Ken'ichi Ohshika
Athanase Papadopoulos *Editors*

In the Tradition of Thurston

Geometry and Topology

Springer

In the Tradition of Thurston

Ken'ichi Ohshika • Athanase Papadopoulos
Editors

In the Tradition of Thurston

Geometry and Topology

Springer

Editors
Ken'ichi Ohshika
Department of Mathematics
Gakushuin University
Tokyo, Japan

Athanase Papadopoulos
Institut de Recherche Mathématique
Avancée
CNRS et Université de Strasbourg
Strasbourg, France

ISBN 978-3-030-55927-4 ISBN 978-3-030-55928-1 (eBook)
https://doi.org/10.1007/978-3-030-55928-1

Mathematics Subject Classification: 32G15, 30G60, 57M25, 57M50, 57M60, 57M07

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This is the first of a series of three volumes consisting of essays on Thurston's contribution to mathematics, its development and its impact.

The present volume contains 16 chapters. Some of them are surveys of Thurston's works on several topics, including knot theory, geometrization of 3-manifolds, Kleinian groups, circle packings, the complex projective geometry of surfaces, and laminar groups. Other chapters are overviews of works that are directly inspired by Thurston's ideas. They include topics such as the dynamical and counting problems for curves on hyperbolic surfaces, the study of surfaces of infinite type and their mapping class groups, the complex-analytic geometry of Teichmüller space, a stratification of moduli spaces of polynomials, and there are two chapters dedicated to the recent activity on anti-de Sitter geometry and quasi-Fuchsian co-Minkowski manifolds, two theories whose development follows closely Thurston's ideas that he introduced in his study of hyperbolic geometry.

All the chapters in this volume are self-contained and peer-reviewed. They are intended to be references for students and researchers who want to learn Thurston's works and ideas.

We take this opportunity to thank the various colleagues and friends who participated in the preparation of this volume, in particular the authors of the various chapters and those who read and refereed them. We are especially grateful to Vincent Alberge for his enthusiasm and help at an early stage of this project. We also thank Elena Griniari for her kind editorial support.

Whereas the present volume carries the subtitle "Geometry and Topology," those of the two forthcoming volumes will be "Geometry and Dynamics" and "Geometry and Groups."

These three volumes are above all a tribute to Bill Thurston, for his unique way of perceiving forms and patterns and of communicating, writing mathematics, and sharing it with others. They are an expression of our gratitude for him and our fascination for his work; he is one of the greatest (if not the greatest) geometers of modern times.

Tokyo, Japan
Strasbourg, France
June 2020

Ken'ichi Ohshika
Athanase Papadopoulos

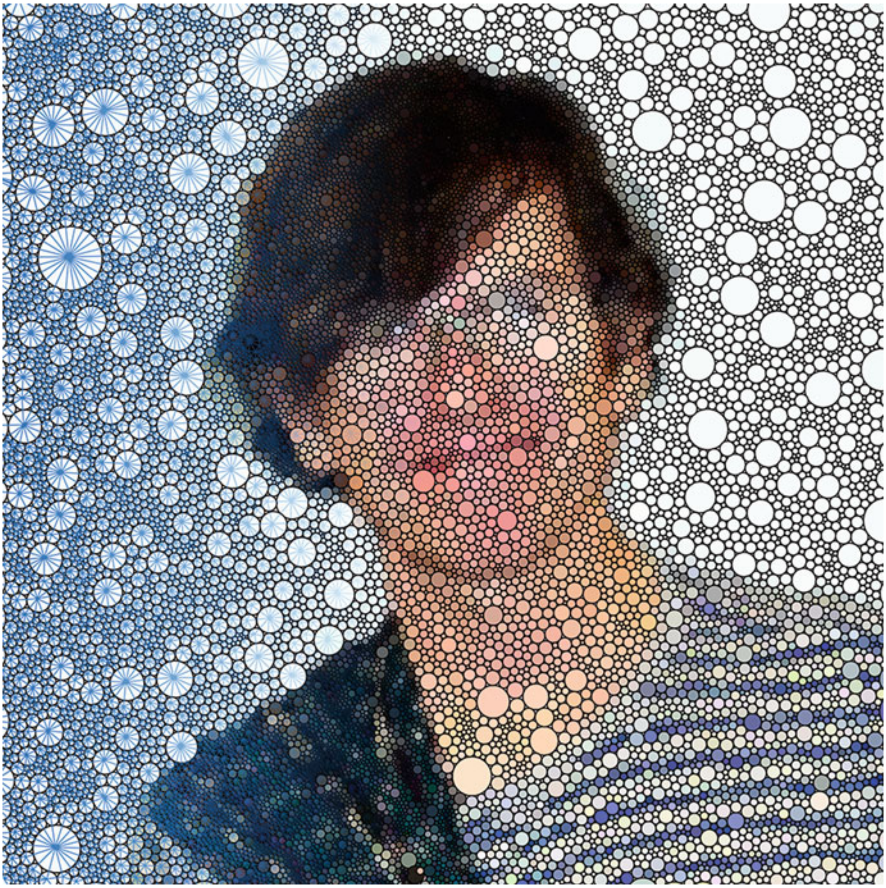


Photo @Philip Bowers

Introduction

A casual mathematician reading the title of this book may ask: "What is the tradition of Thurston?" He may also ask: "In what field is there a tradition of Thurston?"

Let us start by commenting on these questions.

The second question is easier to answer than the first. In fact, with his personal style, Thurston marked indelibly geometry, topology, group theory, and dynamics. His work also had a strong impact on ergodic theory, complex analysis, and discrete geometry. It also influenced combinatorics, algebraic geometry, computer science, and mathematics education. It is not that Thurston passed from one topic to another, but he considered all these topics as a single one; he was able to see the unity of mathematics. Above all, Thurston inaugurated a new way of understanding, communicating mental images, sharing ideas, and writing mathematics. He had a personal and unconventional opinion on what is mathematics about, and on why we do mathematics. He liked to share his mathematics and also his ideas on all sorts of questions. In response to a person asking on MathOverflow: "What can one (such as myself) contribute to mathematics?" Thurston responded:

> It's not mathematics that you need to contribute to. It's deeper than that: how might you contribute to humanity, and even deeper, to the well-being of the world, by pursuing mathematics? Such a question is not possible to answer in a purely intellectual way, because the effects of our actions go far beyond our understanding. We are deeply social and deeply instinctual animals, so much that our well-being depends on many things we do that are hard to explain in an intellectual way. That is why you do well to follow your heart and your passion. [...] The product of mathematics is clarity and understanding. Not theorems, by themselves. [...] The world does not suffer from an oversupply of clarity and understanding (to put it mildly). How and whether specific mathematics might lead to improving the world (whatever that means) is usually impossible to tease out, but mathematics collectively is extremely important. [...] In short, mathematics only exists in a living community of mathematicians that spreads understanding and breathes life into ideas both old and new. The real satisfaction from mathematics is in learning from others and sharing with others. All of us have clear understanding of a few things and murky concepts of many more. There is no way to run out of ideas in need of clarification. The question of who is the first person to ever set foot on some square meter of land is really secondary. Revolutionary change does matter, but revolutions are few, and they are not self-sustaining—they depend very heavily on the community of mathematicians.

Giving an answer to the first question, on what is the tradition of Thurston, is not easy. Skimming the various chapters of the present volume will hopefully give a good idea of what the mathematics underlying this tradition is about. Our wish in publishing this book and the sequel is to contribute to the perpetuation of Thurston's ideas.

The present volume is divided into 16 chapters, of which we now give an outline.

Chapter 1, by Ken'ichi Ohshika and Athanase Papadopoulos, titled *A Glimpse into Thurston's Work*, consists of two parts. The first part is an overview of a portion (but definitely not all) of Thurston's work. The topics considered include foliations, contact structures, symplectic structures, 1-dimensional dynamics, 3-manifolds, geometric structures, geometrization of cone-manifolds, Dehn surgery, Kleinian groups, the Thurston norm, holomorphic dynamics, complex projective geometry, circle packings and discrete conformal geometry, word processing in groups, automata, tilings, computer science, mapping class groups, Teichmüller spaces, and fashion design.

The second part of this chapter is a glimpse into some development and impact of Thurston's work. We start by a short report on the proofs of several conjectures which were either formulated by Thurston or in which Thurston's works played a crucial role. This includes the proofs of the Smith conjecture, of Ahlfors's measure-0 conjecture, of Marden's tameness conjecture, of the ending lamination conjecture, of the density conjecture, of the geometrization conjecture, of Waldhausen's conjecture on the structure of fundamental groups of closed irreducible 3-manifolds, on the virtual-Haken conjecture, of the virtual fibering conjecture, and of the Ehrenpreis conjecture. We then report on other topics whose development used Thurston's ideas in an essential way. They include works on the Cannon–Thurston maps, on Anti-de Sitter geometry, on linkages, on higher Teichmüller theory, and on the Grothendieck–Thurston theory.

Chapter 2, by Ken'ichi Ohshika, titled *Thurston's Influence on Japanese Topologists up to the 1980s*, is a report on the impact of Thurston's work on Japanese school of topology and geometry, in particular on foliations and on hyperbolic manifolds.

Chapter 3, by Makoto Sakuma, titled *A Survey of the Impact of Thurston's Work on Knot Theory* is a survey of the impact of Thurston's work on knot theory. The author chose, as a general theme for this chapter, the rigidity and flexibility in 3-manifolds, a reference to Chapter 5 of Thurston's Princeton lecture notes, titled "Flexibility and Rigidity of Geometric Structures." Hyperbolic structures on knot complements constitute the central object of this theory. This is motivated by the fact that, by Thurston's geometrization theorem for Haken manifolds, almost every knot in the 3-sphere is hyperbolic. In this context, the term "rigidity" refers to the Mostow–Prasad theorem saying that a complete finite-volume hyperbolic structure on any manifold of dimension ≥ 3 is a topological invariant, whereas the term "flexibility" refers to the deformation theory of incomplete or infinite-volume hyperbolic structures.

In his survey, Sakuma reviews successively the main tools that were known before Thurston started working on knot theory, then Thurston's contribution to the subject, then some of the major developments which arose from Thurston's ideas.

Among the classical notions that are mentioned, one can find the Alexander polynomial, tools from finite group theory, ideas arising from works of Reidemeister, Schubert, Conway, and others on knots, as well as some important theorems in 3-manifold topology. Sakuma then reviews Thurston's extensive use of hyperbolic geometry and the associated rigidity theorem, the notions of hyperbolic Dehn filling and orbifold fundamental group, deformation spaces of hyperbolic structures equipped with various topologies, character varieties, the classification of mapping classes, the eight 3-dimensional geometries, Thurston's uniformization theorem for Haken manifolds which led to the proof of the Smith conjecture, group actions on trees, the Thurston norm, Thurston's geometrization theorem for orbifolds, his geometrization conjecture for manifolds which became Perelman's theorem, his virtual fibering conjecture proved by Agol, McMullen's Teichmüller polynomial and its relation with the Alexander polynomial, the use of the homology of finite coverings of knot complements, the use of profinite completions of knot groups as knot invariants, and other notions.

At the same time, Sakuma's article constitutes a survey of the growth and development of knot theory in Japan.

Chapter 4 by Sadayoshi Kojima, titled *Thurston's Theory of 3-Manifolds*, is a survey of Thurston's work on the geometry and topology of 3-manifolds. After mentioning some landmarks from the pre-Thurston era on this subject, in particular, the works of Waldhausen and Jaco–Shalen–Johansson, the review focuses on Thurston's hyperbolic Dehn surgery theorem, his uniformization theorem for Haken manifolds and his work towards the virtual fiber and the virtual-Haken conjectures. The last section of the chapter is a short report on the solution of several conjectures formulated by Thurston, including the geometrization and the virtual fiber conjectures.

Chapter 5, by Philip Bowers, is titled *Combinatorics Encoding Geometry: The Legacy of Bill Thurston in the Story of One Theorem*. The author gives a comprehensive historical and geometrical survey of the so-called Koebe–Andreev–Thurston theorem on circle packings and its impact. This theorem is the main result proved in Chapter 13 of Thurston's Princeton lecture notes. Thurston's work on circle packings turned out to be at the foundations of the field which became known as *discrete conformal geometry*, now an extremely active field.

After a detailed presentation of the Koebe–Andreev–Thurston theorem, Bowers develops the theory of infinite circle packings on non-compact surfaces, reviewing several related topics and results, including the Rodin–Sullivan discrete approach to the Riemann mapping theorem based on circle packings (a result they obtained based on an idea of Thurston) and the characterization of hyperbolic polyhedra using circle packings, a theory which also originates in Thurston's work, leading to the works of several authors, including Hodgson–Rivin, Schramm, Bao–Bonahon, J. Bowers, K. Pratt, the author himself, and others.

In Chap. 6, titled *On Thurston's Parameterization of $\mathbb{C}P^1$-Structures*, Shinpei Baba surveys some ideas of Thurston on complex projective structures on surfaces and their moduli spaces. He reviews in particular Thurston's parametrization of the moduli spaces of marked projective structures by the product of measured lamination space with Teichmüller space. Interestingly, this parametrization of a space of 2-dimensional geometric structures involves a passage to dimension 3, namely the convex hull construction for subsets of the projective plane seen as the boundary of hyperbolic 3-space, in which the theory of pleated surfaces appears as a central object. The author also surveys Goldman's parametrization of the space of marked projective structures with Fuchsian holonomies as well as other notions inspired by Thurston's ideas on complex projective structures.

The convex hull construction mentioned in Chap. 6 is reviewed again in Chap. 7, titled *A Short Proof of an Assertion of Thurston Concerning Convex Hulls*. In this chapter, Graham Smith provides a short proof of the result of Thurston saying that if X is a closed subset of the Riemann sphere considered as the boundary of 3-dimensional hyperbolic space, then the intrinsic metric of the boundary of the convex hull of X is hyperbolic. Thurston stated his result and gave a heuristic proof in Chapter 8 of his Princeton lecture notes. The proof proposed by Smith is clear and short.

Chapter 8 by Cyril Lecuire is titled *The Double Limit Theorem and Its Legacy*. It is a survey of a series of results on deformation spaces of Kleinian groups inspired by two theorems of Thurston: the double limit theorem and the compactness theorem for hyperbolic structures on acylindrical 3-manifolds. These two theorems are used as important steps in Thurston's proof of his uniformization theorem for Haken manifolds. While outlining the proofs of these two theorems, the author reviews several key tools introduced by Thurston in his work on deformation spaces of Kleinian groups (laminations, train tracks, pleated surfaces, etc.) and at the same time he surveys Thurston's theorems known under the names of "Broken windows only theorem" and "Window frame bounded theorem."

Besides reviewing this work of Thurston, the author reports on some historical background, including the classical works of Ahlfors and Bers on quasi-conformal deformations which led to the so-called Ahlfors–Bers coordinates on quasi-conformal deformation spaces. He then reviews several results that are inspired by Thurston's ideas, including works of Culler, Morgan, and Shalen on deformations of hyperbolic manifolds and compactifications of deformation spaces using actions of Λ-trees, Otal's alternative proof of Thurston's double limit theorem, works of Ohshika on deformation spaces of Kleinian groups and their boundaries, and others.

Chapter 9, by Ken'ichi Ohshika and Teruhiko Soma, titled *Geometry and Topology of Geometric Limits I*, is an answer to a question asked by Thurston. More precisely, the authors provide a complete classification of isometry classes of hyperbolic 3-manifolds corresponding to geometric limits of Kleinian surface groups which are isomorphic to the fundamental group of a finite-type hyperbolic surface. They introduce some invariants for hyperbolic 3-manifolds which are geometric limits of Kleinian surface groups, and they prove that the homeomorphism type and the end invariants of such manifolds determine the isometry type. This result is an analogue of the ending lamination theorem for the case of finitely generated Kleinian groups. It gives a possible solution to problem No. 8 in the list of 24 problems raised by Thurston in his paper *Three Dimensional Manifolds,* *Kleinian Groups and Hyperbolic Geometry* (1982),[1] which he formulated as follows: "Analyze limits of quasi-Fuchsian groups with accidental parabolic." As a matter of fact, the result by Ohshika and Soma gives a complete answer to this question in the form it is interpreted by Otal in his survey of Thurston's paper which appeared in 2014,[2] in which he provides an update of Thurston's 24 problems.

Chapter 10, by Hyungryul Baik and Keyeongro Kim, titled *Laminar Groups and 3-Manifolds*, is concerned with a theory that Thurston developed in his 1998 preprint *Three-Manifolds, Foliations and Circles, I*, where he associates to a transversely oriented taut foliation on a closed atoroidal 3-manifold a faithful map from the fundamental group of the manifold to the group of orientation-preserving homeomorphisms of the circle, equipped with an invariant lamination. Here, a lamination on the circle is a collection of pairs of points on this circle with an unlinkedness condition. If one realizes these pairs by straight lines in the disc (or, equivalently, by hyperbolic geodesics, if the circle is seen as the boundary of the Poincaré model of the hyperbolic plane), then one gets a lamination in the usual sense. The authors study more generally the notion of laminar groups, that is, groups acting on the circle with invariant laminations. An elementary example of such a group is a surface group, and some fundamental groups of 3-manifold groups are also laminar. Baik and Kim review Thurston's work on this topic and its development by several authors. They give a condition which insures that a laminar group is not virtually abelian.

