Bulletin
de la SOCIÉTÉ MATHÉMATIQUE DE FRANCE

SPACES OF ALGEBRAIC MEASURE TREES AND TRIANGULATIONS OF THE CIRCLE

Wolfgang Löhr & Anita Winter

Tome 149
Fascicule 1

2021

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

pages 55-117

Le *Bulletin de la Société Mathématique de France* est un périodique trimestriel
de la Société Mathématique de France.

Fascicule 1, tome 149, mars 2021

Comité de rédaction

Christine BACHOC
Yann BUGEAUD
François DAHMANI
Clothilde FERMANIAN
Wendy LOWEN
Laurent MANIVEL
Julien MARCHÉ
Kieran O'GRADY
Emmanuel RUSS
Béatrice de TILIÈRE
Eva VIEHMANN

Marc HERZLICH (Dir.)

Diffusion

Maison de la SMF
Case 916 - Luminy
13288 Marseille Cedex 9
France
`commandes@smf.emath.fr`

AMS
P.O. Box 6248
Providence RI 02940
USA
`www.ams.org`

Tarifs

Vente au numéro : 43 € ($ 64)
Abonnement électronique : 135 € ($ 202),
avec supplément papier : Europe 179 €, hors Europe 197 € ($ 296)

Des conditions spéciales sont accordées aux membres de la SMF.

Secrétariat : Bulletin de la SMF

Bulletin de la Société Mathématique de France
Société Mathématique de France
Institut Henri Poincaré, 11, rue Pierre et Marie Curie
75231 Paris Cedex 05, France
Tél : (33) 1 44 27 67 99 • Fax : (33) 1 40 46 90 96
`bulletin@smf.emath.fr` • `smf.emath.fr`

© *Société Mathématique de France 2021*

Tous droits réservés (article L 122–4 du Code de la propriété intellectuelle). Toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'éditeur est illicite. Cette représentation ou reproduction par quelque procédé que ce soit constituerait une contrefaçon sanctionnée par les articles L 335–2 et suivants du CPI.

ISSN 0037-9484 (print) 2102-622X (electronic)

Directeur de la publication : Fabien DURAND

Bull. Soc. Math. France
149 (1), 2021, p. 55-117

SPACES OF ALGEBRAIC MEASURE TREES AND TRIANGULATIONS OF THE CIRCLE

BY WOLFGANG LÖHR & ANITA WINTER

ABSTRACT. — In this paper, we present with *algebraic trees*, a novel notion of (continuum) trees that generalizes countable graph-theoretic trees to (potentially) uncountable structures. For this purpose, we focus on the tree structure given by the branch-point map, which assigns to each triple of points their branch point. We give an axiomatic definition of algebraic trees, define a natural topology, and equip them with a probability measure on the Borel-σ-field. Under an order-separability condition, algebraic (measure) trees can be considered as tree structure equivalence classes of metric (measure) trees (i.e., subtrees of $\mathbb{R}$-trees). Using Gromov-weak convergence (i.e., sample distance convergence) of the particular representatives given by the metric arising from the distribution of branch points, we define a metrizable topology on the space of equivalence classes of algebraic measure trees.

In many applications, binary trees are of particular interest. We introduce on that subspace with the sample shape and the sample subtree mass convergence two additional, natural topologies. Relying on the connection to triangulations of the circle, we show that all three topologies are actually the same, and the space of binary algebraic measure trees is compact. To this end, we provide a formal definition of triangulations of the circle and show that the coding map that sends a triangulation to an algebraic measure tree is a continuous surjection onto the subspace of binary algebraic nonatomic measure trees.

Texte reçu le 9 avril 2020, modifié le 10 juillet 2020, accepté le 21 juillet 2020.

WOLFGANG LÖHR, University of Duisburg-Essen, Mathematics, 45117 Essen, Germany •
E-mail : `wolfgang.loehr@uni-due.de`
ANITA WINTER, University of Duisburg-Essen, Mathematics, 45117 Essen, Germany •
E-mail : `anita.winter@uni-due.de`

Mathematical subject classification (2010). — 60B10, 05C05; 60D05, 54F50, 57R05, 60C05.

Key words and phrases. — Continuum tree, $\mathbb{R}$-tree, Metric tree, Branch-point map, Convergence of trees, Sample shape convergence, Gromov-weak convergence, Brownian CRT, β-splitting tree, Yule tree, State space.

Research supported by *DFG-RTG 2131* and by *DFG Priority Programme SPP 1590*. Wolfgang Löhr was supported by the *DFG project 415705084*.

