_C) = \emptyset$, and by (CM3) that $\operatorname{at}(\mu) = \emptyset$. There is a unique algebraic measure tree without branch points and atoms, namely the line segment with no atoms (see Example 4.10). We may therefore assume w.l.o.g. that $C \neq \mathbb{D}$, and consequently that $T_C \neq \emptyset$.

We claim that $(V_C, c_V)$ together with $\psi_y(x) := \lambda_S(S \cap \operatorname{comp}_x(y))$ satisfies the assumptions of Proposition 3.13. Indeed, $V_C$ is obviously countable and an algebraic tree by Lemma 4.7; $\psi_y(x)$ depends on $y$ only through its equivalence class w.r.t. $\sim_x$, and the lengths of all the arcs add up to the total length of $\lambda_S(S) = 1$. Furthermore, $\psi_x(y) + \psi_y(x) \geq \lambda_S(S) = 1$, and Proposition 3.13 yields existence and uniqueness of the desired algebraic measure tree.
(ii) Let $\chi = (T, c, \mu) \in \mathbb{T}_2$. We construct a sub-triangulation $C$ such that $\tau(C) = \chi$. Fix $\rho \in \text{If}(T, c)$, and recall that $\rho$ induces a partial order relation $\leq_\rho$. We can extend this partial order to a total (planar) order $\leq$ by picking for every $v \in \text{br}(T, c)$ an order of the two components of $T \setminus \{v\}$ that do not contain $\rho$. That is, we define $S_0(v) := S_v(\rho)$, denote the two remaining components of $T \setminus \{v\}$ by $S_1(v), S_2(v)$, and define

$$v \leq w \iff v \leq_\rho w \text{ or } v \in S_1(c(x, y, \rho)), w \in S_2(c(x, y, \rho)).$$

(4.10)

Identify $\mathbb{S}$ with $[0, 1]$, where the endpoints are glued. For $a \in [0, 1]$ and $b, c > 0$ with $a + b + c \leq 1$, let $\Delta(a, b, c) \subseteq \mathbb{D}$ be the open triangle with vertices $a, a + b, a + b + c \in \mathbb{S}$, $\ell(a, b) \subseteq \mathbb{D}$ the straight line from $a$ to $a + b$, and $L(a, b)$ the connected component of $\mathbb{D} \setminus \ell(a, b)$ containing $a + \frac{b}{2} \in \mathbb{S}$. The first vertex of the triangle or circular segment corresponding to $v \in \text{br}(T, c) \cup \text{H}_{\text{atom}}(\chi)$ is given by the total mass before (w.r.t. $\leq$ defined in (4.10)), i.e. by

$$\alpha(v) := \mu(\{u \in T : u < v\}).$$

Define

$$\mathbb{D} \setminus C := \bigcup_{v \in \text{br}(T, c)} \Delta(\alpha(v), \mu(S_1(v)), \mu(S_2(v))) \sqcup \bigcup_{v \in \text{H}_{\text{atom}}(\chi)} L(\alpha(v), \mu(\{v\}))$$

(4.12)

By definition of $C$, $\text{conv}(C) \setminus C$ consists of open triangles, i.e. condition (Tri1) is satisfied. Furthermore, the extreme points of $\text{conv}(C)$ are contained in $\mathbb{S}$, and for $x, y, z \in \Delta(C) \cup \alpha(C)$ distinct, there are corresponding $u, v, w \in T$, and a triangle $\Delta_{xyz} \in \Delta(C)$ corresponding to $c(u, v, w)$, which satisfies the requirements of (Tri2)'.

We defer the proof of continuity of $\tau$ to the next section, where we prove it in Lemma 5.19.

The following is obvious now.

**Lemma 4.8** (no atoms). A sub-triangulation $C$ of the circle is a triangulation of the circle if and only if, for $(T_C, c_C, \mu_C) := \tau(C)$, the measure $\mu_C$ is non-atomic.

**Corollary 4.9** (finite tree approximation). Let $\chi = (T, c, \mu) \in \mathbb{T}_2$. Then there is a sequence $(\chi_n)_{n \in \mathbb{N}}$ of finite algebraic measure trees in $\mathbb{T}_2$ with $\chi_n \to \chi$ bpd-d-Gromov-weakly. Furthermore, if $\mu$ is non-atomic, then $\chi_n$ can be chosen as a tree with $n$ leaves and uniform distribution on the leaves.

**Proof.** By Theorem 2 there is a sub-triangulation $C \in \mathcal{T}$ with $\tau(C) = \chi$, and by Proposition 4.3 there are finite sub-triangulations $C_n$ with $C_n \to C$. Obviously, $\chi_n := \tau(C_n)$ is a finite algebraic measure tree and by continuity of $\tau$ we have $\chi_n \to \chi$. If $\mu$ is non-atomic, then, by Lemma 4.8, $C$ is a triangulation of the circle, and hence, by Proposition 4.3, $C_n$ can be chosen as triangulation of the $n$-gon, which means that $\chi_n$ has $n$ leaves and uniform distribution on them.

We conclude this section with a few illustrative examples.

**Example 4.10** (coding algebraic measure trees without branch points). Let $\chi$ be an algebraic measure tree without branch points. If $\chi = x_C$ for some sub-triangulation $C$, then $\Delta(C) = \text{br}(x_C) = \emptyset$ and the following five cases can occur (see Figure 9): a) $x_C$ consists of one single
point of mass 1. Then \( C = \{x\} \) for some \( x \in S \). b) \( x_C \) consists of an interval with two leaves, where each carries positive mass adding up to 1, in which case \( C \) is a single line segment dividing the circle into two arcs with length corresponding to the masses of the two leaves. c) \( x_C \) consists of an interval with two leaves, where each has positive mass adding up to \( a < 1 \). In this case, \( C \) is the area of the disc bounded by two distinct line segments and two arcs (possibly one of them degenerated) of \( S \), and the lengths of the remaining two arcs are given by the masses of the leaves. d) \( x_C \) consists of an interval with two leaves, where one has positive mass \( a < 1 \) and the other one has zero mass. Then \( C \) is a circular segment with arc length \( 1 - a \). e) \( x_C \) consists of an interval with no atoms on the leaves, which implies \( C = D \). 

**Example 4.11** (a complete binary tree). Let \( C \) be the sub-triangulation of the circle drawn in the middle of Figure 4. Then \( \#C(C) = \#I_{\text{atom}}(\tau(C)) = 1 \). We refer to this only leaf with positive mass as the root \( \rho \), and obtain \( \mu(\{\rho\}) = \frac{1}{3} \), corresponding to the length of the dotted arc. Moreover, \( \tau(C) \) consists of a complete rooted binary tree in the sense of graph theory (with the convention that the root has degree one), together with an uncountable set of leaves given by the ends at infinity and carrying the remaining \( \frac{2}{3} \) of the mass.

**Example 4.12** (coding the Brownian CRT). Recall the Brownian triangulation \( C_{\text{CRT}} \) from Example 4.5, which is defined as the limit in distribution of uniform random triangulations \( C_n \) of the \( n \)-gon. A realization is shown in the right of Figure 4. It is easy to see that \( \tau(C_n) \) is the uniform binary tree with \( n \) leaves and uniform distribution on the leaves. Thus, by Theorem 2 the uniform binary tree converges bpdd-Gromov-weakly to \( \tau(C_{\text{CRT}}) \). At this point it is not entirely clear that \( \tau(C_{\text{CRT}}) \) is the algebraic measure tree induced by the metric measure Brownian CRT. We will see in Section 6 that this is indeed the case.