Chapter 11, by Viveka Erlandsson and Caglar Uyanik, is titled *Length Functions on Currents and Applications to Dynamics and Counting*. It is concerned with geodesic currents on closed orientable surfaces of finite type. In this setting, geodesic currents are measure-theoretic objects which generalize closed curves. The authors survey various functions on these spaces, for which they use the generic term "length functions." The definitions are motivated by analogously defined functions on spaces of closed curves on surfaces, including the length of a closed geodesic for a given Riemannian (possibly singular) metric, the geometric intersection with a fixed filling closed curve or a collection of curves, the word length in the fundamental group, the stable length, etc. The authors then present an overview of the applications of these length functions to counting problems. In particular, they show how to generalize Mirzakhani's formulae on the asymptotic growth rate of the number of curves of bounded length to the other notions of length. They also study actions of pseudo-Anosov mapping classes on spaces of geodesic currents, showing that such an action has a uniform North-South dynamics, like the one on Teichmüller space equipped with Thurston's boundary.

Chapter 12, by Javier Aramayona and Nicholas Vlamis, titled *Big Mapping Class Groups: An Overview*, is a survey of certain mapping class groups of surfaces of

[1] W. P. Thurston, Three dimensional manifolds, Kleinian groups and hyperbolic geometry, Bull. Amer. Math. Soc. 6 (1982), 357–381.

[2] J.-P. Otal, William P. Thurston: "Three-dimensional manifolds, Kleinian groups and hyperbolic Geometry," Jahresber. Dtsch. Math.-Ver. 116 (2014), 3–20.

infinite type (which the authors call "big mapping class groups"). This includes the group of compactly supported mapping classes, the Torelli groups, the Teichmüller modular groups (or quasi-conformal mapping class groups), and the asymptotic mapping class groups. The authors address questions related to curve graphs and other complexes associated to infinite-type surfaces, surveying several known results on these complexes, and they study their simplicial automorphism groups, reviewing recent works by several authors on these questions, and formulating an extensive number of open questions, on the algebraic and geometric properties of big mapping class groups. They also mention relations with Thompson groups, a subject which was dear to Thurston.

The questions, techniques, and results surveyed in this chapter are all motivated by analogous techniques and results which are known in the context of mapping class groups of surfaces of finite type. The study is based on tools introduced by Thurston on mapping class groups of surfaces of finite type, adapted to this setting.

Chapter 13, by Hideki Miyachi, titled *Teichmüller Theory, Thurston Theory, Extremal Length Geometry, and Complex Analysis* is an exposition of some complex-analytical aspects of Teichmüller space, including extremal length and the Bers boundary, and the Gardiner–Masur boundary, with applications to end invariants and the ending lamination conjecture. One topic discussed in some detail is the use of extremal length geometry in the elaboration of a Poisson integral formula for holomorphic functions and pluriharmonic functions on the Bers compactification that the author developed in his recent works. He emphasizes the fact that this Poisson formula strengthens the connection between the topological and the complex-analytic aspects of Teichmüller theory, promoting a framework in which the combination of these methods sheds a new unified light on this space.

In Chap. 14, titled *Signatures of Monic Polynomials*, Norbert A'Campo constructs a real semi-algebraic cell-decomposition of the space of monic real polynomials of fixed degree. The face operators of this cell-decomposition are explicitly given. The classical discriminant of degree-d polynomials, as well as its complement, appears as a union of cells. Since this complement is a classifying space for the braid group, this work provides a finite complex from which we can compute the group cohomology of the braid group with integral coefficients. The work in this chapter is inspired by Thurston's ideas.

Chapters 15 and 16 concern the impact in pseudo-Riemannian geometry of techniques and tools introduced by Thurston.

Chapter 15 by Francesco Bonsante and Andrea Seppi, titled *Anti-de Sitter Geometry and Teichmüller Theory*, is a survey of n-dimensional anti-de Sitter geometry, with a special emphasis on dimension 3 and on its relation with Teichmüller theory and hyperbolic geometry. This subject was inaugurated in 1990 by G. Mess, who introduced in the study of anti-de Sitter geometry Thurston's ideas and tools that he used in the study of hyperbolic manifolds and Teichmüller spaces. Bonsante and Seppi start by giving a comprehensive introduction to anti-de Sitter geometry. This includes the Klein and quadric models, a study of isometries, geodesics, the universal cover, the boundary at infinity equipped with its conformal structure and Lorentzian metric, polarity, group actions, and Dirichlet domains. The authors emphasize the properties of the special case of dimension 2+1 with the ruling of the boundary surface, surveying Cauchy surfaces and globally hyperbolic timespaces, convexity and convex hull constructions, holonomy and developing maps, and several other notions of anti-de Sitter geometry. They present a self-contained exposition of the results of Mess on the classification of globally hyperbolic anti-de Sitter 3-manifolds containing a Cauchy compact surface and the construction of the Gauss map associated with spacelike surfaces in anti-de Sitter space. They survey the relation between pleated surfaces and the earthquake map, and they recover Mess's proof of Thurston's earthquake theorem. They review recent results on anti-de Sitter 3-dimensional geometry in relation with Teichmüller theory, following the ideas of Mess. They show that the Gauss map is conformal and harmonic. In particular, its image is a minimal surface. They survey the notion of volume in the setting of maximal globally hyperbolic Cauchy compact manifolds, comparing it to the volume of the associated convex core. They describe parameters for various deformation spaces in anti-de Sitter geometry in terms of holomorphic objects, and they apply them to the study of the universal Teichmüller space. They address the question of investigating anti-de Sitter theory with timelike cone singularities and its relation with the Teichmüller theory of hyperbolic surfaces with cone singularities. Some of the results in the last section answer questions raised in the list *Some open questions in anti-de Sitter geometry* (2012).[3]

Chapter 16, by Thierry Barbot and François Fillastre, titled *Quasi-Fuchsian co-Minkowski Manifolds*, is concerned with the geometry of co-Minkowski space, that is, the space of spacelike hyperplanes of Minkowski space. The authors study the geometry of this space equipped with the action of affine deformations of co-compact lattices of hyperbolic isometries, highlighting several facets of an analogy between this geometry and that of hyperbolic space equipped with the action of quasi-Fuchsian groups. The authors survey the known results, and at the same time they present new contributions to the theory.

Let us make a short review of these results.

The authors start by introducing a cylindrical affine model of co-Minkowski space which is analogous to the Klein ball model of hyperbolic space. They survey a convex core construction and the definition of a traceless fundamental form giving rise to a unique mean hypersurface (i.e. a surface with a traceless second fundamental form) contained in the convex core. They show that the mean distance between the lower boundary component of the convex core and this mean hypersurface gives rise to an asymmetric norm defined on the space of affine deformations of a certain lattice. By symmetrization, this asymmetric norm leads to a notion of volume of the convex core and to a "mean distance" between the future complete and the past complete flat globally hyperbolic maximal Cauchy compact spacetimes having the same holonomy. In dimension 2+1, this asymmetric norm turns out to be the earthquake norm defined by Thurston, and it is also the

[3]T. Barbot, F. Bonsante, J. Danciger, W. M. Goldman, F. Guéritaud, F. Kassel, K.Krasnov, J.-M. Schlenker, and A. Zeghib. Some open questions in anti-de Sitter geometry. 2012. arXiv:1205.6103.

total length of the bending lamination of the lower boundary component of the convex core. This allows the authors to give a new proof and a generalization to an arbitrary dimension of a theorem of Thurston saying that the total length function of measured geodesic laminations leads to an asymmetric norm, after identification of a measured lamination with a tangent vector to Teichmüller space using infinitesimal earthquakes. Furthermore, the volume of the convex core turns out to be the sum of the total length of the bending lamination of its boundary, a result which the authors compare with its analogues in hyperbolic and anti-de Sitter spaces.

Finally, Barbot and Fillastre describe a relation between the geometry of co-Minkowski space and the theory of Anosov representations. In particular, they investigate the Anosov character (in the sense of Labourie) of certain representations of hyperbolic groups of isometries of Minkowski space into the isometry group of this space.

We hope that the essays that constitute the various chapters of this volume, besides giving an idea of what Thurston's tradition is, will contribute in keeping it alive. Let us conclude with some words of Thurston, extracted from a brief essay that he wrote on beauty, mathematics, and creativity:[4]

> Many people think of mathematics as austere and self-contained. On the contrary, mathematics is a very rich and very human subject, an art that enables us to see and understand deep interconnections in the world. The best mathematics uses the whole mind, embraces human sensibility, and is not at all limited to the small portion of our brains that calculates and manipulates with symbols. Through pursuing beauty we find truth, and where we find truth, we discover incredible beauty.

Tokyo, Japan — Ken'ichi Ohshika
Strasbourg, France — Athanase Papadopoulos

[4]The essay was distributed on the occasion of a fashion show that took place at the Salon du Carrousel du Louvre in Paris, in March 2010, in which a collection of the Japanese fashion designer Issey Miyake was presented, a collection whose conception was based on Thurston's vision on the eight 3-dimensional geometries.

Contents

Chapter 1
A Glimpse into Thurston's Work

Ken'ichi Ohshika and Athanase Papadopoulos

Abstract We present an overview of some significant results of Thurston and their impact on mathematics.

Keywords Foliation · Contact structure · Symplectic structure · Conformal geometry · Holomorphic motion · Kleinian groups · Circle packing · Automatic groups · Tiling · Mapping class groups · Teichmüller space · Fashion design · Linkage · Anti-de Sitter geometry · Higher Teichmüller theory · Grothendieck–Thurston theory · Asymmetric metric · Schwarzian derivative · Computer science · Ehrenpreis conjecture · Transitional geometry · 3-Manifolds · Geometric structures · (G,X)-structures · Dehn surgery · Hyperbolic geometry · Thurston norm · Smith conjecture · Cannon–Thurston map · Discrete conformal geometry · Discrete Riemann mapping theorem

AMS Codes 57N10, 57M50, 20F34, 20F65, 22E40, 30F20, 32G15, 30F60, 30F45, 37D40, 57M25, 53A40, 57D30, 58D05, 57A35, 00A30, 01A60, 20F10, 68Q70, 57M05, 57M07, 57Q15, 57D15, 58A10, 58F10, 65Y25

1.1 Introduction

In this chapter, we present an overview of some significant results of Thurston and their impact on mathematics.

The chapter consists of two parts. In the first part, we review some works of Thurston, grouped in topics. The choice of the topics reflects our own taste and our

---

K. Ohshika (✉)
Department of Mathematics, Gakushuin University, Tokyo, Japan
e-mail: ohshika@math.gakushuin.ac.jp

A. Papadopoulos
Institut de Recherche Mathématique Avancée, CNRS et Université de Strasbourg, Strasbourg, France
e-mail: papadop@math.unistra.fr

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_1

degree of knowledge. The choice of the order of these topics was almost random. Indeed, it is not clear whether a given topic is more important than another one, and there are interconnections and mutual influences between most of these topics. Furthermore, it was not possible to follow a chronological order because Thurston was thinking about all these subjects simultaneously.

In the second part of this chapter, we report briefly on the proofs of some conjectures which were either formulated by Thurston or whose solution depended in a crucial way on his work. We also discuss a few topics whose development was directly or intellectually influenced by ideas of Thurston.

We have included at some places remarks and quotations which give an idea of Thurston's approach to science in general and to the aesthetics of mathematics.

Our exposition will certainly be too short at some places, for some readers who know little about Thurston's work on the topic discussed, and it will be redundant for readers familiar with this topic (and even more for the experts). We apologize in advance to both categories of readers. We have added here and there some historical notes, whenever we felt this was useful. These notes will probably be beneficial to both groups of readers.

1.2 On Thurston's Works

1.2.1 Foliations and Groups of Homeomorphisms

The first time the word "foliation" was used in a mathematical sense (in its French version, *feuilletage*) took place by the end of the 1940s by Georges Reeb and Charles Ehresmann (who was Reeb's advisor).[1] Reeb, in his dissertation [189], gave the first example of a foliation on the 3-sphere.[2]

When Thurston came into the subject, examples and constructions of foliations on special manifolds were available. To describe the situation in short, one can say that in a lapse of time of 5 years, he obtained all the general existence results that were hoped for. In this section, we briefly review his work on the subject.

[1] In their first papers on the subject, Ehresmann and Reeb used the expression *éléments de contact de dimension p complètement intégrables* ("completely integrable contact elements of dimension *p*.").

[2] The question of the existence of a foliation on the 3-sphere was asked by Heinz Hopf in 1935, who certainly did not use the word "foliation". The first example was given in a joint paper by Ehresmann and Reeb, but Ehresmann always attributed this construction to Reeb.

Thurston's first published paper on foliations[3] is a short note titled *Noncobordant foliations of* S^3 [206], which appeared in 1972. In this note, Thurston proved that any closed 3-manifold carries a family of foliations whose Godbillon–Vey invariant takes all possible real values. This invariant (an element of the real 3-cohomology of the manifold) was discovered the year before, by C. Godbillon and J. Vey, who came across it by manipulating differential forms. The question of whether there exist foliations with non-zero Godbillon–Vey invariant was soon raised, together with the problem of giving it a geometrical interpretation. Thurston's result closed both problems. It is interesting to note that Thurston's existence proof has a strong hyperbolic geometry flavor: he constructed a family of foliations of the 3-sphere which depend on convex polygons in the hyperbolic plane whose area is equal to the Godbillon–Vey invariant of the foliation. Besides hyperbolic geometry, we can find in this proof another ingredient which was soon to become fundamental in Thurston's work, namely, the notion of singular hyperbolic surface. In his paper on foliations, Thurston described these surfaces as "surfaces having a number of isolated corners, with metrics of constant negative curvature everywhere else."

Thurston formulated his result on the Godbillon–Vey invariant as the surjectivity of a certain homomorphism from the group of cobordism classes of foliations onto $\mathbb{R}$. As a corollary, he proved the existence of an uncountable family of non-cobordant foliations on S^3. The precise results are stated in terms of Haefliger's classifying spaces of Γ-structures. These objects, also called Haefliger structures, were introduced by André Haefliger in his thesis published in 1958 [95], as a generalization of the notion of foliation.[4] In short, Haefliger structures, which may be interpreted as singular C^r foliations, are $\mathbb{R}^k$-bundles over n-dimensional manifolds equipped with foliations transverse to the fibers. (Such a Haefliger structure is said to be of codimension k). A natural example of a Haefliger structure is the normal bundle to a foliation. Haefliger structures are the natural setting for the theories of classifying spaces and of characteristic classes of foliations, and Haefliger's theory reduces the question of the existence of certain classes of

[3]With the exception of his PhD thesis, which was defended the same year and which remained unpublished. The thesis, whose title is *foliations of 3-manifolds which fiber over a surface.* was submitted to the Swiss journal *Inventiones*. The referee asked for modifications; Thurston did not comply and withdrew the paper. Haefliger, in the collective Thurston memorial article [85], writes: "The referee suggested that the author should give more explanations. As a consequence, Thurston, who was busy proving more theorems, decided not to publish it."

[4]A posteriori, it is interesting to read Richard Palais' comments on this notion, in his MathSciNet review of Haefliger's paper: "The first four chapters of the paper are concerned with an extreme, Bourbaki-like generalization of the notion of foliation. After some 25 pages and several hundred preliminary definitions, the reader finds that a foliation of X is to be an element of the zeroth cohomology space of X with coefficients in a certain sheaf of groupoids. While such generalization has its place and may in fact prove useful in the future, it seems unfortunate to the reviewer that the author has so materially reduced the accessibility of the results, mentioned above, of Chapter V, by couching them in a ponderous formalism that will undoubtedly discourage many otherwise interested readers." In fact, the notion that Haefliger introduced turned out to be of paramount importance.

foliations to that of certain maps between manifolds and classifying spaces. A consequence of Thurston's work on Haefliger structures is that in some sense (up to a natural condition on normal bundles) the class of Γ-structures is not different from that of foliations.

Two years later, Thurston published three papers in which he proved a series of breakthrough results on foliations. In the paper [210], titled *The theory of foliations of codimension greater than one*, working in the setting of codimension-k Haefliger structures for $k > 1$, he showed the existence of a large class of completely integrable plane fields on manifolds which led to the construction of new classes of foliations. In particular, he obtained that any plane field of codimension greater than one is homotopic to a completely integrable C^0 field. He also proved that for any $n \geq 3$ and any $1 < k \leq n/2$, if the sphere S^n carries a k-plane field then it carries also a C^∞ foliation of dimension k. These results are wide generalizations of existence results obtained in special cases by Reeb, Tamura, Lawson, Phillips, Haefliger, and others.

The second paper by Thurston published in 1974 [208] is titled *A generalization of the Reeb stability theorem*. The theorem, obtained by Reeb in his thesis [189] which we already mentioned, says that if a transversely oriented codimension-1 foliation on a compact manifold has a two-sided compact leaf with finite fundamental group, then all the leaves of this foliation are diffeomorphic. Thurston obtained a much stronger result under the hypothesis that the foliation is of class C^1, namely, he proved that one can replace the hypothesis that the compact leaf has finite fundamental group by the one saying that the first real cohomology group of the leaf is zero. At the same time, he showed that, under the same conditions, the leaves of this foliation are the fibers of a fibration of the manifold over the circle or the interval. Furthermore, he showed by an example that this result does not hold in the C^0 case. Thurston approached the problem by studying the linear holonomy around the compact leaf, and he gave an interpretation of the result in terms of the linearity properties at a fixed point for a topological group acting continuously in the C^1 topology, as a group of C^1 diffeomorphisms of a manifold.

The relation between foliations and groups of diffeomorphisms is announced in the title of the third paper published in the same year [209]: *Foliations and groups of diffeomorphisms*. In this paper, Thurston studied higher codimension Haefliger structures in relation with groups of diffeomorphisms of arbitrary manifolds, generalizing a relation discovered by John Mather between the group of compactly supported diffeomorphisms of the real line and framed codimension-one Haefliger structures. Using the techniques of classifying spaces, Thurston proved that any two C^∞ foliations of a manifold arising from nonsingular vector fields are homotopic as Haefliger structures if and only if their normal bundles are isotopic. At the same time, he announced a result giving a precise relationship between the classifying space of codimension-k Haefliger structures and that of the diffeomorphism group of compact manifolds of dimension k as a discrete group. He used this result to prove that the diffeomorphism group of a compact manifold is perfect (that is, equal to its commutator subgroup), generalizing a result obtained by Mather in the case of 1-dimensional manifolds.