© Société Mathématique de France
doi:10.24033/bsmf.2825

Résumé (*Espaces d'arbres algébriques mesurés et triangulations du cercle*). — Nous présentons dans cet article une nouvelle notion d'arbres (continus), appelés arbres algébriques, qui généralise celle des arbres dénombrables (en théorie des graphes) à des structures (potentiellement) indénombrables. Pour cela, nous nous intéressons uniquement à la structure d'arbre donnée par la fonction de branchement, qui à chaque triplet de points associe leur point de branchement. Nous définissons les arbres algébriques de manière axiomatique et les munissons d'une topologie naturelle ainsi que d'une mesure de probabilité sur la tribu borélienne. Sous une condition de séparabilité de la structure d'ordre, les arbres algébriques mesurés peuvent être considérés comme des classes d'équivalence d'arbres métriques mesurés (i.e. des sous-arbres de $\mathbb{R}$-arbres). À chaque arbre algébrique mesuré on peut associer un arbre métrique en considérant la distance générée par la distribution des points de branchement. En utilisant la convergence Gromov-faible (i.e. la convergence des distances échantillonnées) de ces arbres métriques mesurés associés, nous définissons une topologie métrisable sur l'espace des classes d'équivalence d'arbres algébriques mesurés.

Le cas des arbres binaires est particulièrement intéressant en termes d'applications. Nous introduisons sur ce sous-espace deux autres topologies naturelles, la convergence des cladogrammes engendrés par un échantillon de points de l'arbre et la convergence des masses des sous-arbres associés à un échantillon. En utilisant le lien avec les triangulations du cercle, nous montrons que ces trois topologies sont identiques, et que l'espace des arbres algébriques mesurés binaires est compact. Nous donnons pour cela une définition formelle des triangulations du cercle, et nous montrons que la fonction de codage qui à une triangulation associe un arbre algébrique mesuré est une surjection continue sur le sous-espace des arbres algébriques binaires munis d'une mesure diffuse.

1. Introduction

Graph-theoretic trees are abundant in mathematics and its applications, from computer science to theoretical biology. A natural question is how to define limits and limit objects as the size of the trees tends to infinity. On the one hand, there are *local* approaches yielding countably infinite graphs or generalized so-called graphings with a Benjamini–Schramm-type approach (going back to [10], see [49, Part 4]). On the other hand, if one takes a more *global* point of view, as we are doing here, the predominant approach is to consider graph-theoretic trees as metric spaces equipped with the (rescaled) graph distance. Then the limit objects are certain "tree-like" metric spaces, most prominently the so-called $\mathbb{R}$-trees introduced in [56]. They are also of independent interest, e.g., to study isometry groups of hyperbolic space ([52]) or as generalized universal covering spaces in the study of the fundamental groups of one-dimensional spaces ([30]). The characterization of the topological structures induced by $\mathbb{R}$-trees has received considerable attention ([51, 50, 28]). Here, instead of the topological structures, we are more interested in the "tree structures" induced by $\mathbb{R}$-trees. We formalize the tree structure with a branch point map and call the resulting axiomatically defined objects *algebraic trees.*

While, unlike for metric spaces, we do not know any useful notion of convergence for topological spaces or topological measure spaces, it is essential for us that we can define a very useful convergence of algebraic measure trees.

Our main motivation lies in suitable state spaces for tree-valued stochastic processes. The construction and investigation of scaling limits of tree-valued Markov chains within a metric space setup started with the continuum analogues of the Aldous–Broder algorithm for sampling a uniform spanning tree from the complete graph ([26]) and of the tree-valued subtree-prune and regraft Markov chain used in the reconstruction of phylogenetic trees ([27]). It continued with the construction of evolving genealogies of infinite size populations in population genetics ([37, 20, 44, 54, 38]) and in population dynamics ([35, 45]). Moreover, continuum analogues of pruning procedures were constructed ([2, 1, 48, 42, 43]). All these constructions have in common that they encode trees as metric (measure) spaces or bimeasure $\mathbb{R}$-trees, and equip the respective space of trees with the Gromov–Hausdorff ([39]), Gromov-weak ([34, 36, 46]), Gromov–Hausdorff-weak ([58, 8]), or leaf-sampling weak-vague topology ([48]).

In the present paper, we shift the focus from the metric to the tree structure for several reasons. First, checking compactness or tightness criteria for (random) metric (measure) spaces is not always easy, and some natural sequences of trees do not converge as metric (measure) spaces with a uniform rescaling of edge lengths. At least for the subspace of binary algebraic measure trees that we introduce, the situation is much more favorable, because it turns out to be compact. Second, the metric is often less canonical than the tree structure in situations where it is not clear that every edge should have the same length, e.g., in a phylogenetic tree, where edges might correspond to very different evolutionary time spans. Third, one might want to preserve certain functionals of the tree structure in the limit. For instance, the limit of binary trees is not always binary in the metric space setup, while this will be the case for our algebraic measure trees. Also, the centroid function used in [7] is not continuous on spaces of metric measure trees, but it is continuous on our space.

The starting point of our construction is the notion of an $\mathbb{R}$-tree (see [56, 22, 13, 24]). There are many equivalent definitions, but the following one is the most convenient for us:

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.,

$$r(x_1,x_2)+r(x_3,x_4) \le \max\big\{r(x_1,x_3)+r(x_2,x_4),\ r(x_1,x_4)+r(x_2,x_3)\big\} \tag{1}$$

for all $x_1,x_2,x_3,x_4 \in T$.

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