5 The subspace of binary algebraic measure trees

In this section we introduce in Subsections 5.1 and 5.2 with the sample shape convergence and the sample subtree-mass convergence two more notions of convergence of algebraic measure trees which seem more natural when thinking of algebraic trees as combinatorial objects. We then show in Subsection 5.3 that on \( T_2 \), both of these notions are equivalent to the bpdd-Gromov-weak convergence. The main tools are a uniform Glivenko Cantelli argument, and that the coding map sending a sub-triangulation of the circle to an element in \( T_2 \) is continuous.
5 THE SUBSPACE OF BINARY ALGEBRAIC MEASURE TREES

5.1 Convergence in distribution of sampled tree shapes

The basic idea behind Gromov-weak convergence for metric measure spaces is to sample finite metric sub-spaces with the sampling measure \( \mu \) and then require these to converge in distribution. In this section, we propose a corresponding construction for binary algebraic measure trees, where we sample finite tree shapes with \( \mu \).

First, we have to make precise what we mean by “tree shape”, which we understand to be a cladogram with the peculiarity that leaves may carry more than one label. The multi-label case is necessary to allow for sampling the same point several times due to a possible atom at that point.

**Definition 5.1** \((m\)-labelled cladogram\). For \( m \in \mathbb{N} \), an \( m\)-labelled cladogram is a binary, finite algebraic tree \( C = (C, c) \) together with a surjective labelling map \( \ell: \{1, \ldots, m\} \to \ell f(C) \). Two \( m\)-labelled cladograms \((C, \ell_1)\) and \((C, \ell_2)\) are equivalent if they are label preserving isomorphic i.e., there exists a tree isomorphism \( f: C_1 \to C_2 \) with \( f(\ell_1(i)) = \ell_2(i) \) for all \( i = 1, \ldots, m \).

Define

\[
\mathcal{C}_m := \{\text{isomorphism classes of } m\text{-labelled cladograms}\}.
\]

In the following we will use cladograms to encode the shape of a subtree spanned by a finite sample of leaves.

**Definition 5.2** (tree shape). For a binary algebraic tree \((T, c), m \in \mathbb{N}\), and \( u_1, \ldots, u_m \in T \setminus \text{br}(T, c) \), the tree shape \( s_T(u_1, \ldots, u_m) \) of the \( m\)-labelled cladogram spanned by \( u_1, \ldots, u_m \) in \((T, c)\) is the unique (up to isomorphism) \( m\)-labelled cladogram \( s_T(u_1, \ldots, u_m) = (C, c_C, \ell) \) with \( \ell f(C) = \{u_1, \ldots, u_m\} \) and \( \ell(i) = u_i \) for all \( i = 1, \ldots, m \), and such that the identity on \( \ell f(C) \) extends to a tree homomorphism from \( C \) onto \( c(\{u_1, \ldots, u_m\})^3 \).

**Remark 5.3** (spanned subtree and cladogram are not necessarily isomorphic). The tree homomorphism from \( s_T(u_1, \ldots, u_m) \) onto \( c(\{u_1, \ldots, u_m\})^3 \) does not need to be injective. This is the case if (and only if) \( u_i \in (u_j, u_k) \) for some \( i, j, k \in \{1, \ldots, m\} \). See Figure[7]

**Example 5.4** (shape of a totally ordered algebraic tree). Let \((T, c)\) be a totally ordered algebraic tree, and \( u_1, \ldots, u_m \in T \). Then \( s_T(u_1, \ldots, u_m) \) is a so-called **comb tree** which has a totally ordered spine of binary branch points with attached leaves (see Figure[5]).

In the following, we build a topology on the convergence of tree shapes of \( m \) randomly sampled points. We therefore need the measurability of the shape map.
5 THE SUBSPACE OF BINARY ALGEBRAIC MEASURE TREES

Figure 8: The left shows a totally ordered binary algebraic tree and four different points \( u_1, \ldots, u_4 \). The middle shows the shape \( s_T(u_1, \ldots, u_4) \) of the cladogram which forms a comb tree. The right illustrates what happens if a fifth point is equal to \( u_1 \). Now one of the leaves of \( s_T(u_1, \ldots, u_5) \) has two labels.

Lemma 5.5 (measurability of the shape map). For every binary algebraic tree \((T, c)\) and \( m \in \mathbb{N} \), the tree shape map \( s_T : (T \setminus \text{br}(T, c))^m \to \mathcal{C}_m \) is a measurable function.

Proof. Restricted to the open subset \( \{ v \in (T \setminus \text{br}(T, c))^m : v_1, \ldots, v_m \text{ distinct} \} \), \( s_T \) is locally constant, hence continuous. The same is true on the set \( \{ v \in (T \setminus \text{br}(T, c))^m : v_1 = v_2, v_2, \ldots, v_m \text{ distinct} \} \), which is an intersection of a closed and an open set, hence measurable. We can continue this way to see that \( s_T \) is measurable on \((T \setminus \text{br}(T, c))^m\).

Definition 5.6 (tree shape distribution). For \( \chi = (T, c, \mu) \in \mathbb{T}_2 \) and \( m \in \mathbb{N} \), the \( m \)-tree shape distribution of \( \chi \) is defined by

\[
\mathcal{G}_m(\chi) := \mu^{\otimes m} \circ s_T^{-1} \in \mathcal{M}_1(\mathcal{C}_m). \tag{5.2}
\]

Example 5.7 (shape of the linear non-atomic measure tree). Let \( \chi = (T, c, \mu) \) be the linear non-atomic algebraic measure tree (Example 3.4). Then any sample \( (u_1, \ldots, u_m) \) with \( \mu \) consists of pairwise different points, and \( \mathcal{G}_m(\chi) \) is the mixture of Dirac measures on labelled comb trees where the mixture is over all (up to isometry) permutations of the labels. ◊

We refer to the weakest topology on \( \mathbb{T}_2 \) such that for every \( m \in \mathbb{N} \) the \( m \)-tree shape distribution is continuous as sample shape topology.

Definition 5.8 (sample shape topology). The topology induced on \( \mathbb{T}_2 \) by the set \( \{ \mathcal{G}_m : m \in \mathbb{N} \} \) of tree shape distributions is called sample shape topology.

We say that a sequence \( (x_n)_{n \in \mathbb{N}} \) is sample shape convergent to \( \chi \) in \( \mathbb{T}_2 \) if it converges w.r.t. the sample shape topology, i.e. if \( \mathcal{G}_m(x_n) \) converges to \( \mathcal{G}_m(\chi) \) as \( n \to \infty \) for every \( m \in \mathbb{N} \).

In analogy to the set \( \Pi_\Psi \) of polynomials introduced in Remark 3.10, we also introduce a set of test functions which evaluate the tree shape distributions. We refer to \( \Phi = \Phi^{m, \varphi} : \mathbb{T}_2 \to \mathbb{R} \),

\[
\Phi(\chi) = \int_{\mathcal{C}_m} \varphi \, \mathcal{G}_m(\chi) = \int_{T_m} \varphi \circ s_T \, d\mu^{\otimes m}, \tag{5.3}
\]

where \( m \in \mathbb{N} \) and \( \varphi : \mathcal{C}_m \to \mathbb{R} \), as shape polynomial. We also define

\[
\Pi_\Phi := \{ \text{shape polynomials on } \mathbb{T}_2 \}. \tag{5.4}
\]

Obviously, the sample shape topology is induced by the set \( \Pi_\Phi \) of shape polynomials.