One may recall here that at the time Thurston was working on these topics, the study of the algebraic structure of groups of diffeomorphisms and homeomorphisms of compact manifolds was a very active subject of research, involving mathematicians such as John Mather, David Epstein, Michel Herman, Jean Cerf and others.

In the same year, Thurston gave a talk at the Vancouver ICM (1974) whose title was *On the construction and classification of foliations* [211]. In the paper published in the proceedings of the congress, Thurston reviewed some of the major results he had obtained and he announced the results of his forthcoming paper [213], *Existence of codimension-one foliations*, which appeared in 1976. In this paper, he proved the existence of a C^∞ codimension-one foliation on any closed manifold whose Euler characteristic is zero. This result may be contrasted with a result of Haefliger [95] stating that there is no codimension-one real-analytic foliation on a sphere of any dimension. This displays a striking difference between the C^∞ and the real-analytic cases. In the same paper, Thurston proved that on any closed manifold without boundary, every hyperplane field is homotopic to the tangent plane field of a C^∞-foliation.

Naturally, the ICM paper, which is only 3 pages long, is written in the pure Thurston style, warm, unconventional and appealing to the reader's imagination. It starts with the following:

> Given a large supply of some sort of fabric, what kinds of manifolds can be made from it, in a way that the patterns match up along the seams? This is a very general question, which has been studied by diverse means in differential topology and differential geometry.

It is also not surprising that the definition of a foliation that Thurston gives in this paper is informal and unusual:

> A foliation is a manifold made out of a striped fabric—with infinitely thin stripes, having no space between them. The complete stripes, or "leaves," of the foliation are submanifolds; if the leaves have codimension k, the foliation is called a codimension-k foliation.

ICM talks are intended for a general audience, but very few mathematicians were able, like Thurston was, to describe the objects they were studying in simple words, avoiding notation and formulae.

Thurston's paper [219] which appeared in 1986 and which is titled *A norm for the homology of 3-manifolds*, is the foundational paper on the so-called Thurston norm, and it also contains results and conjectures on foliations of 3-manifolds. The results include a classification of codimension-1 foliations without holonomy that are transverse to the boundary, in terms of the top-dimensional faces of the unit ball of Thurston's norm on homology. We shall talk about this in Sect. 1.2.9 below. Thurston showed that if such a foliation has no Reeb component, then any compact leaf is norm-minimizing in its homology class. A converse was obtained by D. Gabai soon after, who also obtained a general existence result for codimension-1 transversely oriented foliations transverse to the boundary with no Reeb components [84]. In the same paper, Gabai proved several conjectures of Thurston.

In 1976, Thurston published a paper with Joseph Plante on the growth of germs of diffeomorphisms [185]. The study of such germs was motivated by the theory of foliations (the germs appear in the holonomy groups of foliations). The question that Plante and Thurston studied was motivated by the work of Milnor [148] who introduced the notion of growth of a finitely generated group. It was conjectured that a group has polynomial growth if and only if it contains a nilpotent subgroup of finite index. The conjecture was proved by Gromov in a paper which was a major breakthrough [89]. In their article, Plante and Thurston showed that the conjecture is true for germs of diffeomorphisms, and they gave applications of this result to foliations. One of the consequences they obtained is that if a compact manifold with fundamental group of polynomial growth carries a transversely oriented and real-analytic codimension-1 foliation, then its first real homology group is nontrivial. This generalized a result of Haefliger [95].

One should also talk about Thurston's mostly unpublished work on extending the theory of foliations to that of laminations and essential laminations in 3-manifolds which was developed by several authors, following his ideas.

We mention Thurston's manuscript [231] on taut foliations, that is, codimension-1 foliations for which there exists an embedded closed curve that is transverse to the foliation and intersects every leaf.[5] In this paper, Thurston associates to every transversely orientable taut foliation on a closed atoroidal 3-manifold a faithful homomorphism from the fundamental group of the manifold onto the group of orientation-preserving homeomorphisms of the circle which preserves a pair of dense invariant laminations of the circle (in an appropriate sense), and which is universal in some precise sense. This theory gave rise to developments by several authors, see in particular the paper [49] by Calegari and Dunfield in which the authors give a new proof of Thurston's result and where Thurston's ideas are made more precise. In the same paper, the authors show that there are other classes of essential foliations and laminations than those which were considered by Thurston that give rise to faithful actions on the circle. We also refer the reader to the exposition in Chap. 10 by Baik and Kim in the present volume [16]. See also the book [48] by Calegari on foliations which contains several sections explaining Thurston's homomorphism to the group of orientation-preserving homeomorphisms of the circle. As a matter of fact, this books is a valuable reference for many aspects of Thurston's theory of foliations and laminations, including his work on the cohomology of the group of orientation-preserving homeomorphisms of the circle, his stability result for the group of orientation-preserving homeomorphisms of the interval, his construction of foliations on 3-manifolds using a triangulation of the manifold, the Thurston norm in relation with foliations, and several other topics.

Talking about Thurston's work on foliations, one should also mention measured foliations on surfaces and his construction of the space of measured foliations, a space equipped with a topology which makes it homeomorphic to a Euclidean space of the same dimension as Teichmüller space. The space of measured foliations

[5] The terminology *taut foliation* is also used for a foliation such that the ambient manifold admits a Riemannian metric for which the leaves are minimal surfaces. In his paper [201], Sullivan showed that for a C^2 taut foliation (in the above sense) on a closed orientable 3-manifold, there exists a Riemannian metric on the ambient manifold that is taut in this Riemannian sense.

became a central object in low-dimensional geometry and topology. We review this in Sect. 1.2.15 below.

1.2.2 Contact and Symplectic Geometry

A contact structure on a differentiable manifold is a field of hyperplanes in the tangent bundle satisfying a "complete non-integrability" condition that makes it, locally, unrealizable as a hyperplane field tangent to a foliation. (Note that the non-integrability property of these hyperplanes is in contrast with the dimension-1 case: vector fields are locally always integrable.) It is easy, although not trivial, to produce examples of contact structures. Several examples are discussed in Thurston's book on the geometry and topology of 3-manifolds [224].

Contact geometry, like symplectic geometry which we shall discuss below, originates in classical mechanics, and it has applications in geometric optics, thermodynamics and other domains of physics. In fact, this notion can be traced back to the work of Gaston Darboux. One of his results is often quoted, viz. the fact that a contact structure is always locally equivalent to a standard contact structure [67].

The usual definition of a contact structure is algebraic, formulated in terms of differential forms. In fact, a foliation and a contact structure are both defined locally by a differential form α, but in the case of a foliation, α satisfies $\alpha \wedge d\alpha \equiv 0$ whereas in the case of a contact function, it satisfies $\alpha \wedge d\alpha \neq 0$ at every point (and if the relation is replaced by $\alpha \wedge d\alpha > 0$, with respect to a given orientation, we say that we have a "positive" contact form). The problem of classifying contact structures on manifolds arose naturally. Thurston writes in his monograph [224, p. 168]: "[contact structures] give an interesting example of a widely occurring pattern for manifolds that is hard to see until your mind and eyes have been attuned." Several pages of his book [224] are dedicated to the effort of sharing with the reader an intuitive picture of contact structures. On p. 172 of this book, to give a physical sense of the contact structure on the tangent circle bundle of a surface, he uses models related to ice skating and bicycling, dedicating several paragraphs to these images.

Thurston obtained a number of important results on contact structures. His first paper on this subject is a joint paper with H. E. Winkelnkemper [212], titled *On the existence of contact forms*, published in 1975. It contains a very short proof of a result, which was already obtained by Robert Lutz and by Jean Martinet in 1971 [132, 138], saying that every closed orientable 3-manifold carries a contact structure. Thurston and Winkelnkemper deduced this result from a classical result, namely, the so-called "open-book decomposition theorem" of Alexander [9].

Even though, in some sense, a contact structure is the complete opposite of a foliation, the two subjects are related. With Yakov Eliashberg, Thurston introduced the notion of *confoliation* in dimension 3, and he developed a theory which gives a hybrid setting for codimension-1 foliations and contact structures on 3-manifolds [73, 74]. A confoliation in this sense interpolates between a codimension-1 foliation and a contact structure. The techniques that Eliashberg and Thurston developed allowed them to prove that any C^2 codimension-1 foliation on a 3-manifold, except for the product foliation $S^1 \times S^2$, can be approximated in the C^0 sense by positive contact structures. Confoliations appear in a crucial manner in the proof of this result, since the main step consists in the modification of the plane field tangent to a foliation so that it becomes integrable (tangent to a foliation) in some part of the manifold and a positive contact structure in the complement. It is interesting to note that at the same time he was developing confoliations, Thurston developed a theory of foliations of three-manifolds that are hybrids of fibrations over the circle and foliated circle bundles over surfaces, see his 1997 preprint [231].

Contact structures are defined on odd-dimensional manifolds, and their analogues on even-dimensional manifolds are symplectic structures.

This brings us to Thurston's work on symplectic geometry.

When Thurston started working in this field, the questions of the existence of symplectic structures on closed manifolds and that of Kähler metrics on symplectic manifolds were the main problems. In 1976, Thurston gave the first examples of compact symplectic manifolds that do not admit any Kähler metric. He presented his examples in a short note titled *Some simple examples of symplectic manifolds* [214]. The examples were later called Kodaira–Thurston manifolds, since it turned out that the manifolds described by Thurston were already known to Kodaira (who used them for other purposes). At the same time, Thurston gave a counter-example to a claim made by Heinrich Guggenheimer, saying that the odd-dimensional Betti numbers of symplectic manifolds are necessarily even. The examples that Thurston gave have odd first Betti numbers. The odd-dimensional Betti numbers of Kähler manifolds are all even. After Thurston's discovery, the question of characterizing the symplectic manifolds which admit no Kähler structure became a very active research field (works of Robert Gompf, Dusa McDuff, etc.).

One may also mention here a result of Thurston on volume-preserving diffeomorphisms of differentiable manifolds. This theory is related to symplectic geometry, since a symplectomorphism of a $2n$-dimensional manifold preserves the volume form obtained as the n-th power of the symplectic form. In a preprint titled *On the structure of the group of volume preserving diffeomorphisms*, first circulated in 1972 [207], Thurston proved that the group of volume-preserving diffeomorphisms of a manifold is perfect provided the first homology group of the manifold is zero, and he introduced at the same time a certain number of ideas that became later very useful in symplectic geometry. Although Thurston's preprint remained unpublished, the techniques it contains and the questions it raises had a profound impact in symplectic geometry (works of Augustin Banyaga, of Dusa McDuff, etc.). Banyaga, in his paper [17] and in his book [18, p. 125ff], developed many ideas of Thurston and gave detailed proofs of several of his results in symplectic geometry.

1.2.3 One-Dimensional Dynamics

The first published work by Thurston on dynamics is his paper with Milnor *On iterated maps of the interval*, which appeared in 1988 [150]. An early version of the paper, containing more material, was circulated in 1977.[6] Some results stated as a conjecture in the preprint version became theorems in the published version.

In this paper, Milnor and Thurston studied the dynamics of continuous piecewise (strictly) monotone maps of the interval, to which they associated a certain number of naturally and very simply defined invariants. These invariants became at the basis of kneading theory, an important element in the theory of dynamics of unimodal maps of the interval. Let us recall some of the notions they introduced.

Given a map f of an interval I, a *lap* of f is a maximal sub-interval of I on which f is monotone. This leads to the notion of *lap number* $\ell = \ell(f)$ of f. Milnor and Thurston studied the *growth* of the lap number of the iterates of f, that is, the limit $\lim_{k\to\infty} \ell(f^k)^{1/k}$. By a theorem of Misiurewicz and Szlenk, this limit is equal to the topological entropy of f. Milnor and Thurston introduced an invariantly defined "formal coordinate function" $\theta(x)$ which is given for x in I, as a formal power series $\sum \theta_k(x)t^k$, where if $f^k(x)$ belongs to the interior of the j-th lap I_j, the coefficient $\theta_k(x)$ is the formal symbol I_j multiplied by ± 1 or 0 according as to whether f^k is increasing or decreasing, or has a turning point at x. This led them to a basic invariant called the *kneading matrix* of f, an $(\ell + 1) \times \ell$ matrix with entries in the ring $\mathbb{Z}[[t]]$ of integer formal power series, with its associated *kneading determinant*, a power series with odd integer coefficients, $D(t) = 1 + D_1(t) + D_2(t) + \ldots$. There is a close relation between the kneading determinant and the behavior of the periodic points of the map. In the simplest case where f has only one turning point (which is the critical point of the map), the coefficients of $D(t)$ are either $+1$, -1 or 0 according to whether the iterate f^{k+1} has a local minimum or a local maximum at c. In the same paper, Milnor and Thurston gave a method for computing the sequence of lap numbers of the iterates of f in terms of the kneading matrix. They studied the convergence of the kneading determinant, showing for example that for $s > 1$, this power series is holomorphic in the unit disc, and has a smallest zero at $t = 1/s$ where $s = s(f)$ is the growth number of the map. Under the same hypothesis ($s > 1$), they showed that f is topologically semi-conjugate to a piecewise linear map having slope $\pm s$ everywhere. The Artin–Mazur zeta function encodes the periodic orbits of f.

Milnor and Thurston used methods of Julia and Fatou, before these methods found their place in the revival of holomorphic dynamics that took place a few years later. They proved what they call their *main theorem*, which allows a computation of the Artin–Mazur zeta function in terms of the kneading determinant. They gave several applications of their theory.

[6] Leo Jonker, in his Mathscinet review of this paper writes: "If there were a prize for the paper most widely circulated and cited before its publication, this would surely be a strong contender. An early handwritten version of parts of it was in the reviewer's possession as long ago as 1977".

A particularly important class of examples of maps to which the Milnor–Thurston applies is the one of maps of lap number two, *unimodal real maps*. A typical family of such maps is the family of quadratic polynomials $x \mapsto x^2 + c$. Each map in this family has a unique critical point, and the kneading sequence describes the location of the sequence of images of this critical point, to the left or right of this critical point. For the family of quadratic polynomials, Milnor and Thurston gave a characterization of power series that can occur as a kneading determinant, they discussed continuity properties of the growth number $s(f)$, and they obtained a monotonicity result for the entropy. Furthermore, the paper contains several algorithms to compute the entropy of a piecewise monotone map.

This is now the occasion for us to quote Milnor from the preface and the dedication to Thurston that he wrote, in Volume VI of his *Collected Works* [149].[7] In the preface, Milnor writes: "I was introduced to Dynamical Systems by Bill Thurston in the late 1970s and found the field so engaging that it was hard to escape from." In the dedication, Milnor writes:

> My interactions with Bill followed a consistent pattern. He would propose a mathematical statement which I found amazing, but extremely unlikely. However, the evidence would accumulate, and sooner or later I would have to concede that he was completely right. My introduction to the field of dynamics proceeded in exactly this way. Bill had been intrigued by the work of Robert May in theoretical ecology. May had proposed that the population of some insect species in successive years behaved in a chaotic way, which could be described by a very simple mathematical model, in which next years population is expressed as a universal modal function of this years population. Bill developed this idea by constructing symbol sequences associated with unimodal maps. He claimed that many quite different looking one-parameter families of unimodal maps would give rise to the same family of symbol sequences. I didn't believe a word of this, but couldn't find a counter-example. Eventually, I was convinced, and we collaborated on the paper "On iterated maps of the interval."

The notions that Milnor and Thurston introduced in their paper remain until now part of the most important tools for the study of the dynamics of maps of the interval. Their paper continues to be a source of inspiration for the works done in this field. A large volume of literature is devoted to the generalization of their results for unimodal maps to maps with a larger number of laps. Furthermore, kneading sequence theory, as a way of encoding combinatorial information, was applied in the study of complex dynamics, by Milnor and others.

Thurston's last published paper [228][8] is on dynamics. He wrote it before his death, a period where, according to Milnor, "Bill was entering a period of renewed creativity, full of ideas and eager to communicate them." [149, p. IX] The paper is titled *Entropy in dimension one*, and it contains new ideas and results in this field. One of the results that Thurston obtained is a characterization of positive numbers that arise as the topological entropy of postcritically finite self-maps of the interval. Precisely he proved that these are exactly the numbers h such that $\exp(h)$ is an

[7]The volume is dedicated to Thurston.

[8]The paper was published posthumously in 2014.

algebraic integer that is at least equal to the absolute value of any conjugate of $\exp(h)$. He also showed that the map can be chosen to be a polynomial whose critical points are all in the open interval $(0, 1)$. At the same time, the paper makes it clear what are the phenomena of 1-dimensional dynamics that are relevant for entropy.

Thurston used in this paper a number of ideas and notions from his previous works: the central role played by postcritically finite maps, train tracks for graphs together with train track maps and the operations of zipping and splitting of train tracks (ideas originating in his theory of surface dynamics, under the version adapted by Bestvina and Handel in their study of outer automorphism groups of free groups), a generalization of the notion of pseudo-Anosov maps, Perron–Frobenius matrices and Pisot and Salem numbers, notions that appear in Thurston's theory of surface automorphisms.

1.2.4 The Topology of 3-Manifolds

We start with a few words on the pre-Thurston era.

