Advances in Applied Mathematics 67 (2015) 75–95

Contents lists available at ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama

Navigation in tree spaces

Satyan L. Devadoss [a,*], Jack Morava [b]

[a] *Williams College, Williamstown, MA 01267, United States*
[b] *Johns Hopkins University, Baltimore, MD 21218, United States*

A R T I C L E I N F O

Article history:
Received 16 September 2014
Received in revised form 13 March 2015
Accepted 23 March 2015
Available online 13 April 2015

MSC:
14H10
92B10
52B11

Keywords:
Phylogenetics
Configuration spaces
Associahedron
Tree spaces

A B S T R A C T

The orientable cover of the moduli space of real genus zero algebraic curves with marked points is a compact aspherical manifold tiled by associahedra, which resolves the singularities of the space of phylogenetic trees. The resolution maps planar metric trees to their underlying abstract representatives, collapsing and folding an explicit geometric decomposition of the moduli space into cubes, endowing the resolving space with an interesting canonical pseudometric. Indeed, the given map can be reinterpreted as relating the *real* and the *tropical* versions of the Deligne–Knudsen–Mumford compactification of the moduli space of Riemann spheres.

1. Introduction

A classical problem in computational biology is the construction of a phylogenetic tree from gene sequence alignments for n species. Billera, Holmes, and Vogtmann [3] construct a space BHV_n of such metric trees, with nonpositive curvature; this makes geometric methods (geodesics, centroids) available. Unfortunately, their space is not a

* Corresponding author.
E-mail addresses: satyan.devadoss@williams.edu (S.L. Devadoss), jack@math.jhu.edu (J. Morava).

http://dx.doi.org/10.1016/j.aam.2015.03.006

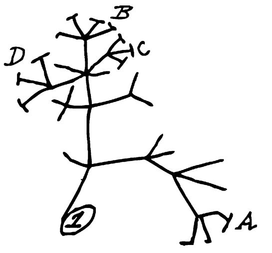

Fig. 1. An example from Darwin's notebook (1837).

manifold but a cone over a relatively singular simplicial complex. This paper introduces a compact orientable hyperbolic manifold $\overline{\mathcal{M}}^{\mathrm{or}}_{0,n+1}(\mathbb{R})$, tiled by $n!$ Stasheff associahedra K_n, which resolves the singularities of BHV_n. Indeed, this map [Theorem 9] can be reinterpreted as relating the *real* and the *tropical* versions of the Deligne–Knudsen–Mumford compactification of the moduli space of Riemann spheres.

Our manifold of interest is the orientation cover of the moduli space $\overline{\mathcal{M}}_{0,n+1}(\mathbb{R})$ of real stable genus zero algebraic curves marked with distinct smooth points; these objects can also be interpreted as compactified moduli spaces of ordered configurations of points on the real projective line [11], but we understand them here as spaces of rooted metric trees with labeled leaves, as in Fig. 1. One of the objectives of this paper is to provide $\overline{\mathcal{M}}_{0,n+1}(\mathbb{R})$ with a pseudometric naturally associated to the geometry of the trees it parametrizes.

These spaces have some remarkable similarities to the *space forms* which captured the imagination of mathematicians in the late nineteenth century [16]: they are $K(\pi, 1)$ manifolds with very pretty tessellations. The work of Klein and other geometers soon found applications in physics and mechanics; we hope our constructions may be useful for the study of big data in general, and biology [12] in particular.

Sections 2 and 3 of this paper summarize work of Billera, Holmes, and Vogtmann on spaces of abstract metric trees, and the corresponding classical work of Boardman and Vogt on spaces of planar trees. Sections 4 through 6 use previous work of Devadoss to construct an *orientable* moduli space of planar trees (a prerequisite for statistical applications), and Section 7 concerns the (equivariant) topology of the resolving map from the space of planar to abstract trees.

2. Spaces of abstract trees

A *metric tree* is a tree with a nonnegative weight assigned to each of its *internal* edges [6]. Billera, Holmes, and Vogtmann [3] have constructed a space BHV_n of isometry classes of metric trees with n labeled leaves (plus a *root*). When such a tree is binary, it has $n - 2$ internal nodes, and hence $2n - 2$ nodes in all. Diaconis and Holmes [13] have shown that such trees, with all nodes labeled, are in bijection with the $(2n-3)!!$ possible (unordered) pairings of those nodes. This enumeration defines coordinate patches for the space of such trees, parametrized by the weights of their internal edges. The resulting