Proposition 5.9 (sample shape implies bpdd-Gromov-weak convergence). On \( \mathbb{T}_2 \), the sample shape topology is stronger than the bpdd-Gromov-weak topology (i.e. any open set in the bpdd-Gromov-weak topology is open in the sample shape topology).
5 THE SUBSPACE OF BINARY ALGEBRAIC MEASURE TREES

**Proof.** The bpdd-Gromov-weak topology is induced by the set \( \Pi \), of polynomials (see Remark 3.10). Because the set of \( \phi \in C_b(\mathbb{R}^{m \times m}) \) which are Lipschitz continuous is convergence determining for probability measures on \( \mathbb{R}^{m \times m} \), the subset of those \( \Psi \in \Pi \) with

\[
\Psi(T, c, \mu) = \int_{T^{3n}} \phi\left(\left(\nu[u_i, u_j] - \frac{1}{2} \nu\{u_i\} - \frac{1}{2} \nu\{u_j\}\right)_{i,j=1,...,m}\right) \mu^{\otimes m}(d\mu) \tag{5.5}
\]

for some \( m \in \mathbb{N} \) and Lipschitz continuous \( \phi \in C_b(\mathbb{R}^{m \times m}) \) also induces the bpdd-Gromov-weak topology. Therefore, it is enough to show that such a \( \Psi \) is continuous on \( \mathbb{T}_2 \) w.r.t. the sample shape topology. We do so by showing that the restriction to \( \mathbb{T}_2 \) of \( \Psi \) is in the uniform topology. Let \( L \) be the Lipschitz constant of \( \phi \) w.r.t. the \( \ell_\infty \)-norm on \( \mathbb{R}^{m \times m} \).

For \( n \in \mathbb{N} \) with \( 3n \geq m \), we define

\[
\Phi_n(T, c, \mu) := \int_{T^{3n}} \phi\left(\left(\nu_{n,T}[u_i, u_j] - \frac{1}{2} \nu_{n,T}\{u_i\} - \frac{1}{2} \nu_{n,T}\{u_j\}\right)_{i,j=1,...,m}\right) \mu^{\otimes 3n}(d\mu), \tag{5.6}
\]

with the empirical branch point distribution

\[
\nu_{n,T} := \frac{1}{n} \sum_{k=0}^{n-1} \delta_{c(u_{3k+1}, u_{3k+2}, u_{3k+3})}. \tag{5.7}
\]

Note that the restriction of \( \Phi_n \) to \( \mathbb{T}_2 \) belongs to \( \Pi_s \) because whether or not \( c(u_{k+1}, u_{k+2}, u_{k+3}) \) lies on \( [u_i, u_j], k \in \{0, ..., n-1\}, i, j \in \{1, ..., m\} \) only depends on the shape \( s_{3n}(u) \).

Finally, we observe

\[
\|\Psi - \Phi_n\|_\infty \leq \sup_{(T, c, \mu) \in \mathbb{T}_2} \int_{T^{3n}} L \cdot 3 \sup_{I \in \mathcal{I}_T} |\nu(I) - \nu_{n,T}(I)| \mu^{\otimes 3n}(d\mu) \leq 3L \cdot \epsilon_n \to 0, \tag{5.8}
\]

with \( \mathcal{I}_T := \{[x, y]; x, y \in T\} \) and \( (\epsilon_n)_{n \in \mathbb{N}} \to 0 \), where we have used a uniform Glivenko-Cantelli estimate which upper bounds the distance of the empirical branch point distribution to the branch point distribution. Such an estimate should be known, but as we could not come up with a reference, we show it in Lemma A.4 in the appendix. We note that \( \dim_{VC}(\mathcal{I}_T) = 2 \) (compare Example A.2).

**Corollary 5.10 (metrizability).** The sample shape topology is metrizable.

**Proof.** Because the sample shape topology is induced by a countable family of functions \((\mathcal{S}_m)_{m \in \mathbb{N}}\) with values in metrizable spaces, it is pseudo-metrizable. By Proposition 5.9, it is stronger than the bpdd-Gromov weak topology, hence a Hausdorff topology. Therefore, it is metrizable.

### 5.2 Convergence in distribution of sampled subtree masses

In this subsection, we introduce yet another notion of convergence of algebraic measure trees which, in contrast to sampling tree shapes, is based on sampling branch points and evaluating the masses of the subtrees that are joined at these branch points. This approach might be more similar to the case of metric measure spaces and distance matrix distributions, because we sample a tensor of real numbers (masses of subtrees) as opposed to a combinatorial object (tree shape). Thus, the typical tools of analysis are more readily applicable for the corresponding class of test functions.
Let \((T, c, \mu) \in \mathbb{T}_2\) and recall from (2.21) for \(u, v, w \in T\) the subtree components \(S_{c(u,v,w)}(x)\) of \(T \{ c(u, v, w) \}\) which contain \(x \neq c(u, v, w)\). Here, we always take the component containing \(x = u\) and use the notation
\[
S(u, v, w) := S_{c(u,v,w)}(u) \cup \{c(u,v,w)\}.
\] (5.9)

**Lemma 5.11** (measurability of the subtree masses). For every binary algebraic measure tree \(\chi = (T, c, \mu) \in \mathbb{T}_2\) and \(m \in \mathbb{N}\), the function \(m_\chi: T^m \to [0,1]^{m \times m}\), given for \(u = (u_1, \ldots, u_m) \in T^m\) by
\[
m_\chi(u) := (\mu(\tilde{S}(u_i, u_j, u_k)))_{i,j,k=1,\ldots,m}
\] (5.10)
is measurable.

**Proof.** First, we claim that the map \(\psi: T^2 \to [0,1]\),
\[
\psi(u,v) := 1_{v \neq u} \cdot \mu(S_u(u))
\] (5.11)
is lower semi-continuous. Indeed, let \((u_n, v_n)\) be a sequence converging to \((u, v)\). We may assume w.l.o.g. that \(v \neq u\), \(u_n \in S_v(u)\), and either \(v_n \notin S_v(u)\) for all \(n \in \mathbb{N}\), or \(v_n \in S_v(u)\) for all \(n \in \mathbb{N}\). In the first case, \(S_v(u) \subseteq S_v(u_n)\), and hence \(\psi(u, v) \leq \psi(u_n, v_n)\). In the second case, for every \(x \in S_v(u)\) and \(n \geq n_x\) sufficiently large, we have \(u \in S_v(u_n)\) and \(v_n \notin [x, u]\). This means \(x \in S_v(u) = S_v(u_n)\) and hence
\[
\psi(u,v) - \liminf_{n \to \infty} \psi(u_n, v_n) \leq \lim_{n \to \infty} \mu(S_u(u) \setminus S_v(u_n)) = 0.
\] (5.12)

Therefore, \(\psi\) is lower semi-continuous.

Because \(\mu\) is a finite measure, the map \(g: T \to [0,1], v \mapsto \mu(\{v\})\) is upper semi-continuous, and the branch point map \(c\) is continuous due to Lemma 2.16. Therefore,
\[
u \mapsto \mu(\tilde{S}(u_i, u_j, u_k)) = \psi(u_i, c(u_i, u_j, u_k)) + g(c(u_i, u_j, u_k))
\] (5.13)
is measurable, and the same applies to \(m_\chi\).