In the 1960s, a new direction of research in 3-manifolds was started by Haken and Waldhausen [97, 238]. Among the objects of their research is the class of compact irreducible 3-manifolds containing incompressible surfaces. These manifolds were called by Waldhausen sufficiently large 3-manifolds; now they are called Haken manifolds. In particular, Waldhausen proved that any homotopy equivalence between two closed Haken manifolds is homotopic to a homeomorphism. In the 1970s, Jaco, Shalen and Johannson developed a theory of decomposing Haken 3-manifolds along incompressible tori and annuli, setting the basis of a theory now called Jaco–Shalen–Johannson theory [106, 107]. They showed that any Haken manifold can be decomposed along (a possibly empty) union of disjoint incompressible tori in such a way that each piece is either a Seifert fibered manifold or an atoroidal manifold, i.e. a 3-manifold which contains no non-peripheral immersed incompressible tori. (Here an immersed surface is said to be incompressible when the map induces a monomorphism between the fundamental groups.)

Through the work of Andreev, Riley and Jørgensen, Thurston already noticed that there are many 3-manifolds that admit complete hyperbolic metrics. He considered that this should be the case in much more generality. Indeed he proved a "uniformization theorem for Haken manifolds," stating that any atoroidal Haken manifold which is closed or which has torus boundaries carries a complete hyperbolic metric of finite volume. His proof of this theorem is very intricate and long. The argument is divided, at a large scale, into two cases: the first is when the manifold is not a surface bundle over the circle and the second is when it is. The first case is proved by induction involving Maskit's combination theorem. The second case is proved using his own theorem called the "double limit theorem," which is itself an important contribution to the theory of Kleinian groups, and which we describe in Sect. 1.2.8. One of the remarkable consequences of the uniformization theorem is the resolution of the Smith conjecture, which we shall review in Sect. 1.3.1.

Thurston conjectured that the same kind of uniformization theorem should hold for all closed 3-manifolds, and not only for Haken ones. This conjecture was formulated in the form of a "geometrization conjecture" which includes the Poincaré conjecture as a very special case. The geometric structures to which Thurston referred are locally homogeneous metrics. He gave the list of eight kinds of three-dimensional geometric structures. Six among them can be carried only by Seifert fibered manifolds. The two remaining ones are the hyperbolic geometry and the solvable geometry. Only torus bundles over the circle can carry a solvable geometry. The geometrization conjecture says that every compact 3-manifold is decomposed along incompressible tori into 3-manifolds having geometric structures. In the case of a homotopy sphere, this is equivalent to the Poincaré conjecture.

At the end of his the expository paper [216], Thurston gave a list of problems on 3-manifolds and Kleinian groups. The problems on 3-manifolds contain the above-mentioned geometrization conjecture (and in particular the Poincaré conjecture), and the virtual-Haken conjecture, which says that every closed irreducible 3-manifold has a finite cover which is Haken. This question was first posed by Waldhausen (see [239]). The list also contains quite a new and unexpected conjecture, now called the "virtual fibering conjecture," which was proved by Agol more than a quarter of a century later.

The list as a whole has been the driving force of all research in 3-manifolds and Kleinian groups for more than 30 years after its appearance.

1.2.5 (G, X)-Structures and Geometric Structures

In the Erlangen programme, Klein proposed a new way of thinking geometry. According to him, geometry consists of a base space and a group acting on it. Although he did not think of general manifolds (the notion of manifold did not exist yet), we can regard his work as the origin of (G, X)-structures. A formal definition of a (G, X)-structure first appeared in Ehresmann's work. For a geometrico-historical exposition, we refer the reader to Goldman's article [88].

Given a space X and a group G acting on X by homeomorphisms, a manifold is said to have a (G, X)-structure when it is equipped with an atlas whose charts are maps into X with transition maps being restrictions of elements of G. Thurston made this notion central in low-dimensional topology by giving many important and interesting examples of (G, X)-structures on manifolds. By his work, this notion moved to the forefront of research.

Geometric structures are typical (G, X) structures, where X is a homogeneous space and G is its isometry group. As we mention in Sect. 1.2.4, Thurston showed that there are eight kinds of geometric structures in dimension 3, and conjectured that every compact 3-manifold can be decomposed along incompressible tori into submanifolds having geometric structures.

Geometric structures are Riemannian structures, i.e., the stabilizer of G at a point in X is compact. Thurston also studied non-Riemannian (G, X)-structures, above all complex projective structures on surfaces. This is the case where $X = \mathbb{C}P^2$ and $G = \mathrm{PSL}(2, \mathbb{C})$. The space of (marked) complex projective structures modulo isotopies on a closed surface Σ had been studied from the viewpoint of complex analysis using Schwarzian derivatives by Bers, Kra and Marden among others. Thurston gave a new parametrization for this space based on a more geometric approach, which has the form of a homeomorphism between this space and the product $\mathcal{T}(\Sigma) \times \mathcal{ML}(\Sigma)$, where $\mathcal{T}(\Sigma)$ denotes the Teichmüller space of Σ and $\mathcal{ML}(\Sigma)$ the space of measured laminations on Σ. This more geometric approach to the space of complex projective structures opened up a new flourishing field, which should be called a topological study of projective structures. The reader is referred to Sect. 1.2.11 of the present chapter.

1.2.6 Geometrization of Cone-Manifolds

After proving the uniformization theorem for Haken manifolds, Thurston tackled the general geometrization problem by a quite different approach. Since non-Haken manifolds do not contain incompressible surfaces, there is no way to cut them into simpler ones. Instead, Thurston introduced the technique of deforming the structure of a cone manifold by increasing its cone angle. For instance, to prove that a non-Haken atoroidal 3-manifold M has a hyperbolic structure, we would take a hyperbolic knot K in M (i.e. a knot K such that $M \setminus K$ has a complete hyperbolic metric, which is guaranteed to exist by virtue of the uniformization theorem for Haken manifolds combined with Myers' theorem [166]), and consider a deformation of the complete hyperbolic structure on $M \setminus K$ to a cone hyperbolic structure whose singular locus is the knot K, with cone angle θ. If we could deform the cone-hyperbolic structure until the cone angle becomes 2π, then we would be able to show that M is hyperbolic. Of course, this strategy should break down in general, for K may not be isotopic to a closed geodesic in a hyperbolic 3-manifold.

What Thurston really proved can be expressed as follows, if we only consider the case where the singularity is a knot and there is no incompressible torus disjoint from the singularity. Suppose that M is a closed irreducible 3-manifold containing a prime knot K, and that we are given an angle $\theta = \pi/n$ on K. We consider a 3-orbifold $(M, K(\theta))$ whose underlying space is M and whose singular set is K with cone angle θ. Thurston proved that in this situation $(M, K(\theta))$ has some (possibly empty) disjoint incompressible Euclidean 2-sub-orbifolds which decompose M into geometric 3-orbifolds. For simplicity, we consider the case where $M \setminus K$ is atoroidal. To prove the theorem (in this case), Thurston considered a deformation of the hyperbolic cone structure by increasing the cone angle on K, starting from the complete hyperbolic metric on M, which is regarded as the cone angle 0. If the angle reaches θ without degeneration, then $(M, K(\theta))$ is a hyperbolic orbifold. Thurston showed that if the degeneration occurs, then $(M, K(\theta))$ admits either a decomposition along an incompressible Euclidean 2-sub-manifold or a geometric structure other than the hyperbolic one. To prove the last step, in a special case, Thurston made use of the Ricci flow and of Hamilton's theorem [98] to get a spherical structure in the limit. Thurston's geometrization conjecture, including the Poincaré conjecture, was solved later by Perelman using precisely these Ricci flows, based on the idea of Hamilton. It is noteworthy that Thurston already noticed the usefulness of Ricci flows back in the 1980s.

The geometrization theorem of 3-orbifolds implies that if a closed prime 3-manifold has a finite group action with one-dimensional fixed point set then it has a geometric structure which is invariant under the action. This is (quite a huge) generalization of the Smith conjecture.

1.2.7 Dehn Surgery

Besides the Smith conjecture which we mentioned above, Thurston's work had a great impact on knot theory, through his theory of hyperbolic Dehn surgery. Dehn surgery is a classical tool in knot theory. The definition is simple: performing a Dehn surgery along a knot means that we take a tubular neighborhood of the knot and glue back the removed solid torus in such a way that the boundary of the meridian is in a homotopy class (called the meridian slope) on the boundary of the tubular neighborhood different from the original. In this way, we get a new 3-manifold. Lickorish [128] proved that if we consider a link instead of a knot, we can get any 3-manifold from the 3-sphere in this way. This tool has been heavily used in both knot theory and 3-manifold topology.

Thurston introduced hyperbolic geometry into the theory of Dehn surgery. First of all, his uniformization theorem for Haken manifolds implies that any knot that is neither a satellite knot nor a torus knot has a complement which has a complete hyperbolic metric. Such knots are called hyperbolic knots. For a hyperbolic knot K, Thurston considered Dehn surgeries along K, and proved that except for finitely many slopes, the manifolds obtained by surgeries are all hyperbolic. (This theorem is called the hyperbolic Dehn surgery theorem.) An interesting feature in the proof of this result is that it does not use the uniformization theorem, once we know the complement of K has a complete hyperbolic metric. Also among those manifolds obtained by hyperbolic Dehn surgery, there are non-Haken manifolds, whose hyperbolicity cannot be shown by the uniformization theorem.

More generally, by using the same techniques as those used in the proof of the hyperbolic Dehn surgery theorem, Thurston showed that for any complete hyperbolic 3-manifold of finite volume having a torus cusp, one can obtain a hyperbolic 3-manifold by attaching a solid torus to a cusp. Such an operation gives a hyperbolic 3-manifold except for finitely many homotopy classes of the attaching disc. Combining this theorem with his result on the Gromov invariant, which we shall present in Sect. 1.2.9, he obtained the fact that the set of the volumes of hyperbolic 3-manifolds constitutes an ordered subset of $\mathbb{R}$ isomorphic to ω^ω. The volumes of hyperbolic 3-manifolds are important objects in 3-manifold topology and there is still a large amount of activity taking place on this topic.

Hyperbolic Dehn surgery gives a better framework to understand deformations of hyperbolic cone structures. The homotopy classes of simple closed curves on a torus can be regarded as co-prime lattice points on $\mathbb{R}^2$. Therefore, the set of Dehn surgeries on a hyperbolic knot K can be identified with such lattice points. In this picture hyperbolic cone structures on S^3 whose singularities are isotopic to K can be identified with a subset of the x-axis. In this respect, Thurston considered a hyperbolic Dehn surgery space which contains both hyperbolic cone structures and hyperbolic Dehn surgeries.

1.2.8 Kleinian Groups

The notion of Kleinian group was first introduced by Poincaré as a generalization of the notion of Fuchsian group. Kleinian groups were extensively studied from the viewpoint of complex analysis in 1960–1975 by Ahlfors, Bers, Kra, Maskit, Marden among others. In particular, Bers considered the compactification of a slice in quasi-Fuchisan space, which is now called the Bers compactification of Teichmüller space [26], and Marden considered the deformation space of convex cocompact representations [134].

In the process of proving the uniformization theorem for Haken manifolds, Thurston needed to develop the theory of deformation spaces of Kleinain groups and ends of hyperbolic 3-manifolds. Thurston's proof of the uniformization theorem is largely divided into two cases, the one where the manifold is fibered over the circle and the other when it is not. In the first case, treated in [232], he proved what is called the double limit theorem. In the second case, he showed the compactness theorem for deformation spaces of hyperbolic structures on acylindrical 3-manifolds, proved in [226], and its relative version, proved in [233]. All these theorems are very important in the study of deformation spaces. They were generalized and became the fundamental tools in Kleinian group theory. The reader may refer to Chap. 8 of the present volume, by Lecuire [127].

Another important part in Thurston's proof is the analysis of geometrically infinite Kleinian groups, in which Thurston introduced the notion of geometric tameness using pleated surfaces. A geometrically infinite (torsion-free and finitely generated) Kleinian group is said to be geometrically tame when every geometrically infinite end of the corresponding hyperbolic 3-manifold has a sequence of pleated surfaces tending to it. Thurston showed that if a Kleinian group is geometrically tame, then the corresponding hyperbolic 3-manifold can be compactified by adding a boundary component to each of its ends. This property is called the topological tameness.

Thurston considered that one can prove Marden's tameness conjecture [134] saying that every complete hyperbolic 3-manifold with finitely generated fundamental group is topologically tame, by showing that every Kleinain group is geometrically tame. He gave a proof of this conjecture in the special case of groups that are algebraic limits of quasi-Fuchsian groups.

In [216], Thurston gave a list of 13 problems on Kleinian groups, one of which is Marden's tameness conjecture described above. We shall say more about these problems in Sect. 1.3.2.

1.2.9 The Thurston Norm, the Gromov Norm and the Gromov Invariant

As the work of Waldhansen, Haken, Jaco, Shalen and Johannson show, incompressible surfaces are important tools to study 3-manifolds. In the same vein, understanding Seifert surfaces is essential for knot theory. For every knot in the 3-sphere, the complement has first homology group isomorphic to $\mathbb{Z}$, and its Poincaé dual is represented by a Seifert surface.

Thurston introduced in [219] a pseudo-norm on the second homology groups (or on the first cohomology groups) of 3-manifolds, which is called today the Thurston norm. Given a second homology class c of the manifold, its Thurston norm is defined to be $x(c) = \max\{\min_S(-\chi(S)), 0\}$, where S ranges over all surfaces representing this homology class. (In the case of 3-manifolds with boundary, it is more reasonable to consider their relative homology groups.) It should be noted that a surface realizing the Thurston norm is always incompressible, but the converse does not hold in general.

This simple idea led to very interesting results. The norm can be extended first to the second homology with rational coefficient by defining $x(c/r)$ to be $x(c)/r$ for any integral homology class c, and by continuity to homology with real coefficients. The unit ball in the second (relative) homology group with real coefficients constitutes a convex polytope with vertices on lattice points when the manifold is irreducible, atoroidal, and acylindrical. This is called the Thurston norm polytope.

In a 3-manifold which fibers over the circle with fibers having negative Euler characteristic, the Euler class of the bundle of planes tangent to the fibers defines a second cohomology class. Thurston proved that the Thurston dual norm of such an Euler class is always equal to 1. This implies that for any fibering over the circle, the second homology class represented by a fiber lies in the interior of a facet of the Thurston norm polytope. As a corollary, this implies that a 3-manifold fibered over the circle (whose fibers have negative Euler characteristic) contains an incompressible surface which cannot be a fiber for any fibering over the circle. More generally, if a 3-manifold has a transversely oriented codimension-one foliation, we can consider the Euler class of the bundle of planes tangent to the leaves. Thurston showed that if this foliation does not have Reeb components, then the Thurston dual norm of the Euler class is less than or equal to 1, and that in particular if such a foliation has a compact leaf, then the dual norm is equal to 1. He also showed that a compact leaf of such a foliation realizes the Thurston norm of its second homology class. The notion of Thurston norm and these results gave the foundation of further studies of fibrations and foliations in 3-manifolds.

We now consider the Gromov norm.

To introduce the notion of bounded cohomology in [90], Gromov defined a pseudo-norm on homology groups as follows. Given a singular chain $s = \sum a_i c_i$ with real coefficients, define its norm $\|s\|$ to be $\sum_i |a_i|$. Then for any homology class σ, its Gromov norm is defined to be $\inf\{\|s\| \mid [s] = \sigma\}$. The same definition works also for relative homology groups. In particular, for a closed manifold, the Gromov norm of its fundamental class is called its Gromov invariant. Gromov proved that for a closed manifold of dimension n, its Gromov invariant is equal to its volume divided by a constant v_n depending only on n, where v_n is equal to the supremum of the volumes of n-dimensional hyperbolic simplices.

Thurston took up Gromov's invariant as the topic of Section 6 of his lecture notes [215]. (This was before the publication of Gromov's paper [90], in which we can find Thurston's influence, both explicitly and implicitly.) Thurston generalized Gromov's result to negatively curved manifolds, where the Gromov norm is bounded below by the volume divided by a constant depending only on n, and to manifolds with geometric structures, where equality between the Gromov norm and volume holds, but with a constant depending on the geometry. Thurston also considered a relative version of the Gromov invariant for manifolds with boundary. Using this in the case of hyperbolic 3-manifolds with torus cusps, he showed that the operation of hyperbolic Dehn surgery decreases the volume. He also proved that a torus cusp corresponds to an accumulation point of the distribution of volume, and that this is the only way for volumes of hyperbolic 3-manifolds to accumulate.

The Gromov norm for the second homology group of a 3-manifold is related to the Thurston norm. Thurston conjectured that if we change embedded surfaces in the definition of the Thurston norm to immersed surfaces, then the norm obtained should coincide with the Gromov norm. Gabai proved in [84] that this is indeed the case.

1.2.10 Conformal Geometry and Holomorphic Dynamics

Conformal geometry, since its birth, is intertwined with topology. The relation between the two subjects started in Riemann's doctoral dissertation (1851) [190] in which he introduced the concept of Riemann surface (as we call it today), as a branched coverings of the complex plane or of the Riemann sphere. His work was partly motivated by the problem of describing a general method for finding domains of definitions for multivalued functions $w(z)$ of a complex variable z defined by algebraic equations of the form $f(w, z) = 0$, so that a multi-valued function becomes single-valued. (This is the original meaning of the word "uniformization".) Thanks to the work of Riemann, analytic functions became objects that are no more necessarily defined on the complex plane or on subsets of it. With him, the concept of Riemann surface, with the closely related notion of analytic continuation, were born. Riemann further developed his ideas on this topic in his paper on Abelian functions [191]. At the same time, he introduced a number of topological notions that he used in the theory of functions of a complex variable: connectedness, degree of connectivity, genus, the classification of closed orientable surfaces, etc. One should also remember that there was still no notion of manifold in those times, and a surface could not be simply defined as a 2-manifold. Riemann also studied moduli of Riemann surfaces. He discovered that the number of such moduli, for a surface of genus g, is $3g-3$. He also proved the famous *Riemann mapping theorem*, saying that any simply connected open subset of the complex plane, provided it is not the whole plane, is biholomorphically equivalent to the open unit disc. The problem of characterizing topologically analytic functions arose (this was called later on the "Brouwer problem"). This problem was also a motivation for the development of quasiconformal mappings. Indeed, Grötzsch, Lavrentieff and Teichmüller, the three founders of the theory, tried to prove for quasiconformal mappings some results that were known to hold for conformal mappings, with the idea that it was only the topological form of the mapping that matters and not the fact that it was conformal.

It was natural that Thurston got attracted by this field. One of the first problems that he asked, when he joined the community of MathOverflow is related to Riemann surfaces and rational functions on the Riemann sphere. He wrote the following (posted on September 10, 2010): "I would like to understand and compute the shapes of rational functions, or equivalently, ratios of two polynomials, up to Moebius transformations in both domain and range." He also formulated the following more precise problem: "Given a set of points to be the critical values [in the range], along with a covering space of the complement homeomorphic to a punctured sphere, the uniformization theorem says this Riemann surface can be parametrized by S^2, thereby defining a rational function. Is there a reasonable way to compute such a rational map?"

On holomorphic dynamics, we shall mention two results of Thurston. The first one has a discrete character; it is Thurston's topological characterization of rational maps among branched coverings of the sphere. The second one has a continuous character; it is his result with Sullivan on holomorphic motions.

Thurston's theorem on the characterization of rational maps, that he proved at the beginning of the 1980s was a preliminary (but huge) step towards the program he formulated 30 years later on MathOverflow. It gives a necessary and sufficient condition for a branched covering of the sphere which is postcritically finite, that is, such that the union of the forward orbits of the critical points is finite, to be homotopy equivalent to a rational map. Here, homotopy equivalence is defined in an appropriate and natural sense; the relation is called now Thurston's equivalence. Thurston's criterion is given in terms of the action of the covering map (by taking inverse images) on systems of homotopy classes of essential simple closed curves on the sphere with the postcritical set deleted. The necessary and sufficient condition refers to this action, and it now carries the name "absence of a Thurston obstruction."

Thurston obtained this theorem in 1982. He lectured on it on several occasions and he wrote notes that were widely circulated [218]. Although it was announced at some point that a final version of these notes will be published in the CBMS conference series of the AMS, the notes remained unpublished. Adrien Douady and John Hubbard wrote a proof of Thurston's theorem, following his outline, and they circulated it in preprint form in 1984. Their paper was eventually published in 1993 [68].

The proof of Thurston's theorem, like the proofs of some others of his geometrical important results, involves the construction of a weakly contracting self-map of a certain Teichmüller space (which in the case at hand is the one of the sphere with the post-critical set deleted). The main step in this proof is to show that in the absence of a Thurston obstruction, this map has a unique fixed point. The fixed point, when it exists, is the desired rational function.

There are several analogies between this theory of Thurston and his classification theory of surface homeomorphisms: the use of hyperbolic geometry, the construction of an action on Teichmüller space, the study of an action on the collection of homotopy classes of simple closed curves, the existence of invariant laminations, the use of quasiconformal mappings, the utilization of linear algebra in the study of the action on curves, and in particular the Perron–Frobenius theorem for nonnegative matrices, etc.

Many applications of Thurston's theorem were obtained by various authors. We mention as an example its use in the theory of mating of two polynomials (getting a rational map whose Julia set is obtained by gluing those of two postcritically finite polynomials). We refer the reader to the paper [47] by Xavier Buff, Guizhen Cui, and Lei Tan for a survey of Thurston's theorem, including a self-contained proof of a slightly generalized version of this theorem and an overview of its applications.

As another aspect of holomorphic dynamics that was touched upon by Thurston, we review now his result with Sullivan on holomorphic motions.

A holomorphic motion of a subset X of the complex plane $\mathbb{C}$ is a family of mappings $f_t : X \to \mathbb{C}$ parametrized by a complex number t (considered as a complex time parameter) varying in a domain T containing the origin and satisfying the following three properties: (1) for each fixed x, $f_t(x)$ is holomorphic in t; (2) for each fixed t, $f_t(x)$ is injective in x; (3) f_0 is the identity mapping of X. The motivation behind this definition is the wish to adapt the topological notion of isotopy to a holomorphic context. The main question addressed is to know whether a holomorphic motion of the subset X extends to a holomorphic motion of the complex plane $\mathbb{C}$. (This is a holomorphic analogue of the topological problem of extending an isotopy to an ambient isotopy.) Holomorphic motions were introduced by Mañé, Sad and Sullivan in [133].

The main result of the paper [204] by Sullivan and Thurston says that there exists a universal constant $a > 0$ such that any holomorphic motion of any subset X of $\mathbb{C}$ parameterized by the unit disk {t: $|t| < 1$} can be extended to a holomorphic motion of the complex plane with time parameter the disc {t: $|t| < a$}. With his logician bias, Thurston did not fail to notice a close relation between the problem of extending holomorphic motions and a "holomorphic axiom of choice". This result was later improved by Slodowski [198], who showed in particular that one may take $a = 1$ in the above statement, answering a question raised by Sullivan and Thurston in their paper, and proving a conjecture they formulated, precisely, related to the holomorphic axiom of choice.

In the same paper, Thurston and Sullivan introduced the notion of quasiconformal motion. They noticed that the map f_t in the above definition is necessarily quasiconformal and extends to a quasiconformal map of $\mathbb{C}$. They also proved a general extension theorem for quasiconformal motions over an interval. In proving their results, Sullivan and Thurston introduced an averaging procedure for pairs of probability measures defined on the Riemann sphere.

It was known, since the work of Grötzsch, Lavrentieff and Teichmüller that quasiconformal mappings are useful in the study of conformal mappings. From the work of Sullivan and Thurston, the notion of quasiconformal motion became useful in the study of holomorphic motion.

There are applications of holomorphic motion to the theory of Kleinian groups, where the subset X in the above definition of holomorphic motion is the limit set of the group action. There are also applications in the theory of iterations of rational maps (where X is taken to be the Julia set of the map), in the theory of invariant metrics in complex geometry, in the study of holomorphic families of Riemann surfaces, in the theory of quasiconformal mappings and in the study of Teichmüller spaces.

1.2.11 Complex Projective Geometry

Complex projective geometry is a classical topic, rooted in the nineteenth-century work of Klein, Poincaré and their contemporaries. One must add that in the 1960s, Bers and his collaborators were thoroughly involved in the relation between the complex projective geometry of surfaces and Teichmüller theory. In particular, the Bers embedding of Teichmüller space is defined in the setting of complex projective structures.

In the late 1970s, Thurston reconsidered this theory. He did not publish any paper on this topic, but many authors wrote on it, following Thurston's outline. He highlighted a profound analogy between the complex projective geometry of surfaces and the theory of Kleinian groups, opening a new perspective in 3-manifold topology, and motivating important later works by Sullivan, Epstein, Marden and others.

Thurston introduced metrics that are conformal to the complex structures on complex projective surfaces. In the simplest case, such a metric is obtained by grafting a Euclidean annulus on a hyperbolic surface, after cutting it along a simple closed geodesic. (In fact, it is the projective structure underlying the annulus, and not its Euclidean structure, which matters.) The general definition needs Thurston's theory of projective laminations, and the metric is obtained as a limit of metrics when a sequence of simple closed curves along which the grafting is made converges, in Thurston's topology, to a measured lamination. This metric is now called Thurston's metric on the complex projective surface. Thurston also gave a description of this metric as a Finsler metric, in fact, as a solution of an extremal problem, in the spirit of the Kobayashi metric on a complex space: the distance between two points is the infimum of the length of piecewise C^1 paths joining them. Here, the length of a C^1 path is computed as the integral of the norms of vectors tangent to this path, the norm of a vector v being the infimum of the norms of vectors v' in the tangent space of the unit disc equipped with the hyperbolic structure, such that there exists a projective map from the unit disc to the surface whose differential sends v' to v.

Before Thurston came into the subject, grafting, in its simplest form, was studied by Bers [26], Maskit [139] and Hejhal [102]. There are also relations with the theory of harmonic maps between surfaces, in particular, between the extremal length of a grafted surface and the energy of a harmonic map, see [205]. The grafting operation makes connections between projective structures and hyperbolic geometry. Such connections were already known to Cayley, Klein, Study and others, and traces of the elementary grafting operation can be found in the work of Klein.

Using the notion generalized grafting, Thurston discovered a geometric parametrization of the moduli space of marked projective structures on a surface, establishing a precise relationship between this moduli space and the Teichmüller space of the surface. The parametrization takes the form of a homeomorphism between the moduli space of projective structures and the product of measured lamination space with Teichmüller space. The result says that the fiber at each point of the natural forgetful map from the moduli space of projective structures to Teichmüller space is naturally identified with the space of measured laminations on the surface.

There are several surveys of Thurston's work on complex projective geometry, and they give complementary points of view on the subject. The paper [111] by Kamishima and Tan is concerned with grafting and Thurston's parametrization in the setting of the theory of geometric structures and locally homogeneous spaces. Goldman, in his paper [87], sets the foundations of the theory of complex projective structures on surfaces as geometric structures, using the notions of holonomy and developing map, in the tradition of Thurston, and with ideas originating in the work of Ehresmann [72]. He refers, for Thurston's parametrization of the moduli space of projective structures by the product of measured lamination space with Teichmüller space, to a course given by Thurston at Princeton University in 1976–1977 (Goldman was an undergraduate student there). We also refer the reader to the paper [70] by Dumas, and to his survey [71], for the parametrization of the moduli space of complex projective structures on surfaces. Chapter 6 in the present volume, by Shinpei Baba [15], is a survey of Thurston's work on complex projective structures on surfaces, and it contains other references.

We also mention the paper [203] by Sullivan and Thurston, in which these authors provide a series of examples that show the subtleties of higher-dimensional real and complex projective structures, together with other kinds of geometric structures (inversive and affine).

Now we wish to talk about a classical parametrization of the moduli space of complex projective structures on surfaces based on the Schwarzian derivative, a differential operator invariant under Möbius transformations. This parametrization first appeared in the nineteenth century in the works of Schwarz, Klein and others. Thurston used it in his study of projective geometry, and we briefly discuss this.

In his paper [225] published in a special volume of the AMS at the occasion of the proof of the Bieberbach conjecture, Thurston introduced a topology on the set of univalent mappings from the unit disk into the Riemann sphere using the topology of uniform convergence of Schwarzian derivatives. The uniformity refers to the hyperbolic metric of the disk. To see that this is a natural object of study, one may recall that the Schwarzian derivative was classically used to obtain Riemann mappings of some special domains of the complex plane: regular polygons, domains with circular edges, etc.; generalizing, this makes the set of Schwarzian derivatives is a candidate for a natural parameter space for projective structures.

The usual definition of the Schwarzian derivative, involving third-order complex derivatives, makes it at first sight quite obscure. It is interesting to see how this object was described by Thurston in his paper. He writes: “For the benefit of people to whom the Schwarzian derivative may seem a mystery, we will set the stage by discussing the Schwarzian derivative.” He continues a few lines below: “The Schwarzian derivative is very much like a kind of curvature: the various kinds of curvature in differential geometry measure deviation of curves or manifolds from being flat, while the Schwarzian derivative measures the deviation of a conformal map from being a Möbius transformation.” Then, after a page of explanations, he writes: “A formula for the Schwarzian derivative can be readily determined from the information above, or it may be looked up—someplace else. Like the formula for the curvature of a curve in the plane, the formula looks somewhat mystical at first, and in a quantitative discussion the formula tends to be a distraction from the real issue.” Reading Thurston’s paper gives a clear intuition of what the Schwarzian derivative is.

Responding to a question by Paul Siegel on MathOverflow, on September 9, 2010, asking: “Is there an underlying explanation for the magical powers of the Schwarzian derivative?”, Thurston writes, on the next day: “Like many people (but not all people), I have trouble thinking in terms of formulas such as that for the Schwarzian. For me, a geometric image works much better [...]”. He then gives a nice description based on hyperbolic geometry and quadratic differentials and measured foliations. On the next day (September 11, 2010), Thurston writes to Siegel who thanked him for his response: “I appreciated the question, which resonated with my thoughts. I’m new to MO, but it seems like a rich environment. I understand MO is not intended for extended threads, but I’d like to leave a pointer forward to my first question, which I posted partly as a followup to this, since it indicates the immediate source for my interest in Schwarzians.” In this follow up, Thurston asks several questions, including the following:

> Given a set of $2d-2$ points on $\mathbb{C}P^1$ to be critical points [in the domain], it has been known since Schubert that there are Catalan rational functions with those critical points.
>
> – Is there a conceptual way to describe and identify them?
> – Is there a complete characterization of the Schwarzian derivative for a rational map, starting with the generic case of $2d-2$ distinct critical points?
> – What planar graphs occur for Schwarzian derivatives of rational functions? What convex (or other) inequalities do they satisfy?

1.2.12 Circle Packings and Discrete Conformal Geometry

The study of circle packings, that is, configurations of circles that are tangent to each other, is classical and can be traced back to the work of Apollonius of Perga (third century B.C.), see [14].[9] In the nineteenth century, Paul Koebe proved the existence of some circle packings, and considered the idea of using them to prove the Riemann Mapping Theorem [120]. For a review of the work of Koebe, we refer the reader to the chapter by Philip Bowers in the present volume [40]. One may recall in passing that Koebe (and independently Poincaré) proved a wide generalization of the Riemann Mapping Theorem, namely, the uniformization theorem.

Thurston's work on circle packings inaugurated the notion of discrete Riemann mapping theorem, and more generally, the study of discrete conformal mappings. At the same time, it shed a new light on several geometric ideas that are rooted in classical mathematics.

Let us first recall that the (classical) Riemann mapping theorem, proved by Riemann in his doctoral dissertation [190], says that an arbitrary simply connected open subset of the complex plane which is not the whole plane is conformally equivalent to the unit disc by a mapping which is unique up to composition by a Möbius transformation.

Thurston conjectured the existence of a discrete version of the Riemann mapping theorem as a limit of a sequence of circle packings. The intuition behind this is that a conformal mapping between two open subsets of the plane is characterized by the fact that it sends infinitesimal circles to infinitesimal circles (recall that at the level of tangent planes, it sends circles centered at the origin to circles centered at the origin). Therefore one might hope that finding circle packings with smaller and smaller radii on a given domain and a sequence of homeomorphisms that send them to circle packings of the unit disc leads, by taking limits, to a Riemann mapping. A "discrete Riemann mapping" is one of these mappings used in the approximating sequence.

Thurston's conjecture was proved by Rodin and Sullivan in their paper *The convergence of circle packings to the Riemann mapping* [194]. In their introduction, the authors recall the setting:

> In his address,[10] *The finite Riemann Mapping Theorem*, Bill Thurston discussed his elementary approach to Andreev's theorem and gave a provocative, constructive, geometric approach to the Riemann mapping theorem. This method is quite beautiful and easy to implement on a computer.

[9]For Apollonius' works, the main reference is Rashed's critical edition of the Arabic manuscripts (many Greek texts do not survive), published by de Gruyter in 5 volumes (more than 2500 pages) between 2008 and 2010. Apollonius' problems are discussed in the volume [14].

[10]International symposium in Celebration of the Proof of the Bieberbach Conjecture. Purdue University, March 1985.

They then recall Thurston's strategy of the proof:

> Almost fill a simply connected region R with small circles from a regular hexagonal circle packing. Surround these circles by a Jordan curve. Use Andreev's theorem to produce a combinatorially equivalent packing of the unit disc—the unit circle corresponding to the Jordan curve. The correspondence between the circles of the two packings ought to approximate the Riemann mapping.

Following Thurston's ideas, Rodin and Sullivan develop in an appendix to their paper an algorithm to obtain the discrete Riemann mapping.

In his Princeton lectures, Thurston studied circle packings in the midst of a discussion of orbifolds and of an existence theorem for hyperbolic polyhedra. When he started lecturing on the subject, he was aware neither of Koebe's nor of Andreev's work; see the interesting historical remarks in Bowers' review [39]. He realized at some point that some of the results he obtained were generalizations of results contained in two papers by Andreev [12, 13]. He then called the existence theorem for circle packings that is contained in Chapter 13 of his Princeton notes [215] Andreev's Theorem. The result is now called the Koebe–Andreev–Thurston theorem. This theorem states that for any triangulation of a closed surface (of any genus) which lifts to a simple triangulation of the universal cover (that is, a triangulation which has no pair of edges connecting the same vertices, and no edge connecting the same vertex), there exists a unique metric of constant curvature on the surface with a circle packing that is modeled on it. Furthermore, the packing is unique up to a conformal map isotopic to the identity (which implies, in the hyperbolic case, that the map is the identity). Thurston deduced the uniqueness result from Mostow's rigidity theorem. In his notes, he considered in detail the genus 1 and ≥ 2 cases. The genus 0 case was treated by Andreev and was attributed to him by Thurston. Marden and Rodin wrote a paper showing that Thurston's method also gives a proof for the genus 0 case [135].

Thurston also proved an existence theorem for patterns of circles that generalizes a result of Koebe, allowing an overlap among the pattern of circles (and not only tangency), and he used this result in his proof of the generalized Andreev theorem. In §13.4 of his notes [215], he gave algorithms for constructing circle packings and more general circle patterns. His algorithms allow computations.

For an overview of Thurston's discrete Riemann mapping theorem and its impact, we refer the reader to the comprehensive survey by Bowers in the present volume [40]. We also refer to Luo's paper [131] and to Kojima's survey [121].