**Definition 5.12** (subtree-mass tensor distribution). For \(\chi = (T, c, \mu) \in \mathbb{T}_2\) and \(m \in \mathbb{N}\), the 
\(m\)-subtree-mass tensor distribution of \(\chi\) is defined by
\[
\vartheta_m(\chi) := \mu^\otimes m \circ m_\chi^{-1} \in \mathcal{M}_1([0,1]^{m \times m \times m}).
\] (5.14)

**Example 5.13** (symmetric binary tree). Let for each \(n \in \mathbb{N}\), \(x_n = (T_n, c_n, \mu_n)\) the symmetric binary tree with \(N = 2^n\) leaves and the uniform distribution on the set of leaves. Then the 
\(3\)-subtree-mass tensor distribution of \(x_n\) is equal to
\[
\vartheta_3(x_n) = \mu_n^\otimes 3 \circ m_\chi^{-1}
\] (5.15)

\[
= \sum_{k=1}^{n-1} \frac{1}{2^{k(n+1)}} (1 - \frac{1}{2^k}) \left( \delta \left(\frac{1}{2^{k+1}}, \frac{1}{2^{k+1}}, (1 - \frac{1}{2^k}) \right) + \delta \left(\frac{1}{2^{k+1}}, (1 - \frac{1}{2^k}), \frac{1}{2^{k+1}} \right) + \delta \left((1 - \frac{1}{2^k}), \frac{1}{2^{k+1}}, \frac{1}{2^{k+1}} \right) \right)
\] (5.15)

\[
+ \frac{1}{2^n} (1 - \frac{1}{2^n}) \left( \delta \left(\frac{1}{2^n}, \frac{1}{2^n}, 1 \right) + \delta \left(\frac{1}{2^n}, 1, \frac{1}{2^n} \right) + \delta \left(1, \frac{1}{2^n}, \frac{1}{2^n} \right) + 2^{-3n} \delta \left(\frac{1}{2^n}, \frac{1}{2^n}, \frac{1}{2^n} \right) \right)
\] (5.15)

\[
\overset{n \to \infty}{\longrightarrow} \sum_{k=1}^{\infty} \frac{1}{2^{k(n+1)}} (1 - \frac{1}{2^k}) \left( \delta \left(\frac{1}{2^{k+1}}, \frac{1}{2^{k+1}}, (1 - \frac{1}{2^k}) \right) + \delta \left(\frac{1}{2^{k+1}}, (1 - \frac{1}{2^k}), \frac{1}{2^{k+1}} \right) + \delta \left((1 - \frac{1}{2^k}), \frac{1}{2^{k+1}}, \frac{1}{2^{k+1}} \right) \right).
\] (5.15)
Remark 5.14 (3-subtree-mass tensor distribution is not enough). It is not enough to consider only the 3-subtree-mass tensor distribution. Indeed, $\vartheta_3$ cannot distinguish all non-isomorphic binary algebraic measure trees, i.e. it does not separate the points of $T_2$. To see this, take the tree from Figure 9 with uniform distribution on its 12 leaves, and the same tree with the subtrees marked by $\times$ and $\circ$, respectively, exchanged. These two trees are clearly non-isomorphic, and because the two marked subtrees have the same number of leaves, every vertex in one tree corresponds to a vertex in the other with the same value for $m_x$.  

We consider the weakest topology on $T_2$ such that for every $m \in \mathbb{N}$ the $m$-subtree-mass tensor distribution is continuous. Here, as usual, we equip $\mathcal{M}_1([0,1]^{m\times m\times m})$ with the weak topology.

Definition 5.15 (sample subtree-mass topology). The topology induced on $T_2$ by the set $\{\vartheta_m : m \in \mathbb{N}\}$ of subtree-mass tensor distributions is called sample subtree-mass topology.

We say that a sequence $(x_n)_{n \in \mathbb{N}}$ is sample subtree-mass convergent to $x$ in $T_2$ if it converges w.r.t. the sample subtree-mass topology, i.e. if $\vartheta_m(x_n)$ converges to $\vartheta_m(x)$ as $n \to \infty$ for every $m \in \mathbb{N}$.

To see that the sample subtree-mass topology is a Hausdorff topology on $T_2$, we need the following reconstruction theorem.

Proposition 5.16 (reconstruction theorem). The set of subtree-mass tensor distributions $\{\vartheta_m : m \in \mathbb{N}\}$ separates points of $T_2$, i.e., if $x_1, x_2 \in T_2$ are such that $\vartheta_m(x_1) = \vartheta_m(x_2)$ for all $m \in \mathbb{N}$, then $x_1 = x_2$.

Proof. We always assume that the representative $(T,c,\mu)$ of an algebraic measure tree is chosen such that $\mu(S_w(u)) > 0$ whenever $u, v \in T$, $u \neq v$.

Because the set $\{\mathcal{S}_m : m \in \mathbb{N}\}$ of tree shape distributions separates points by Corollary 5.10 it is enough to show that $\mathcal{S}_m$ is determined by the $m$-subtree-mass tensor distribution $\vartheta_m$ for every $m \in \mathbb{N}$. We do so by showing that there exists a (non-continuous) function $h: [0,1]^{m\times m\times m} \to \mathcal{C}_m$ such that for every $\chi = (T,c,\mu) \in T_2$ we have $s_T = h \circ m_{\chi}$ on $(T \setminus \text{br}(T,c))^m$.

This is enough, because $\mu(\text{br}(T,c)) = 0$ by countability of $\text{br}(T,c)$ and the assumption that $\text{at}(\mu) \subseteq \text{lf}(T,c)$. Fix $u_1, \ldots, u_m \in T \setminus \text{br}(T,c)$ and set $C = (C,c_C,\ell) := s_T(u_1, \ldots, u_m)$. The $m$-labelled cladogram $C$ is uniquely determined by the set of pairs $(x_1, x_2)$ with $x_i = (x_{i,1}, x_{i,2}, x_{i,3}) \in \{u_1, \ldots, u_m\}^3$, $x_{i,j} \neq x_{i,k}$ for $j \neq k, i = 1, 2$, such that

$$c_C(x_{1,1}, x_{1,2}, x_{1,3}) = c_C(x_{2,1}, x_{2,2}, x_{2,3}).$$

(5.16)

We claim that (5.16) holds if and only if we can reorder the three entries of $x_2$ such that we can replace every entry of $x_1$ by the corresponding entry of $x_2$ and obtain the same masses.
of subtrees. More precisely,
\[ m_\chi(x_{i,1}, x_{j,2}, x_{k,3}) = m_\chi(x_{i,1}, x_{1,2}, x_{1,3}) \quad \forall i, j, k \in \{1, 2\}. \]
Indeed, if \( C_c(x_1) = C_c(x_2) \) then \( c(x_1) = c(x_2) \) by definition of \( \mathfrak{s}_T \). Because none of the \( u_i \) is a branch point, every component of \( T \setminus \{c(x_1)\} \) contains precisely one of the \( x_{1,i} \), as well as one of the \( x_{2,i} \), and we can reorder the entries of \( x_2 \) such that \( x_{1,i} \) is in the same component as \( x_{2,i}, i = 1, \ldots, 3 \). Then it is easy to check that \((5.17)\) holds.

Conversely, assume that \( C_c(x_1) \neq C_c(x_2) \). Because the restriction of the tree homomorphism \( C \to c(\{u_1, \ldots, u_m\}) \) to the branch points of \( C \) is injective, this implies \( v_1 := c(x_1) \neq c(x_2) =: v_2 \). Assume w.l.o.g. that \( v_2 \) is in the same component of \( T \setminus \{v_1\} \) as \( x_{1,3} \). Let \( v_3 := c(x_{1,2}, x_{2,2}, x_{1,3}) \). We distinguish the two cases \( v_3 = v_2 \) and \( v_3 \neq v_2 \) (see Figure 10). If \( v_3 = v_2 \), then the fact that \( (T, c) \) is binary implies \( w := c(x_{1,1}, x_{2,2}, x_{1,3}) \in (v_1, v_2) \). Hence \( m_\chi(x_{1,3}, x_{1,1}, x_{2,2}) < m_\chi(x_{1,3}, x_{1,1}, x_{1,2}) \). If \( v_3 \neq v_2 \), then \( m_\chi(x_{2,1}, x_{2,2}, x_{1,3}) \neq m_\chi(x_2) \). In both cases, \((5.17)\) does not hold.