1.2.13 Word Processing in Groups, Automata and Tilings

Besides the name of Thurston, two names will be highlighted in this section: Jim Cannon and John Conway.

We start with groups and automata, to which the name of Cannon is attached.

In 1984, Cannon published a paper in which he showed that Cayley graphs of cocompact discrete groups of isometries of n-dimensional hyperbolic space can be given finite recursive descriptions [53]. He wrote in the introduction that he was inspired by Thurston, who showed that a large number of groups that are of interest to topologists cannot be dealt with using the standard methods of combinatorial group theory, but can be attacked by "a return to geometric consideration", that is, the classical methods of Dehn and Cayley. It is significant that Cannon's paper contains an appendix on elementary properties of hyperbolic space, for which, at that time, no modern exposition was available, except for Thurston's unpublished notes [215].

Thurston noticed that Cannon's results can be formulated in the language of finite state automata, and may be applied to a wider class of groups. This led him to the introduction of the notion of automatic group. This is a group equipped with a simple algorithm for the word problem, that is, an automaton can tell when two words (in a given system of generators) represent the same element in this group.

After their discovery by Thurston, automatic groups found applications in a wide class of domains including low-dimensional topology and geometry, geometric and combinatorial group theory, algorithmics, decision theory, computer vision, mathematical logic, etc. Furthermore, the theory of automatic groups is closely related to that of finite state automata, which has applications in computer science. Thurston was interested in all these applications. He developed with his collaborators computer programs to carry out what he called "word processing on groups." There is also a relation with self-similar tilings. Thurston writes in [230]: "An automatic structure for a group in general produces a kind of self-similar tiling of a certain 'sphere at infinity' for the group; in particular examples, this space is actually a 2-sphere."

Soon after their discovery, automatic groups became a central part of geometric group theory. Thurston collaborated with Cannon and others on this theory in connection to his eight geometries of 3-manifolds. In the paper [55] written with Cannon, Floyd end Grayson, he showed that no cocompact discrete group based on solvgometry, **Sol**, is almost convex. As a consequence, the Cayley graph of such a group cannot be efficiently constructed by finitely local replacement rules. After recalling Thurston's geometrization conjecture, the authors write that "any package of decision algorithms designed to compute within the fundamental groups of low-dimensional manifolds and orbifolds must be able to deal with the groups from each of the standard geometries." The theory of automatic groups is developed in the comprehensive textbook [75] that Thurston wrote with Epstein, Cannon, Holt, Levy and Paterson, which appeared several years after he started working on this topic.

We now pass to tilings.

The study of tilings is closely related to discrete group actions, a theory that plays an essential role in Thurston's work on 3-manifolds. Thurston was fascinated not only by tilings in dimension 3, but also by the theory of plane tilings, in particular by the theory of quasiperiodic tilings. These are tilings where finite patters appear regularly, without being necessarily periodic. Let us quote an excerpt from a set of questions that Thurston distributed at the beginning of his course on "Geometric topology" at Princeton, during the Spring Semester of 1983 [217, Question 9]:

> Is there a general mathematical theory for Penrose-like tilings, where one specifies certain combinatorial relationship and then deduces that certain shapes of tiles exist which satisfy these relations? Are there many essentially different such tilings, or just few?

Thurston was stimulated on this subject by ideas of Conway, who was working at the same university and who made major contributions to group theory, sphere packings, tilings and cellular automata. The latter, together with Jeffrey Lagarias, developed a method, based on combinatorial group theory, to tackle the problem of tiling some finite region of the plane using a certain number of regular tiles [58]. This method involved the encoding of the edges of the tiles by elements of a finitely-presented group in such a way that a tile can be interpreted as a relator in the group. The problem was then reduced to that of deciding whether some group element, describing the boundary of the plane region, is the trivial element.

In 1990, Thurston published a paper in the American Mathematical Monthly [223] in which he re-interpreted Conway's construction using the tools of geometric group theory. In the same paper, he gave a necessary and sufficient condition for a simply-connected region of the plane which is the union of unit squares, to have a tiling by dominos, that is, rectangles which are the union of two squares. He also gave several constructions of tessellations of planar regions by given tiles.

In the same year, Thurston and Conway, together with Peter Doyle, started a new course at Princeton, called "Geometry and the imagination." Thurston writes [222]: "The course came alive, qualitatively more than any course we had taught before. Students learned a lot of mathematics and solved problems we wouldn't have dared ask in a conventional college class."

Thurston's collaboration with Conway includes the paper [59] by Conway, Delgado Friedrichs, Huson and Thurston in which these authors gave a new enumeration of n-dimensional crystallographic groups, that is, cocompact discrete subgroup of the isometry group of Euclidean 3-space. The enumeration is based on a description of these groups as fibrations over the plane crystallographic groups, when the enumeration is possible. The "old" enumeration, due to Barlow, Fedorov, and Schönflies, dates back to the 1890s.

We mention another paper on tilings (although this word is used in a slightly different meaning), namely, the paper [63], by Coven, Geller, Silberger and Thurston, concerned with the symbolic dynamics of tiling the integers. Here, a finite collection of finite sets of integers is said to "tile the integers" if the set of all integers can be written as a disjoint union of translates of elements of this finite set. These elements are called tiles. To such a set of tiles, the authors associate a bi-infinite sequence of elements of tiles. They show that the set of all such sequences is a sofic system, and that every shift of finite type can be realized (up to a power) as a tiling system.

The paper [230] contains results on self-similar tilings, in particular, constructions of such tilings from algebraic integers λ whose Galois conjugates, except λ and $\overline{\lambda}$, are smaller. More generally, Thurston introduced the notion of complex expansion constants for self-similar tilings, and he gave a characterization of these constants. He obtained a characterization of the set of similarities for self-similar tilings of the plane or of higher-dimensional spaces, making an analogy with the construction of Markov partitions from classical dynamical systems. Beyond the results he obtained, Thurston emphasized the aesthetical side of the topic. He writes: "What is interesting about this subject is the particular constructions—at issue is how simple and how nice can a self-similar tiling can be."

Automata and tilings were part of the subject of a series of lectures which Thurston gave at a summer AMS colloquium in the summer of 1989. The title of the series was *Groups, tilings and finite state automata*. A preprint carrying the same title [230] was distributed at the meeting, and it was later included in the Research Reports of the Geometry Center preprint of the University of Minnesota, a center co-founded by Thurston. The paper, which may be considered as semi-expository, remained in a preprint form. In this domain, and like many other topics which he considered, Thurston had a huge amount of ideas bubbling in his brain, and it was certainly difficult for him to sort out what was new and what was known in some sort or another.

1.2.14 Computers

In the preceding section, we were led to mentioning computers quite a few times. We give here a quick overview on other works of Thurston on this subject, and of his collaboration with computer scientists. We highlight the fact that Thurston's collaboration with computer scientists was twofold. On the one hand, he used methods of geometry, in particular 3-dimensional hyperbolic geometry, to solve problems in computer science, and on the other hand, motivated by questions that arose from computer science, he developed new topics and opened up new ways of research in geometry.

Thurston was heavily involved in computing and computer graphics since the 1980s. Let us quote a question from his list addressed to his students that we already mentioned [217, Question 18]:

> What is the information content of text? How well can one model the sequence of letters in a novel as a dynamical system? That is, suppose you forget that you know anything about language and meaning, and just try to analyse it from a statistical point of view; how could you do in automatically guessing what the next letter would be?
>
> This relates to the question of how much space it takes to store such a string of text in a computer. Given a model process, one could make up a coding scheme. In one direction, it would be possible to feed in a random set of bits and have the code produce a more-or-less plausible stream of text (depending on the complexity of the process which one allows); and in the other direction, one would feed in a text and have it compressed into a much shorter stream of symbols. One would try not to be prejudiced too much by the meaning of the words, but still use knowledge of English (or whatever language) to figure out a good reasonably small set of data which are useful in predicting what next occurs.
>
> Similar questions can be asked about many other human-generated processes (e.g., music), many of them with obvious applications (e.g., the stock market, sequences of answers to multiple choice texts, …). How much entropy do these processes have? Are there families of dynamical systems which do well the modeling?

Thurston collaborated with computer scientists on geometric problems he formulated, but also on problems that were asked by computer scientists themselves. We mention first his paper with Sleator and Tarjan, *Rotation distance, triangulations, and hyperbolic geometry* [197], published in 1988, in which a distance, called *rotation distance*, is defined on the set of binary trees, as the minimal number of rotations that may be used to convert one of these trees into the other. The term "rotation" denotes here the operation of collapsing an internal edge of a binary tree to a point and expanding the node, obtaining a new binary tree. The authors show in this paper that for binary trees with n nodes with $n \geq 11$, the maximum rotation distance is at most $2n - 6$. The motivation for this problem comes from a problem used in data structuring and network algorithms, and more precisely, from a conjecture called the *splaying conjecture*. The authors explain this conjecture as follows: "Splaying is a heuristic for modifying the structure of a binary search tree in such a way that repeatedly accessing and updating the information in the tree is efficient." The methods used in this paper are based on hyperbolic geometry, in the pure Thurston tradition. The rotation operation between binary trees is converted to an equivalent operation of flipping a diagonal in a polygon then passing one dimension higher which permits the rotation distance problem to be reduced to a 3-dimensional problem of dissecting hyperbolic polyhedra into tetrahedra. The volume of hyperbolic polyhedra appears in various forms as a fundamental object in this study. The last section of the paper contains open questions, asking in particular for more calculations of triangulations and volumes for polyhedra. A relation with the Gromov norm, as a measure of how efficiently a homology class in a hyperbolic manifold can be represented by simplices (Chapter 6 of Thurston's Princeton notes [215]) is also mentioned in this section on open problems.

There are other papers of Thurston on computer science and algorithmic problems in which Thurston's geometrical methods are used as en essential tool. We mention the three papers in collaboration with Gary Miller, Shang-Hua Teng and Stephen Vavasis [146], *Automatic mesh partitioning* [145], *Separators for sphere-packings and nearest neighborhood graphs* and *Geometric separators for finite-element meshes* [147], and his paper with Bob Riley, *The absence of efficient dual pairs of spanning trees in planar graphs* [192].

Finally, we mention Thurston's paper *Shapes of polyhedra and triangulations of the sphere* [227] motivated by the question of classifying the combinatorial classes of triangulations of the sphere with at most 6 triangles at a vertex, in which he was led to endow the moduli space of polyhedra with n vertices with given total angles less than 2π at each vertex (that is, Euclidean cone metrics of nonnegative curvature) with a Kähler metric which is locally isometric to complex hyperbolic space $\mathbb{CH}^{n-3}$. This paper had an enormous influence and several generalizations of the results were attempted by many authors.

Regarding his collaboration with computer scientists, Thurston writes his 1987 *Notices* article [220]:

> Recently, through circumstances, I have spent time with computer scientists. I find myself talking and thinking about computer science problems, and analyzing them with modes of thought sometimes foreign to the culture of computer science. I enjoy this. My experience

> would be similar if I were to spend time with physicists, biologists, economists, chemists, engineers ....

One should emphasize the fact that Thurston since the 1970s has been constantly programing, computing, implementing lists of knots, of 3-manifolds, of volumes of hyperbolic manifolds, of tilings, etc.

1.2.15 Surfaces, Mapping Class Groups and Teichmüller Spaces

In 1975–1976, Thurston gave a course at Princeton on the geometry and dynamics of homeomorphisms of surfaces. He presented there a complete theory which had to have a major and everlasting impact on low-dimensional topology and geometry. A major role in this theory was given to Teichmüller space, the space of isotopy classes of metrics of constant curvature -1 on a surface. Thurston's results included a compactification of this space by the space of projective classes of measured foliations, the latter seen as a completion of the set of homotopy classes of simple closed curves on the surface. The results also included a natural action of the mapping class group on this compactified Teichmüller space, and the classification of mapping classes into periodic, reducible and pseudo-Anosov, obtained by analyzing the fixed point set of the action of a mapping class on this compactified space.

Copies of a set of notes on Thurston's course, taken by Bill Floyd and Mike Handel, were circulated, and in particular they arrived to Orsay where they gave rise to the famous seminar *Travaux de Thurston sur les surfaces* which took place during the year 1976–1977; see the paper [126] for the history of this seminar. It appears that Thurston was already thinking about surfaces, and in particular how simple closed curves approach a foliation, at the time he was a PhD student in Berkeley, see Sullivan's account in [126].

A couple of years after the Orsay seminar, the book [77] was written and became the major reference on Thurston's theory on surfaces. In the meanwhile, Thurston wrote a research announcement of his results, which he submitted without success to a few journals. The research announcement eventually appeared in the Bulletin of the AMS, in 1988 [221], 12 years after Thurston wrote it. The paper contains new bibliographical references and a new preface in which Thurston gives a few notes on the history of the manuscript and of the subject.

Shortly after Thurston obtained his classification theorem for mapping classes, Lipman Bers gave a new proof of that theorem in a complex analysis setting, and using the techniques of quasiconformal mappings [27]. Bers's proof also uses the action of the mapping class group on Teichmüller space, but unlike Thurston's proof, it is based on an analysis of the translation length of an element of the mapping class group with respect to Teichmüller's metric. In fact, in Bers' classification, there are four sorts of mapping classes, according to whether the translation length is zero or positive, and in each case, according to whether this translation length (which is defined as an infimum) is attained or not by a point in Teichmüller space.

The book *Travaux de Thurston sur les surfaces* did not include Thurston's theory of geodesic laminations and train tracks, which turned out to be very efficient tools in low-dimensional topology. These notions were expanded on in the courses that Thurston gave the following years at Princeton, and they are included in his notes [215]. Several books appeared on the subject, see e.g. the notes by Casson and Bleiler that arose from a course that Casson gave on Thurston's theory of surfaces at the University of Texas at Austin, [56] and the book [181] by Penner and Harer on train tracks.

One consequence of Thurston's work was the revival of nineteenth-century hyperbolic geometry, a subject which was almost forgotten. Thurston's notes [215], together with the books [77] and [56], were for some time the main references on this topic. (In particular, [77] contains all the hyperbolic trigonometry formulae that are useful in the theory). At the time Thurston started working on the subject, there were practically no modern treatments of the subject. Of course, one could look into Lobachevsky's works, but this was very unlikely. The textbooks by Beardon, Ratcliffe, Anderson and others appeared several years later. The classical books *Elementary geometry in hyperbolic space* [78] by Fenchel and *Discontinuous groups of isometries in the hyperbolic plane* [79] by Fenchel and Nielsen, which existed in the form of lecture notes and had trouble in being published, appeared in 1989 and 2003, after Thurston's work made them famous. The so-called Fenchel–Nielsen parameters for hyperbolic surfaces, associated with geodesic pairs of pants decompositions, acquired all their strength after Thurston used them in his work. Works of Abikoff [1], Wolpert [240] and others on this deformation appeared after Thurston revived the subject.

Pseudo-Anosov homeomorphisms, which appeared in Thurston's classification, turned out to be a major ingredient in the geometry and topology of 3-manifolds; we mention for instance their role in the construction of hyperbolic manifolds which fiber over the circle, explained in Sect. 1.2.4.

Before continuing on Thurston's theory of surfaces, we make a small digression concerning Nielsen's contribution to this subject.

Jakob Nielsen, in several long papers published between 1927 and 1944 [168–171], studied questions related to automorphisms of surfaces, using hyperbolic geometry. Thurston writes in the introduction of his paper [221]:

> At the time I originally discovered the classification of diffeomorphism of surfaces, I was unfamiliar with two bodies of mathematics which were quite relevant: first, Riemann surfaces, quasiconformal maps and Teichmüller's theory; and second, Nielsen's theory of the dynamical behavior of surface at infinity, and his near-understanding of geodesic laminations.

In the same preface, Thurston writes: "Dennis Sullivan first told me of some neglected articles by Nielsen which might be relevant." In a paper he wrote with M. Handel, titled *New proofs of some results of Nielsen* [99], Thurston gave a new proof of his classification theorem using techniques from Nielsen's program. The relationship between the works of Thurston and Nielsen is also examined in the papers by Jane Gilman [86], Richard Miller [144] and Joan Birman [28].

Talking about Nielsen, we are led to Nielsen's realization problem and the use of earthquakes in its solution.

Earthquakes generalize the Fenchel–Nielsen deformation operation of cutting a hyperbolic surface along a simple closed geodesic and gluing it again after a twist. They are limits of sequences of such operations when the sequence of simple closed geodesics converges in Thurston's topology to a measured geodesic lamination. They became the canonical deformations between two hyperbolic structures after Thurston proved that for any two hyperbolic structures on a given surface, there is a unique left earthquake joining them. His proof is contained as an appendix in Steve Kerckhoff's paper [114]. They were the essential ingredient in Kerckhoff's proof of the Nielsen realization problem, which we review below.

Thurston wrote a paper on earthquakes on the hyperbolic plane [234]. In this paper, earthquakes are more naturally defined by cutting the hyperbolic plane along geodesics, taking limits of such operations, and studying the action on the unit circle at infinity. Considering this action on the universal covering and on the circle at infinity solves the problems caused by the discontinuities of the map. At the same time, Thurston placed his theory in the setting of the universal Teichmüller space, the natural setting for deformations of the hyperbolic disc. He obtained a new and elementary proof of the earthquake theorem. He described this fact by the expression "geology is transitive." In a set of notes he distributed in October 1987 on this new proof of the earthquake theorem, while he constructs the earthquake map using a homeomorphism of the circle at infinity of the hyperbolic plane, he writes:

> This is closely connected to basic properties of convex hulls of sets in 3-space. Intuitively, imagine having disks with rubber arrows representing the identifications. Imagine some physical device which forces all the arrows to point counterclockwise: they bump against some barrier if you try to rotate them too far. You are allowed to move one of the disks by any isometry of the hyperbolic plane. You can kind of roll the disk around on the barriers through many different positions. This is very much like rocking a plane around on top of a wire which projects to a circle on a table. In the latter case, pushing straight down above one point finds the flat of the convex hull lying above a point inside the circle; in the former case, twisting at one point finds the stratum of the earthquake.
>
> The earthquake theorem can be proven by formalizing this argument.