**Corollary 5.17** (metrizability). The sample subtree-mass topology is metrizable.

**Proof.** Because the sample subtree-mass topology is induced by a countable family of functions \((\vartheta_m)_{m \in \mathbb{N}}\) with values in metrizable spaces, it is pseudo-metrizable. By Proposition 5.16 it is a Hausdorff topology, hence it is metrizable.

In analogy to the sets \( \Pi_\chi \) and \( \Pi_\Psi \) of polynomials and shape polynomials, respectively, the sample subtree-mass topology also comes with a canonical set of test functions. We call \( \Psi : \mathbb{T}_2 \to \mathbb{R} \) a **subtree-mass polynomial** if there is \( m \in \mathbb{N} \) and \( \psi \in C_b([0,1]^{m\times m}) \) with
\[
\Psi(\chi) = \int_{[0,1]^{m\times m}} \psi \, d\vartheta_m(\chi) = \int_{T^m} \psi \circ m_\chi \, d\mu^{\otimes m} \quad (5.18)
\]
We also define
\[ \Pi_m := \{ \text{subtree-mass polynomials on } \mathbb{T}_2 \}. \quad (5.19) \]
Obviously, the sample subtree-mass topology is induced by the set \( \Pi_m \) of subtree-mass polynomials.

**Proposition 5.18** (sample shape convergence implies sample subtree-mass convergence). The sample shape topology is stronger than the sample subtree-mass topology.

**Proof.** The proof is similar to that of Proposition 5.9. We will show that each subtree-mass polynomial in \( \Psi \in \Pi_m \),
\[
\Psi(T, c, \mu) = \int_{T^m} \psi((\mu(S(u_i, u_j, u_k)))_{i,j,k=1,\ldots,m})^{\otimes m}(\mathfrak{d}u), \quad (5.20)
\]
with \( m \in \mathbb{N} \) and \( \psi \in C([0,1]^{m \times m \times m}) \) Lipschitz continuous w.r.t. the \( \ell_\infty \)-Norm on \([0,1]^{m \times m \times m}\) is in the uniform closure of \( \Pi_s \). Let \( L \) be the Lipschitz constant of \( \Psi \). For \( n \in \mathbb{N} \) with \( n \geq m \), we define

\[
\Phi_n(T, c, \mu) := \int_{T^n} \psi((\mu_{n,y}(\tilde{S}(u_i, u_j, u_k)))_{i,j,k=1,...,m}) \mu^n(dy)
\]

with the empirical sample distribution

\[
\mu_{n,y} := \frac{1}{n} \sum_{\ell=1}^{n} \delta_{u_\ell}.
\]

Note that \( \Phi_n \in \Pi_s \) as whether or not \( u_\ell \in \tilde{S}(u_i, u_j, u_k) \) for some \( \ell = 1,...,n \), \( i,j,k \in \{1,...,m\} \) depends only on the shape \( s_T(y) \).

Finally, applying the uniform Glivenko-Cantelli estimate Lemma [A.4], we have

\[
\|\Psi - \Phi_n\|_\infty \leq \sup_{(T,c,\mu) \in T_2} \int_{T^n} L \cdot \sup_{S \in \tilde{S}_T} \|\mu(S) - \mu_{n,y}(S)\|_\infty \mu^n(dy) \leq L \epsilon_n \quad \text{as } n \to \infty,
\]

where \( \tilde{S}_T := \{S \cup \{v\} : u, v \in T\} \) and \( (\epsilon_n)_{n \in \mathbb{N}} \to 0 \). We note that \( \dim_{VC}(\tilde{S}_T) \leq 3 \) (compare Example [A.3]).

5.3 Equivalence and compactness of topologies

In this section, we show that sample shape convergence (Definition 5.8), sample subtree-mass convergence (Definition 5.15) and branch point distribution distance Gromov-weak convergence (Definition 3.7) on \( T_2 \) are equivalent. While spaces of metric measure spaces are usually far from being locally compact, \( T_2 \) is in this topology even a compact metrizable space.

Theorem 3 (equivalence of topologies and compactness). The sample shape topology, the sample subtree-mass topology, and the bpdd-Gromov weak convergence coincide on \( T_2 \). Furthermore, \( T_2 \) is compact and metrizable in this topology.

As a preparation of the proof we show that binary algebraic measure trees depend continuously on their encoding as sub-triangulations of the circle. Recall therefore the space \( \mathcal{T} \) of sub-triangulations of the circle equipped with the Hausdorff metric topology from (4.3), and the coding map \( \tau : \mathcal{T} \to T_2 \) from Theorem 2 (note that we have not proven continuity yet).

Lemma 5.19 (continuity of the coding map). Let \( T_2 \) be equipped with the sample shape topology, and \( \mathcal{T} \) with the Hausdorff metric topology. Then the coding map \( \tau : \mathcal{T} \to T_2 \) is continuous.

**Proof.** Fix \( C \in \mathcal{T} \) and \( m \in \mathbb{N} \). By definition of the sample shape topology, it is enough to show that \( \mathcal{C}_m \circ \tau : \mathcal{T} \to \mathcal{M}_1(\mathbb{C}_m) \) is continuous at \( C \). Let \( U_1,...,U_m \) be i.i.d. points on the circle \( \mathbb{S} \) chosen with the Lebesgue measure.

Recall from (4.2) the set \( \mathcal{C}(C) \) of connected components of \( \mathbb{D} \setminus \text{conv}(C) \), from (4.7) the connected component \( \text{comp}_x(y) \) of \( \mathbb{D} \setminus \partial_\mathbb{D} x \) which contains \( y \), where \( x \in \mathcal{C}(C) \cup \mathcal{C}(C) \), and \( y \subseteq \mathbb{D} \) connected and disjoint from \( \partial_\mathbb{D} x \). Furthermore, recall the set \( \square(C) \) from (4.8), and the subtree components \( S_x(y) \) from (4.9).

For \( \epsilon > 0 \), there exists \( N = N_{C,m,\epsilon} \in \mathbb{N} \) and \( v_1,...,v_N \in \mathcal{C}(C) \cup \mathcal{C}(C) \) distinct such that with probability at least \( 1 - \epsilon \) the following holds:
• if $\{U_1, ..., U_m\} \cap v \neq \emptyset$ for $v \in \Delta(C)$, then $v \in \{v_1, ..., v_N\}$, and
• if $\{U_1, ..., U_m\} \cap \text{comp}_v(w) \neq \emptyset$ for some $v \in \Delta(C)$ and all $w \in \Delta(C) \cup \circ(C) \cup \Box(C)$ with $w \neq v$, then $v \in \{v_1, ..., v_N\}$.

Put $c' := c \cdot (12mN)^{-1}$. Then
\[
P\left(\{d(U_i, \partial v_j) \geq c', \forall i = 1, ..., m; j = 1, ..., N\}\right) \geq 1 - \epsilon. \tag{5.24}
\]

There is a $\delta = \delta(\epsilon) > 0$ sufficiently small such that for any $C' \in \mathcal{T}$ with Hausdorff metric $d_H(C, C') < \delta$ there are distinct $v'_1, ..., v'_N \in \Delta(C') \cup \circ(C')$ such that $d_H(v_i, v'_i) \leq c'$ for $i = 1, ..., N$. Let $\chi = (T, c, \mu) := \tau(C)$, and $V_1, ..., V_m$ be i.i.d. $\mu$-distributed, coupled to $U_1, ..., U_m$ such that $V_k \in S_v(w)$ if and only if $U_k \in \text{comp}_v(w)$, which is possible due to the properties of $\tau$ established in Theorem 2. Define $\chi'$ and $V'_1, ..., V'_m$ similarly with $C'$ instead of $C$. Then
\[
P\left(\{\mathcal{E}_m(V_1, ..., V_m) = \mathcal{E}_m(V'_1, ..., V'_m)\}\right) \geq 1 - 2\epsilon, \tag{5.25}
\]
which implies that $d_{\mathcal{E}}(\mathcal{E}_m(\tau(C)), \mathcal{E}_m(\tau(C'))) \leq 2\epsilon$ (with $d_{\mathcal{E}}$ denoting the Prokhorov distance). This shows that $\mathcal{E}_m \circ \tau$ is continuous at $C$ and, since $m$ and $C$ are arbitrary, that $\tau$ is continuous.

Now we are in a position to combine our results to a proof of the main theorem of Section 5.

Proof of Theorem 3. The space $\mathcal{T}$ of sub-triangulations of the circle with Hausdorff metric topology is compact according to Lemma 1.2. The coding map $\tau: \mathcal{T} \to \mathcal{T}_2$ is surjective by Theorem 2 and continuous when $\mathcal{T}_2$ is equipped with the sample shape topology by Lemma 5.19. Therefore, the sample shape topology is a compact topology on $\mathcal{T}_2$. Moreover, the sample shape topology is Hausdorff by Corollary 5.10. As the sample subtree-mass topology is a weaker Hausdorff topology by Proposition 5.18 and Corollary 5.17 it coincides with the sample shape topology. The same is true for the bpdd-Gromov-weak topology by Proposition 5.9.

Recall from Remark 3.10 that the set of distance polynomials is convergence determining for measures on $\mathcal{T}_2$. It directly follows from the construction that the same is true for the sets of shape polynomials and subtree-mass polynomials. This property is very useful for proving convergence in law of random variables.

Corollary 5.20 (Convergence determining classes of functions). The sets $\Pi_s \subseteq \mathcal{C}_b(\mathcal{T}_2)$ (defined in (5.3)) and $\Pi_m$ (defined in (5.18)) are convergence determining for measures on $\mathcal{T}_2$ with bpdd-Gromov-weak topology.

Proof. $\mathcal{T}_2$ is a compact metrizable space, and both $\Pi_s$ and $\Pi_m$ induce the bpdd-Gromov-weak topology on $\mathcal{T}_2$ by Theorem 3. Furthermore, each of $\Pi_s$ and $\Pi_m$ is closed under multiplication. Thus the claim follows by the Stone-Weierstrass theorem.

6 Examples

Consider a family $(T_n, c_n)_{n \in \mathbb{N}}$ of random, finite binary (algebraic) trees, where $(T_n, c_n)$ has $n$ leaves. Let $K_n$ be the Markov kernel that takes such a tree and removes a leaf uniformly chosen at random, together with the branch point it is attached to, thus obtaining a binary tree with $n - 1$ leaves. We say that the family is sampling consistent if $K_n(T_n, \cdot) = \mathcal{L}(T_{n-1})$, where $\mathcal{L}$ denotes the law of a random variable.
Example 6.1 (β-splitting trees). For every \( \beta \in [-2, \infty] \), let \( T_n^\beta \) be the \( \beta \)-splitting tree on \( n \) leaves from \cite{Ald96b} (with forgotten labels). Then \( (T_n^\beta)_{n \in \mathbb{N}} \) is sampling consistent. Note the special cases \( \beta = -2 \) which is the comb tree, \( \beta = -\frac{3}{2} \) which is the uniform cladogram, \( \beta = 0 \) which is the Yule tree and \( \beta = \infty \) which is the symmetric binary tree. See Figure 11 for triangulations of a realization of \( \beta \)-splitting trees for different values of \( \beta \) and large \( n \). The Aldous Brownian CRT, which is the limit for \( \beta = -\frac{3}{2} \), is shown in Figure 4.

Lemma 6.2 (convergence of sampling consistent families). Let \( ((T_n, c_n, \mu_n))_{n \in \mathbb{N}} \) be a sampling consistent family of random binary trees, and \( \mu_n \) the uniform distribution on \( \mathcal{L}(T_n, c_n, \mu_n) \). Then we have the convergence in law

\[
(T_n, c_n, \mu_n) \xrightarrow{n \to \infty} (T, c, \mu) \quad \text{on } \mathbb{T}_2 \text{ with bpdd-Gromov-weak topology}
\]

(6.1)

for some random algebraic measure tree \( (T, c, \mu) \in \mathbb{T}_2 \) with non-atomic measure \( \mu \).

Proof. Recall the m-tree shape distribution \( \mathcal{S}_m \) from Definition 5.8. Let \( n, m \in \mathbb{N} \) with \( m < n \) and define

\[
\epsilon_{n,m} := \mu_n \{ x \in T^m : x_1, \ldots, x_m \text{ not distinct} \} \leq \frac{m^2}{n}.
\]

(6.2)

Because \( (T_n) \) is sampling consistent, we obtain for the annealed shape distribution

\[
\mathbb{E}(\mathcal{S}_m(T_n, c_n, \mu_n)) = (1 - \epsilon_{n,m}) \mathcal{L}(T_n^m) + \epsilon_{n,m} \mu_{n,m}.
\]

(6.3)

where \( T_n^m \) is obtained from \( T_n \) by randomly labelling the leaves, and \( \mu_{n,m} \in \mathcal{M}_1(\mathcal{C}_m) \) is some law of \( m \)-labelled cladograms supported by cladograms where at least one leaf has more than one label. This shows that, for every fixed \( m \), the expected \( m \)-tree shape distribution converges as \( n \to \infty \). Because the \( m \)-tree shape distribution is convergence determining for the bpdd-Gromov-weak topology by Corollary 5.20, all limit points of \( \mathcal{L}(T_n, c_n, \mu_n) \) in \( \mathcal{M}_1(\mathbb{T}_2) \) coincide. According to Theorem 3, \( \mathbb{T}_2 \), and hence \( \mathcal{M}_1(\mathbb{T}_2) \), is compact and thus a unique limit exists. That the limiting measure is non-atomic is obvious, because the probability that a sampled shape is single-labelled tends to one by (6.3).

Example 6.3 (β-splitting trees continued). By Lemma 6.2 for every \( \beta \in [-2, \infty] \), the sequence \( (T_n^\beta, c_n^\beta, \mu_n^\beta)_{n \in \mathbb{N}} \) of increasing \( \beta \)-splitting trees converges in distribution to some limiting random algebraic measure tree \( (T^\beta, c^\beta, \mu^\beta) \). In the case of the uniform cladogram \( \beta = -\frac{3}{2} \), the limit is the Brownian algebraic continuum random tree which can be obtained as tree...
\(\tau(C_{\text{CRT}})\) coded by the Brownian triangulation (see Example 4.5), or as the algebraic measure tree induced by the metric measure Brownian CRT which is known to have uniform shape distribution \((\text{Ald93})\). In the case of the comb tree \((\beta = -2)\), the limit is the unit interval with Lebesgue measure (a coding triangulation is shown in the very right of Figure 6).  

\[\diamondsuit\]

A Appendix

In Subsections 5.1 and 5.2 we made use of uniform estimates of the speed of convergence in the approximation of the branch point distribution and the measure of an algebraic measure tree by empirical distribution. Such uniform Glivenko-Cantelli estimates under a bound on the Vapnik-Chervonenkis dimension (VC-dimension) of the type presented below should be well-known. As we did not find it explicitly in sufficient generality in the literature, we will present it here.