One of the first applications of Thurston's earthquake theorem was the proof of the Nielsen realization problem in 1980 by Kerckhoff [114, 115]. The problem, formulated by Nielsen in 1932, asked whether any finite subgroup of the mapping class group of a surface can be realized as a group of homeomorphisms of this surface. In 1942, Nielsen gave an affirmative answer in the case of finite cyclic groups.[11] Fenchel extended Nielsen's result to the case of finite solvable groups.

[11] Thurston and Handel note in their paper [99] that there should be a gap in Nielsen's proof of the fact that if a mapping class is periodic, then it contains a periodic homeomorphism of the surface. For the fact that Nielsen's proof is incorrect, they refer to Zieschang [242], and they declare that the known proofs of this fact use more sophisticated methods than those of Nielsen, e.g. actions

There were several failed attempts to solve the Nielsen realization problem, namely, by Kravetz in 1959, based on the false assumption that the Teichmüller metric is negatively curved. But this failed proof had the advantage of putting the action of mapping classes on Teichmüller space at the center of the discussion. Kerckhoff's proof is based on a convexity argument and a result of Thurston saying that any two points in Teichmüller space can be joined by a left earthquake. In Sect. 1.3.7 of this chapter, we shall talk about the work of Geoffrey Mess in the early 1990s, who established a profound relation between earthquakes and the geometry of the convex core in anti-de Sitter manifolds. Besides the realization of finite subgroups, it was natural to address the same question for arbitrary subgroups. Thurston asked the question of the lift of the whole group (see Problem 2.6 in Kirby's list [118]), and he conjectured that the answer is no. The conjecture was proved by Markovic, for closed surfaces of genus ≥ 5, in his paper [136], after Morita [165] obtained the same result for diffeomorphisms, using a more algebraic approach (cohomological obstructions). A year later, Markovic and Šarić completed the proof of Thurston's conjecture for the cases of genus 2 to 4 [137].

Talking about Thurston's work on Teichmüller space, we mention now his approach to the Weil–Petersson metric.

Thurston introduced a Riemannian metric on Teichmüller space where the scalar product of two tangent vectors at a point represented by a hyperbolic surface is the second derivative of the length of a uniformly distributed sequence of closed geodesics on the surface. Thurston was motivated by the wish to have a metric defined using only the hyperbolic geometry of the surface, in contrast to the Teichmüller metric, which is based on quasiconformal theory, and to the Weil–Petersson metric, whose definition used the Petersson Hermitian product, defined in the context of modular forms and used by number theorists.

Wolpert showed that Thurston's metric coincides with the Weil–Petersson metric [241]. The consequence is that Thurston gave a purely hyperbolic-geometric interpretation of the Weil–Petersson Riemannian metric on Teichmüller space.

We now pass to Thurston's asymmetric metric.

In 1985–1986, Thurston circulated a preprint titled *Minimal stretch maps between hyperbolic surfaces* [229] in which he introduced a non-symmetric metric on Teichmüller space which now bears the name *Thurston metric*. The distance between two hyperbolic structures on a given surface is taken to be the logarithm of the infimum of the Lipschitz constants of homeomorphisms that are homotopic to the identity, where the distances used for the computation of the Lipschitz constant are the one of the first metric in the domain and the second metric in the range. Thurston's paper is based on first principles (there is no appeal to any theory or any theorem except basic hyperbolic geometry). The paper was submitted to the journal *Topology*. The referee sent a long report asking for clarifications and references, and

on Teichmüller space or Smith theory. Thus, they consider Nielsen's proof of the cyclic case as incomplete.

Thurston withdrew the paper. In 1998, Thurston posted the article on the arXiv, and it remained unpublished.

Thurston's motivation was to develop a theory of Teichmüller space which is purely geometric, and which, like in his approach to the Weil–Petersson metric, does not rely on quasiconformal mappings and quadratic differentials, but only on elementary hyperbolic geometry. He described a class of distinguished geodesics for this metric which he called stretch lines, he showed that any pair of points in the Teichmüller space can be joined by a concatenation of such lines, he showed that the metric is Finsler, and he described the dual unit sphere of the associated norm at each point of the cotangent space as an embedding in this space of projective lamination space. In the same paper, Thurston introduced his *shear coordinates* for a surface decomposed into ideal triangles, coordinates which have had an enormous impact, in the so-called higher Teichmüller theory and elsewhere, see e.g. [123]; see also the generalization of these coordinates to the context of decorated Teichmüller theory [180]. Thurston's paper [229] also contains the definition of an asymmetric norm on the tangent space to Teichmüller space which he called the *earthquake norm*, which leads to another asymmetric metric on Teichmüller space. This norm is considered in the chapter by Barbot and Fillastre in the present volume [20].

In the first years after the preprint was released, little progress was made on this topic, one reason being that it took some time for the geometers to understand the ideas and the proofs contained in it. A survey of the results obtained in the first 20 years after Thurston's preprint was circulated, appeared in 2007, see [177]. A set of open problems on Thurston's metric appeared in 2015 [200], after a conference held on this topic at the American Institute American Institute for mathematics in Palo Alto. We also mention the recent survey [176].

Thurston's metric led to the definition and study of analogous metrics in other settings: surfaces with boundary [129], Euclidean structures on surfaces, [23], the Culler–Vogtmann outer space [82], geometrically finite representations of fundamental groups of surfaces in higher-dimensional Lie groups [94], and there are many others developments, see e.g. [8, 105].

The next section, on fashion design, could have been included in the present one; it is also about Thurston's work on surfaces.

1.2.16 Fashion Design

Let us start this section by mentioning a paper by Thurston and Kelly Delp titled *Playing with surfaces: Spheres, monkey pants, and zippergons* [69], one of the last papers of Thurston, written in 2011. In this paper, the authors describe a process, inspired by clothing design, of smoothing an octahedron to form a round sphere. They mention in the introduction several workshops and series of encounters they organized on devising schemes for designing pattern pieces to fit arbitrary shapes, including the human body. They declare: "It was a very interesting but humbling experience, because our initial assumption that familiar theoretical principles of differential geometry would do most of the work was misleading."

Thurston was interested in the geometrical theory of clothes and the fitting of garments since his early work on surfaces. In the introduction to his paper [229], he writes, after giving the definition of the best Lipschitz constant of maps in a given homotopy class:

> This is closely related to the canonical problem that arises when a person on the standard American diet digs into his or her wardrobe of a few years earlier. The difference is that in the wardrobe problem, one does not really care to know the value of the best Lipschitz constant—one is mainly concerned that the Lipschitz constant not be significantly greater than 1. We shall see that, just as cloth which is stretched tight develops stress wrinkles, the least Lipschitz constant for a homeomorphism between two surfaces is dictated by a certain geodesic lamination which is maximally stretched.

Thurston was not the first mathematician to think of the mathematical question of fitting a piece of fabric to some surface. Pafnuty Chebyshev, back in the nineteenth century, thoroughly investigated the problem of fitting garments to a part of the human body. In particular, he thought about questions regarding the flexibility for a piece of fabric in order for it to approximate in the best exact form the part of body to which it is designated and he established relations between this problem and other mechanical problems he was studying, including the theory of linkages. He wrote an article on this topic, [57]. A review of Chebyshev's work, making relations with his work and Euler's on geography and other problems related to surfaces, is contained in the paper [175].

Thurston became eventually involved in fashion design. He worked with Dai Fujiwara, the creative director of the Japanese fashion designer Issey Miyake, creating in 2010 a beautiful collection inspired by his eight geometries. On the occasion of the fashion show that took place at the Salon du Carrousel du Louvre in Paris, in March 2010, in which this collection was exhibited, Thurston wrote a brief essay, distributed during the show, on beauty, mathematics and creativity. Here is an excerpt:

> Many people think of mathematics as austere and self-contained. To the contrary, mathematics is a very rich and very human subject, an art that enables us to see and understand deep interconnections in the world. The best mathematics uses the whole mind, embraces human sensibility, and is not at all limited to the small portion of our brains that calculates and manipulates with symbols. Through pursuing beauty we find truth, and where we find truth, we discover incredible beauty.
>
> The roots of creativity tap deep within to a place we all share, and I was thrilled that Dai Fujiwara recognized the deep commonality underlying his efforts and mine. Despite literally and figuratively training and working on opposite ends of the earth, we had a wonderful exchange of ideas when he visited me at Cornell. I feel both humbled and honored that he has taken up the challenge to create beautiful clothing inspired by the beautiful theory which is dear to my heart.

In another article written on that occasion for the fashion magazine *Idoménée*, Thurston made the following comment about the collection:

> The design team took these drawings as their starting theme and developed from there with their own vision and imagination. Of course it would have been foolish to attempt to literally illustrate the mathematical theory— in this setting, it's neither possible nor desirable. What

> they attempted was to capture the underlying spirit and beauty. All I can say is that it resonated with me.

Fashion design was for Thurston a ground for the combination of mathematics, art and practical applications, where the aesthetic component is pre-dominant. In an interview released on the occasion of that fashion show, in which Thurston recounted how he came to contribute to the collection, he declared: "Mathematics and design are both expressions of human creative spirit." About the aesthetical aspect of mathematics, Thurston had already written, in this 1990 *Notices* article [222]:

> My experience as a mathematician has convinced me that the aesthetic goals and the utilitarian goals for mathematics turn out, in the end, to be quite close. Our aesthetic instincts draw us to mathematics of a certain depth and connectivity. The very depth and beauty of the patterns makes them likely to be manifested, in unexpected ways, in other parts of mathematics, science, and the world.

1.3 On Thurston's Impact

1.3.1 The Proof of the Smith Conjecture

The resolution of the Smith conjecture was the occasion for the first major application of Thurston's uniformization theorem for Haken manifolds. This conjecture says that if a cyclic group acts on S^3 by diffeomorphisms with one-dimensional fixed points, then it is topologically conjugate to the standard orthogonal action. The conjecture can be paraphrased as follows: a branched cyclic cover of a closed 3-manifold M along a knot K can be homeomorphic to the 3-sphere only if M is also homeomorphic to the 3-sphere and K is unknotted.

The proof of the Smith conjecture was published as a book [161], which also contains a very comprehensive survey of Thurston's uniformization theorem written by Morgan. The proof is divided into two cases depending on whether $M \setminus K$ contains a closed incompressible surface or not. When it does, then results of Meeks–Yau and Gordon–Litherland, contained in the book, show that the branched cover must also contain an incompressible surface, and hence cannot be the 3-sphere. When it does not, then by the uniformization theorem, $M \setminus K$ is either a Seifert fibered manifold or hyperbolic. In the former case, it is easy to see that the only possibility is that $M \setminus K$ is homeomorphic to $S^1 \times \mathbb{R}^2$. In the latter case, an algebraic argument due to Bass implies that the holonomy representation of $\pi_1(M \setminus K)$ can be taken to lie in a ring consisting of algebraic integers. Then an argument of commutative algebra shows that the only possibility is that K is unknotted.

After the publication of this book, Culler and Shalen [64] gave an alternative, more geometric approach to the algebraic argument in the last part of the proof. They considered the algebraic set of characters of representations of the fundamental group of a hyperbolic manifold into $SL_2\mathbb{C}$, called the character variety. Their alternative proof is obtained by considering the points at infinity of the character variety of $M \setminus K$, which gives a decomposition of $\pi_1(M \setminus K)$ by way of Bass–Serre theory. This work of Culler–Shalen was generalized to a theory of compactification of character varieties by Morgan–Shalen [162–164], which also gives an alternative proof of Thurston's compactness theorem for deformation spaces of acylindrical manifolds.

1.3.2 The Proofs of Ahlfors' Conjecture, Marden's Tameness Conjecture, the Ending Lamination Conjecture, and the Density Conjecture

The following four conjectures on Kleinian groups are contained in Thurston's list of unsolved problems [216].

1. For any finitely generated Kleinian group, its limit set in the Riemann sphere either has measure 0 or coincides with the entire sphere. This conjecture is originally due to Ahlfors [7], and therefore called Ahlfors' conjecture.
2. Any hyperbolic 3-manifold with finitely generated fundamental group is homeomorphic to the interior of a compact 3-manifold. This appeared first in Marden's paper [134]. The property is called the topological tameness for the hyperbolic 3-manifold and also for the corresponding Kleinian group. The conjecture is called Marden's tameness conjecture.
3. If two hyperbolic 3-manifolds are homeomorphic and have the same parabolic locus, the same conformal structures at infinity, and the same ending laminations, then they are isometric. This is called the ending lamination conjecture.
4. Any finitely generated Kleinian group is an algebraic limit of geometrically finite Kleinian groups. This is called the (Bers–Sullivan–Thurston) density conjecture.

The resolutions of these four conjectures proceeded in an intertwined way. Thurston himself showed that algebraic limits of quasi-Fuchsian groups are geometrically tame. Geometric tameness implies topological tameness, but is a stronger condition. Thurston also showed that for geometrically tame Kleinian groups, Ahlfors' conjecture is true. Bonahon [31] clarified Thurston's notion of geometric tameness, and proved that any finitely generated Kleinian group that is not decomposed into a free product (i.e. any freely indecomposable Kleinian group) is geometrically and topologically tame, implying that Ahlfors' conjecture is also true for such Kleinian groups.

For freely decomposable Kleinian groups, Canary [51] proved that topological tameness implies geometric tameness, and hence that Ahlfors' conjecture holds for topologically tame Kleinian groups. Ohshika [173] showed that any purely loxodromic algebraic limit of geometrically finite Kleinian groups is topologically tame unless the limit set is the entire sphere, and hence that Ahlfors' conjecture holds for any such algebraic limit. Canary–Minsky [52] proved topological tameness for strong limits of topologically tame Kleinian groups. Brock–Bromberg–Evans–Souto [44] proved that every algebraic limit of geometrically finite Kleinian groups is topologically tame. Finally, Agol [5] and Calegari–Gabai [50] resolved Marden's tameness conjecture completely.

The ending lamination conjecture was proved by Minsky [151] for freely indecomposable Kleinian groups having a positive lower bound for the translation lengths (Kleinian groups are then said to have bounded geometry). Ohshika [172] proved that the assumption of free indecomposability can be removed still under the assumption of bounded geometry. Minsky [152] proved the ending lamination conjecture for once-punctured torus Kleinian groups. The general ending lamination conjecture was resolved by Minsky [153] and Brock–Canary–Minsky [45] using the work of Masur–Minsky [140, 141] on the geometry of curve complexes. The proof relies on Thurston's idea of approximating the geometry of a neighborhood of an end by pleated surfaces. The point is that how pleated surfaces tend to the end is governed by a hierarchical structure of the curve complex, which was investigated in [141].

The density conjecture was proved for Kleinian surface groups by Bromberg [46], Brock–Bromberg [43] using Minsky's resolution of the ending lamination conjecture in the bounded geometry case. The general density conjecture was solved by Ohshika [174] and Namazi–Souto [167] relying on the full resolution of the ending lamination conjecture.

1.3.3 The Proof of the Geometrization Conjecture

The geometrization conjecture says that every closed irreducible 3-manifold can be decomposed into geometric pieces by (Jaco–Shalen–Johannson) torus decomposition. Thurston's uniformization theorem says that this is true for Haken manifolds, but of course there are non-Haken manifolds: closed 3-manifolds with finite fundamental groups are clearly non-Haken, but there are also non-Haken manifolds with infinitely fundamental groups.

The geometrization theorem was resolved by Perelman using the Ricci flow. A Ricci flow is a deformation of Riemannian metric in the direction to reduce the variation of its Ricci curvature over the manifold, i.e. to average the Ricci curvature. Hamilton [98] proved that any closed Riemannian 3-manifold with positive Ricci curvature is diffeomorphic to the 3-sphere, making use of Ricci flows. Perelman considered Ricci flows for general closed irreducible 3-manifolds [182–184]. In contrast to Hamilton's case, the flow may encounter singularities. Perelman showed that even in such cases, the deformation can be continued by rescaling and surgeries, and finally get to either constantly curved manifolds or Seifert fibrations. This proves the geometrization conjecture, and in particular the Poincaré conjecture posed by Poincaré in 1904.

1.3.4 *The Waldhausen Conjectures and the Virtual Fibering Conjecture*

In [239], Waldhausen posed the following two long-standing conjectures on 3-manifolds.

(1) The fundamental group of every closed irreducible 3-manifold either is finite or contains a closed surface group.
(2) Every closed irreducible 3-manifold with infinite fundamental group is finitely covered by a Haken manifold. This conjecture is now called the virtual-Haken conjecture.

 The affirmative resolution of the second conjecture implies that of the first.

 In his list of open questions in [216], Thurston took up the second conjecture again and added the following two stronger conjectures.

(3) Every closed irreducible 3-manifold with infinite fundamental group is finitely covered by a 3-manifold with positive first Betti number.
(4) Every closed hyperbolic 3-manifold is finitely covered by a surface bundle over the circle.

Since it is easy to see that any Seifert fibered manifold with infinite fundamental group has a finite cover with positive Betti number, the affirmative resolution of (4) implies that of (3).

Conjecture (1) was known to hold for Seifert fibered manifolds. Therefore we have only to consider hyperbolic 3-manifolds. The conjecture was proved in the case of arithmetic hyperbolic manifolds by Lackenby [125], and was solved in general form by Kahn and Markovic [108]. Conjectures (2) and (4) (and hence also (3)) were solved by Agol, assuming the result of Kahn–Markovic, after partial results of Cooper–Long–Reid [60] and Bergeron–Wise [25] among others. These two works rely on quite different types of mathematics.