We recall the definition of VC-dimension, going back to the seminal work of Vapnik and Chervonenkis, \([\text{VC71}]\). Let \(E\) be a non-empty set and \(I\) a non-empty collection of subsets of \(E\). For \(n \in \mathbb{N}\) and \(x \in E^n\), put

\[I(x) := \{(1_{I(x_1)}, \ldots, 1_{I(x_n)}) : I \in I\} \subseteq \{0, 1\}^n.\]  

\[\text{(A.1)}\]

Then obviously, \(1 \leq \#I(x) \leq 2^n\).

**Definition A.1** (Vapnik-Chervonenkis dimension). The Vapnik-Chervonenkis dimension of \(I\) is defined as

\[\dim_{\text{VC}}(I) := \sup \{ n \in \mathbb{N} : \max_{x \in E^n} \#I(x) = 2^n \}.\]  

\[\text{(A.2)}\]

**Example A.2** (collection of intervals of an algebraic tree). Let \((T, c)\) be a separable algebraic tree with \(#T > 2\), and

\[I := I_T := \{ [u, v] : u, v \in T \}.\]  

\[\text{(A.3)}\]

For \(x_1, x_2, u \in T\) distinct, we have \(\#I(x) \geq \#([u, u], [x_1, x_1], [x_2, x_2], [x_1, x_2]) = 2^2\), hence \(\dim_{\text{VC}}(I_T) \geq 2\). Conversely, for \(x \in T^3\), either there is \(u, v \in T\) with \(x_1, x_2, x_3 \in [u, v]\).

Then w.l.o.g. \(x_2 \in [x_1, x_3]\) and \((1, 0, 1) \notin I_T(x)\). Or there is no such \(u, v \in T\), which means \((1, 1, 1) \notin I_T(x)\). Therefore,

\[\dim_{\text{VC}}(I_T) = 2.\]  

\[\text{(A.4)}\]

\[\diamondsuit\]

Recall the notion \(S_x(y)\) of the equivalence class of \(T \setminus \{x\}\) containing \(y\).

**Example A.3** (collection of subtrees branching of a branch point). Let \((T, c)\) be a separable algebraic tree, and

\[I := S_T := \{ S_v(u) \cup \{v\} : u, v \in T \}.\]  

\[\text{(A.5)}\]

We claim that

\[\dim_{\text{VC}}(S_T) \leq 3.\]  

\[\text{(A.6)}\]

For this upper bound, let \(x = (x_1, x_2, x_3, x_4) \in T^4\). By the 4-point condition of the branch point map, we can assume w.l.o.g. that

\[c(x_1, x_2, x_3) = c(x_1, x_2, x_4).\]  

\[\text{(A.7)}\]
In this case, it is not possible to cover \{x_1, x_3\} but neither \(x_2\) nor \(x_4\) with a single subtree in \(\mathcal{S}_T\), which proves the claim.

The constant in front in the following Glivenko-Cantelli lemma is clearly not optimal. For us it is only important that it is universal and not depending on the measure space \((E, \mu)\).

**Lemma A.4** (rate of convergence in Glivenko-Cantelli). Let \(E\) be a Polish space, \(\mu\) a probability measure on \(E\), \((X_n)_{n \in \mathbb{N}}\) i.i.d. \(\mu\)-distributed, and \(\mu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_k}\) the empirical measure. Then, for every \(I \subseteq \mathcal{B}(E)\) with \(\dim_{\text{VC}}(I) < \infty\) and \(n > 1\),

\[
\mathbb{E}(\sup_{I \in \mathcal{I}} |\mu(I) - \mu_n(I)|) \leq 96 \sqrt{\frac{\dim_{\text{VC}}(I)}{n}}. \tag{A.8}
\]

**Proof.** By the Kuratowski isomorphism theorem, all uncountable Polish spaces are Borel-isomorphic. Therefore, we may assume w.l.o.g. that \(E = \mathbb{R}\). Theorem 3.2 in [DL01] yields

\[
\Delta := \mathbb{E}(\sup_{I \in \mathcal{I}} |\mu(I) - \mu_n(I)|) \leq \frac{24}{\sqrt{n}} \sup_{x \in \mathbb{R}^n} \int_0^1 \sqrt{\log(2 N(r, I(x)))} \, dr, \tag{A.9}
\]

where \(N(r, I(x))\) is the covering number of \(I(x)\) w.r.t. the metric \(\frac{1}{\sqrt{n}} \cdot d_{\ell^2}\), where \(d_{\ell^2}\) is the Euclidean metric on \(\{0, 1\}^n\). This covering number can be upper-bounded in terms of the separation number \(M(r, I)\) w.r.t. the metric \(\frac{1}{n} \cdot d_{\ell^1}\) used by Haussler [Hau95], and Theorem 1 there yields

\[
N(r, I(x)) \leq M(r^2, I(x)) \leq e(\dim_{\text{VC}}(I) + 1) \left(\frac{2e}{r^2}\right)^{\dim_{\text{VC}}(I)}, \tag{A.10}
\]

provided that \(nr^2 \in \mathbb{N}\). For \(r^2 \leq \frac{1}{n}\), we use the trivial estimate \(M(r^2, I(x)) \leq 2^n\). For general \(r^2 \geq \frac{1}{n}\), we estimate \(M(r^2, I(x)) \leq M(\frac{1}{n} |nr^2|, I(x))\), and inserting (A.10) into (A.9) yields

\[
\Delta \leq \frac{24}{\sqrt{n}} \left(\frac{2e}{r^2}\right)^{\dim_{\text{VC}}(I)} \left(\int_0^{\frac{1}{\sqrt{n}}} \sqrt{\log(2e(\dim_{\text{VC}}(I) + 1)) + \dim_{\text{VC}}(I) \log(2e(r^2 - \frac{1}{n})^{-1})} \, dr \right)
\]

\[
\leq \frac{24}{\sqrt{n}} \sqrt{\dim_{\text{VC}}(I)} \left(\int_0^{\frac{1}{\sqrt{n}}} \sqrt{3 + \log(2e) - 2 \log(r)} \, dr \right), \tag{A.11}
\]

where we used that \(\log(2e(d + 1)) \leq 3d\) for \(d \geq 1\), and \(r^2 - \frac{1}{n} \geq (r - \frac{1}{\sqrt{n}})^2\). The last bracket is less than 4 for \(n > 1\), and the claim follows. \(\Box\)

**References**

[AD12] Romain Abraham and Jean-François Delmas, *A continuum-tree-valued Markov process*, Ann. Probab. **40** (2012), no. 3, 1167–1211.

[ADH13] Romain Abraham, Jean-François Delmas, and Patrick Hoscheit, *A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces*, Electron. J. Probab. **18** (2013), no. 14, 1–21.

[ADV10] Romain Abraham, Jean-François Delmas, and Guillaume Voisin, *Pruning a Lévy continuum random tree*, Electron. J. Probab. **15** (2010), no. 46, 1429–1473.

[Ald93] David Aldous, *The continuum random tree III*, Ann. Probab. **21** (1993), 248–289.
Aldous, David, *Probability distributions on cladograms*, Random discrete structures (Minneapolis, MN, 1993), IMA Vol. Math. Appl., vol. 76, Springer, New York, 1996, pp. 1–18. MR 1395604

Aldous, David, *Recursive self-similarity for random trees, random triangulations and Brownian excursion*, Ann. Probab. 22 (1994), no. 2, 527–545. MR 1288122

Aldous, David, *Triangulating the circle, at random*, Amer. Math. Monthly 101 (1994), no. 3, 223–233. MR 1264002

Aldous, David, *Probability distributions on cladograms*, Random discrete structures (Minneapolis, MN, 1993), IMA Vol. Math. Appl., vol. 76, Springer, New York, 1996, pp. 1–18. MR 1395604

Athreya, Siva, Wolfgang Löhr, and Anita Winter, *Invariance principle for variable speed random walks on trees*, Ann. Probab. 45 (2017), no. 2, 625–667, arXiv:1404.6290. MR 3630284

Blount, Douglas and Michael A. Kouritzin, *On convergence determining and separating classes of functions*, Stochastic Process. Appl. 120 (2010), no. 10, 1898–1907.

Broutin, Nicolas and Henning Sulzbach, *The dual tree of a recursive triangulation of the disk*, Ann. Probab. 43 (2015), no. 2, 738–781. MR 3306004

Chiswell, Ian, *Introduction to Λ-trees*, World Scientific Publishing Co. Inc., River Edge, NJ, 2001.

Curien, Nicolas, Bénédicte Haas, and Igor Kortchemski, *The CRT is the scaling limit of random dissections*, Random Structures Algorithms 47 (2015), no. 2, 304–327. MR 3382675

Curien, Nicolas and Igor Kortchemski, *Random non-crossing plane configurations: a conditioned Galton-Watson tree approach*, Random Structures Algorithms 45 (2014), no. 2, 236–260. MR 3245291

Curien, Nicolas, Bénédicte Haas, and Igor Kortchemski, *Percolation on random triangulations and stable looptrees*, Probab. Theory Related Fields 163 (2015), no. 1-2, 303–337. MR 3405619

Curien, Nicolas and Jean-François Le Gall, *Random recursive triangulations of the disk via fragmentation theory*, Ann. Probab. 39 (2011), no. 6, 2224–2270. MR 2932668

Dasgupta, Abhijit, *Set theory*, Birkhäuser/Springer, New York, 2014, With an introduction to real point sets. MR 3157351

Depperschmidt, Andréj, Andreas Greven, and Peter Pfaffelhuber, *Tree-valued Fleming-Viot dynamics with selection*, Annals of Appl. Probability 22 (2012), no. 6, 2560–2615.

Devroye, Luc and Gábor Lugosi, *Combinatorial methods in density estimation*, Springer Series in Statistics, Springer-Verlag, New York, 2001. MR 1843146

Dress, Andreas, Vincent Moulton, and Werner Terhalle, *T-theory: an overview*, European J. Combin. 17 (1996), no. 2-3, 161–175, Discrete metric spaces (Bielefeld, 1994). MR 97e:05069

Evans, Steven N., Rudolf Grübel, and Anton Wakolbinger, *Doob-Martin boundary of Rényi’s tree growth chain*, Ann. Probab. 45 (2017), no. 1, 225–277. MR 3601650

Evans, Steven N., Jim Pitman, and Anita Winter, *Rayleigh processes, real trees, and root growth with re-grafting*, Probab. Theo. Rel. Fields 134 (2006), no. 1, 81–126.

Evans, Steven N., *Probability and real trees*, Lecture Notes in Mathematics, vol. 1920, Springer, Berlin, 2008, Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR 2351587

Evans, Steven N. and Anita Winter, *Subtree prune and re-graft: A reversible real-tree valued Markov chain*, Ann. Probab. 34 (2006), no. 3, 918–961.
[For] Noah Forman, *Mass-structure of weighted real trees*, arXiv:1801.02700.

[Fuk87] Kenji Fukaya, *Collapsing of Riemannian manifolds and eigenvalues of Laplace operators*, Invent. Math. 87 (1987), 517–547.

[Glo12] Patric Karl Gloede, *Dynamics of Genealogical Trees for Autocatalytic Branching Processes*, Ph.D. thesis, FAU Erlangen, 2012, [http://www.min.math.fau.de/fileadmin/users/gloede/Dissertation/PatricKarlGloedeDissertation.pdf](http://www.min.math.fau.de/fileadmin/users/gloede/Dissertation/PatricKarlGloedeDissertation.pdf).

[GPW09] Andreas Greven, Peter Pfaffelhuber, and Anita Winter, *Convergence in distribution of random metric measure spaces (A-coalescent measure trees)*, Probab. Theo. Rel. Fields 145 (2009), 285–322. MR 2520129

[GPW13] ______, *Tree-valued resampling dynamics: Martingale Problems and applications*, Probab. Theory Rel. Fields 155 (2013), no. 3–4, 789–838.

[Gro99] Misha Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Progress in Mathematics, vol. 152, Birkhäuser Boston Inc., Boston, MA, 1999.

[GSW16] Andreas Greven, Rongfeng Sun, and Anita Winter, *Continuum space limit of the genealogies of interacting Fleming-Viot processes on Z*, Electron. J. Probab. 21 (2016), Paper No. 58, 64. MR 3563886

[Hau95] David Haussler, *Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik-Chervonenkis dimension*, J. Combin. Theory Ser. A 69 (1995), no. 2, 217–232. MR 1313896

[HWa] Hui He and Matthias Winkel, *Gromov-Hausdorff-Prokhorov convergence of vertex cut-trees of n-leaf Galton-Watson trees*, arxiv:1704.00044.

[HWb] ______, *Invariance principles for pruning processes of GW-trees*, arxiv:1409.1014.

[KW] Sandra Kliem and Anita Winter, *Evolving phylogenies of trait-dependent branching with mutation and competition; Part I: Existence*, arxiv:1705.03277.

[Löh13] Wolfgang Löh, *Equivalence of Gromov-Prohorov- and Gromov’s _λ_-metric on the space of metric measure spaces*, Electron. Commun. Probab. 18 (2013), no. 17, 1–10.

[LVW15] Wolfgang Löh, Guillaume Voisin, and Anita Winter, *Convergence of bi-measure R-trees and the pruning process*, Ann. Inst. H. Poincaré Probab. Statist. 51 (2015), no. 4, 1342–1368, arXiv:1304.6035.

[Pio10] Sven Piotrowiak, *Dynamics of Genealogical Trees for Type- and State-dependent Resampling Models*, Ph.D. thesis, FAU Erlangen, 2010, [http://d-nb.info/1009840509/34](http://d-nb.info/1009840509/34).

[SS78] Lynn Arthur Steen and J. Arthur Seebach, Jr., *Counterexamples in topology*, second ed., Springer-Verlag, New York-Heidelberg, 1978. MR 507446

[Tit77] J. Tits, *A “theorem of Lie-Kolchin” for trees*, Contributions to algebra (collection of papers dedicated to Ellis Kolchin), Academic Press, New York, 1977, pp. 377–388. MR 0578488

[VC71] V. N. Vapnik and A. Ya. Chervonenkis, *On the uniform convergence of relative frequencies of events to their probabilities*, Theor. Probability Appl. 16 (1971), no. 2, 264–280, English translation of Russian original.

[Vil09] Cédric Villani, *Optimal transport*, Grundlehren der mathematischen Wissenschaften, vol. 338, Springer, Berlin-Heidelberg, 2009.