The resolution of (1) by Kahn and Markovic took a rather analytic approach. One first considers a pair of pants with geodesic boundaries in the hyperbolic manifold, pastes a pair of pants (with geodesic boundaries) to each of the boundary components, and then goes on pasting a pair of pants to each free boundary. By a measure-theoretic argument, it is shown that after finitely many steps, the pair of pants comes back very close to the original one. Pasting up all these pairs of pants, an immersed incompressible closed surface is obtained.

Agol's resolution of Conjectures (2) and (4) relies essentially on the study of CAT(0)-cube complexes started by Wise. A cube complex is a complex made of finite-dimensional cubes, $[0, 1]^n$, with isometric pasting maps. A cube complex has a metric induced from the standard metrics on cubes, and is called CAT(0) when it is non-positively curved in the sense of triangle comparison. Bergeron–Wise [25] showed, using the work of Sageev [195], that the result of Kahn–Markovic cited above implies that the fundamental group of every closed hyperbolic 3-manifold acts freely and cocompactly on a CAT(0)-cube complex by isometries. Haglund–Wise [96] considered " hyperplanes" in cube complexes, and introduced the notion of "specialness" for CAT(0)-cube complexes. They then proved that if a hyperbolic group acts freely and cocompactly on a special CAT(0)-cube, then every quasi-convex subgroup is separable. This implies that Conjecture (2) can be proved once we can show that every hyperbolic 3-manifold has a finite-sheeted cover whose fundamental group acts on a special CAT(0)-cube complex freely and cocompactly. Agol proved that this is indeed the case. Conjecture (4) was also resolved by combining this line of argument with previous work of Agol [6].

1.3.5 The Ehrenpreis Conjecture

The Ehrenpreis conjecture for Riemann surfaces states that any two compact Riemann surfaces have finite sheeted unramified covers that are of the same genus and that are arbitrarily close to each other in the Teichmüller metric. It is not clear to the authors of this essay where and when exactly this conjecture was formulated for the first time. In their paper [29], Biswas and Nag refer to it as an "old conjecture which, we understand, is due to L. Ehrenpreis and C. L. Siegel."

The conjecture was proved by J. Kahn and V. Markovic in 2011 (the paper [110] was published in 2015). We mention this here because the proof depends heavily on the geometric methods introduced by Thurston in the topology of surfaces and 3-manifolds. A crucial step in the proof is the construction of what the authors call a "good" geodesic pair of pants decomposition of the surface, that is, a decomposition into pants whose cuff lengths are equal to some fixed large number. Another major ingredient in the proof is an appeal to the proof of the surface subgroup theorem and its proof by the same authors, which also makes heavy use of ideas inaugurated by Thurston. Thurston's influence on the subject is touched upon in the Kahn–Markovic ICM talk [109].

Sullivan and Thurston himself tried to prove this conjecture. In his approach to this question, Sullivan introduced in the early 1990s [202] the notion of *solenoid*, the inverse limit of the system of finite-sheeted branched covers of a fixed closed Riemann surface of negative Euler characteristic. He introduced the Teichmüller space and the mapping class group of this object, and studied their geometry and dynamics. The solenoid became an object of study in itself, see the reviews [179, 196].

In a memorial article on Thurston [126], Sullivan recalls the following, from Milnor's 80th fest at Banff: "I recall a comment whispered by Bill who sat next to me during a talk by Jeremy Kahn about the Kahn–Markovic proof of the subsurface conjecture from decades before. Bill whispered: 'I missed the offset step'." (The "offset step" referred to here is a step in the proof of Kahn–Markovic which concerns the construction of pairs of pants with large cuffs).

1.3.6 The Cannon–Thurston Maps

As an important step in the uniformization theorem for Haken manifolds, Thurston proved that any closed surface (S-)bundle over the circle with a pseudo-Anosov monodromy has a hyperbolic metric. His proof of this result uses the double limit theorem, which gives the Kleinian group corresponding to the fiber. The limit set of such a Kleinian surface group G is the entire sphere, for it is a normal subgroup of a cocompact Kleinian group. Cannon and Thurston [54] proved that there is a continuous map from S^1, which is the limit set of the Fuchsian group Γ isomorphic to $\pi_1(S)$, to the limit set S^2 of G which is equivariant under the action of Γ on S^1 and G on S^2. Thus what they got is a $\pi_1(S)$-equivariant Peano map.

Thurston conjectured that such a map, called the Cannon–Thurston map, from the limit set of a convex cocompact Kleinian group to the limit set of an isomorphic, possibly geometrically infinite group, which is invariant under the group action exists in general. The existence of Cannon–Thurston maps was proved for freely indecomposable Kleinian groups with bounded geometry by Mitra(=Mahan Mj) [154] and Klarreich [119], based on the work of Minsky [151]. This was generalized to the case of punctured surface groups, still with the assumption of bounded geometry outside cusps, by Bowditch [38] and Mahan Mj [155]. McMullen, relying on [152], proved the existence of Cannon–Thurston maps for once-punctured torus Kleinian groups [142]. Finally, the general affirmative resolution of the conjecture was obtained by Manah Mj [156, 157] using the technique of "electrocuting" some parts of the manifold keeping the Gromov hyperbolicity.

Thurston also asked in the same list of unsolved problems if there is continuity of movement of Cannon–Thurston maps with respect to the deformation of the Kleinian group. Mahan–Series [159, 160] proved that when geometrically finite freely indecomposable Kleinian groups converge to a geometrically finite group algebraically, Cannon–Thurston maps converge pointwise, but not necessarily uniformly. They also proved that even if the limit is geometrically infinite, the uniform convergence is obtained provided that the limit is strong (i.e. it is both an algebraic and geometric limit). In the case where the convergence is not strong, they gave an example when even the pointwise convergence fails. Mahan–Ohshika [158] gave a necessary and sufficient condition for the pointwise convergence in the case where the sequence consists of quasi-Fuchsian groups.

1.3.7 Anti-de Sitter Geometry and Transitional Geometry

We start by recalling that for every $n \geq 2$, the n-dimensional Anti-de Sitter (AdS) space is a complete Lorentzian space of constant sectional curvature -1. For $n \geq 2$, the model space of Anti-de Sitter space is the vector space $\mathbb{R}^{n+1}$ equipped with the bilinear form of signature $(n-1, 2)$

$$< x, y >= -x_1 y_1 - x_2 y_2 + x_3 y_3 + \ldots + x_{n+1} y_{n+1}.$$

AdS space is the Lorentzian analogue of hyperbolic space. In 1990, Geoffroy Mess wrote a breakthrough paper[12] in which he gave a completely new approach to Lorentzian geometry in dimension 2+1 and in which he proved a classification theorem for AdS spacetimes, that is, complete Lorentzian manifolds of constant negative curvature, obtained by taking a quotient of anti-de Sitter space by a discrete group of isometries acting freely. Mess's theory heavily uses techniques from Thurston's theory of low-dimensional geometry and topology, which he adapted to the Lorentzian setting. This includes a Lorentzian version of the grafting operation for complex projective surfaces, the parametrization of the moduli space of complex projective structures as a bundle over the Teichmüller space of a surface whose fibers are measured laminations space, actions on $\mathbb{R}$-trees, the holonomy map, convex hull and convex core constructions and a study of the geometry of the boundary of the convex core, bending and bending laminations, representations of surface groups, the analogue in AdS geometry of quasi-Fuchsian representations, the parametrization of moduli spaces of such representations by two copies of Teichmüller space (an analogue of the Bers double uniformization theorem), and earthquakes, with a new proof of Thurston's earthquake theorem.

The notions of Cauchy hypersurface and of globally hyperbolic AdS manifold turned out to be central in this context: A Cauchy hypersurface is a space-like hypersurface which intersects all inextendable time-like lines in the manifold in exactly one point. An AdS manifold (or more generally a Lorentzian manifold) is said to be globally hyperbolic if it contains a Cauchy hypersurface. The notion of Cauchy surfaces was first introduced in the context of general relativity.

Mess, in his paper, gave a complete description of globally hyperbolic spacetimes of constant curvature with compact Cauchy surfaces in dimension 2+1. One of the results he obtained is the classification of proper isometric actions of discrete groups on Minkowski space.

The introduction by Mess of Thurston's techniques in the setting of AdS geometry had an important impact on later research, and we mention some works done in this direction.

Bonsante and Schlenker, in their paper [33] studied a space of AdS manifolds with cone singularities, and showed that this space is parametrized by the product of two copies of the Teichmüller space of the surface with marked points (corresponding to the cone singularities). From this result they deduced an analogue of Thurston's theorem on the transitivity of earthquakes for closed hyperbolic surfaces with cone singularities with total angle less than π. In the paper [34], the same authors showed that it is possible to prescribe any two measured laminations filling a surface, to be the upper and lower measured bending laminations of the convex core of a globally hyperbolic AdS manifold, answering positively a question raised by Mess in his paper.

Bonsante in his paper [32] extended the study of globally hyperbolic flat spacetimes to higher dimensions. Among the tools he introduced is a notion

[12] The paper was published in 2007 [143]; see also the accompanying notes [11].

of measured geodesic stratification which extends to higher dimensions that of measured geodesic lamination. Fillastre in [80] studied Fuchsian polyhedra in such spaces, extending to this setting results of A. D. Alexandrov in [10], Rivin–Hodgson in [193], and Labourie–Schlenker in [124] in which these authors study convex Fuchsian surfaces in Lorentz spaces of constant curvature.

In the paper [21], Barbot and Mérigot established a relation between quasi-Fuchsian and AdS representations which are Anosov in the sense of Labourie [122].

The work of Mess and later developments in AdS geometry are surveyed in Chapter 15 of the present volume, by Bonsante and Seppi [35], and in Chapter 16 by Barbot and Fillastre [20]. The reader may also refer to the survey [24] by Benedetti and Bonsante and the works [19, 36, 37].

The list of open questions [22] is another indication of the direction that this field took in the last few years, motivated by Thurston's ideas.

Talking about AdS geometry in relation with Thurston's work, we are led to transitional geometry, a topic also introduced by Thurston.

A *transition* between two geometries is a continuous path in the space of metrics on a manifold, parametrized by an interval, say $(-1, 1)$, where on the sub-interval $(-1, 0)$ the manifold carries the first geometry (say hyperbolic), on the sub-interval $(0, 1)$ it carries the other geometry (say AdS), and at 0, the geometry is from a third type (say Euclidean). Thurston introduced the notion of transition between his eight 3-dimensional geometries in his proof of the orbifold theorem (see the comments by Cooper, Hodgson and Kerckhoff in [61]),[13] and used the technique of Ricci flow in this process. More details on this topic are given in Sect. 1.2.6 of the present article.

Following Thurston's ideas, there has been a recent activity in dimension 3, on a continuous transition between the eight geometries, and also on varying continuously between Riemannian and Lorentzian geometries on orbifolds. Let us mention a few works on this subject.

Transitions between spherical and hyperbolic geometry, passing through Euclidean geometry, were studied by Cooper, Hodgson and Kerckhoff in [61], Hodgson in [104], Boileau–Porti [30] and Porti in [186]. In the last paper, Porti investigated the appearance of orbifolds with geometry **Nil** as limits of rescaled hyperbolic cone manifolds. In his paper [187], he developed a theory of degeneration/regeneration between hyperbolic 2-orbifolds and hyperbolic cone 3-orbifolds. In his paper with Weiss [188], he developed a transition theory between Euclidean cone manifolds and spherical or hyperbolic ones, with applications to questions of rigidity of Euclidean cone structures. We also mention work on combinatorial transitions, by Kerckhoff and Storm in [116].

In the more recent paper [62], Cooper, Danciger and Wienhard studied transitions between Thurston's geometries in the setting of projective geometry, They gave a

[13]The authors write in particular: "Thurston outlined his proof on two occasions in courses at Princeton; in 1982 and again in 1984. On both occasions, due to running out of time, the outline was incomplete in certain aspects at the end of the proof in the collapsing case. In particular the Euclidean/spherical transition in the case of vertices was treated in a few sentences."

complete classification of limits of three-dimensional hyperbolic geometry inside projective geometry. They showed that the three Thurston geometries $\mathbb{E}^3$, **Nil**, and **Sol** appear among the limits, but the other Thurston geometries do not.

In the papers [65, 66], Danciger studied a smooth transition between the hyperbolic and AdS geometry of 3-manifolds, passing through a transversely hyperbolic 1-dimensional foliation of the manifold. In particular, in the first paper, he introduced, in a study of the transition geometry between hyperbolic and AdS geometry, a transitional projective geometry he called half-pipe geometry.

Two-dimensional transitional geometry, from a completely different point of view, based on the notion of "coherent geometry", is studied by A'Campo and Papadopoulos [2, Chapter 9] and [3].

1.3.8 Linkages

The theory of linkages is a classical subject that combines topology, real-analytic geometry and mechanical constructions. It is not surprising that this topic attracted Thurston's attention. Although he wrote very little on it, he influenced the works of several authors, in private conversations and in lectures. Among them, we mention Henry King, Misha Kapovich, John Millson and Alexei Sossinsky.

In the survey [117] on planar linkages, King recalls that he first heard of the subject in a talk by Thurston at the Institute for Advanced Study, in the mid-1970s. He remembers that Thurston gave a proof of what is called now the *Thurston signature theorem.*[14] The statement can take different forms, one of them being that for any signature, one can construct a planar linkage that approximates it arbitrarily closely. In other words, one can find a linkage such that the locus of one of its vertices (or of a set of vertices), when it runs through all its possible positions, is arbitrarily close to the signature. Another (related) result of Thurston mentioned in King's paper concerns the realization of any compact smooth manifold as a configuration space of a linkage. King writes:

> As far as I can tell, Thurston never wrote these results up, so [the Thurston signature theorem] must remain vague. Occasionally since then I have been contacted by an engineer interested in these results, but I could not recall anything about Thurston's proof so I could not help them. Then recently, Millson started asking me lots of questions on real algebraic sets. He and Kapovich were writing up proofs of the results [of Thurston] above. In the course of doing so, they discovered and solved some problems overlooked by previous literature.

Kapovich and Millson, in the paper they wrote on the subject [112], make a reference to the work of the nineteenth century mathematician Alfred Kempe [113] who studied linkages and obtained weak forms of Thurston's results. This fact is

---

[14]The word "signature" refers here to a person signing her name; this is not to be confused with notions like Rokhlin's signature of a manifold.

recurrent in Thurston's work: He used to develop theories from scratch, and it happened that he realized that some of his ideas or results were discovered by others, in general, several years, and sometimes decades before him. In fact, Thurston was reviving classical subjects. We already mentioned such instances in the section concerning his work on surface diffeomorphisms (Sect. 1.2.15) and in that on circle packings (Sect. 1.2.12). There are many other examples.

Kapovich and Millson write in their paper:

> The first precise formulation of a theorem of the above type was given by W. Thurston who stated a version of Corollary C about 20 years ago and has given lectures on it since. He realized that such a theorem would follow by combining the 19th century work on linkages (i.e. Kempe's theorem) with the work of Seifert, Nash, Palais and Tognoli. However, Thurston did not write up a proof so we have no way of knowing whether he overcame the problems discussed above in the 19th century work on linkages. There is also ambiguity concerning which theorem Thurston formulated in his lectures, we heard three different versions from three sources.

Sossinsky writes in his survey [199] that he was introduced to the theory of linkages by Alexander Kirillov, after the latter returned to Moscow from a stay in the US during which he had heard one of Thurston's talks on the subject. In turn, Sossinsky introduced the subject to several Russian mathematicians who started working on it. He writes that Thurston was mainly interested in the topology of configuration spaces of planar linkages, and that he considered two types of problems, which were the main problems in the field: (1) the so-called direct problem ("configuration"): given a planar linkage (or a class of planar linkages), to ask for a description of the corresponding configuration space(s); (2) the inverse problem ("universality"): given a topological space or an algebraic variety (or a class of such objects), to find a planar linkage whose configuration space is this space (or a class of linkages whose configuration spaces are in the given class). On problem (2), Sossinsky mentions a version of Thurston's signature theorem saying that for any real-algebraic curve in the plane there exists a planar linkage which draws it. Sossinsky's article [199] contains a beautiful historical introduction to the subject.

In a post on MathOverflow, Kevin Walker recalls that as an undergraduate student of Thurston at Princeton, the latter told him about the strategy of the proof of the signature theorem. This proof included a use of Nash's theorem saying that any smooth manifold is diffeomorphic to a real algebraic set, which reduces the problem to that of devising planar linkages implementing addition and multiplication of real numbers and showing how to combine these linkages. Walker wrote his bachelor thesis on linkages. He posed there a conjecture, about recovering the relative lengths of the bars of a linkage from intrinsic algebraic properties of the cohomology algebra of its configuration space. The conjecture was proved several years later by Farber, Hausmann and Schütz [76].

Thurston's popular science article with Jeff Weeks [235], published in 1984, contains several passages on linkages. In particular, we find there the description of a simply defined linkage (which was studied later by several authors under the name Thurston–Weeks triple linkage) whose configuration space has a very interesting topology.

In a correspondence with the second author of this chapter, Bill Abikoff wrote: "Thurston was characteristically terse in his discussion of spaces formed by flexible linkages. His response to the question of which topological spaces appear as the configuration space of a flexible linkage was: *all*."

1.3.9 Higher Teichmüller Theory

Thurston is a forerunner of higher Teichmüller theory. He was the first to emphasize the importance of the study of connected components of the representation variety of the fundamental groups surfaces into Lie groups other than the group $\mathrm{PSL}(2, \mathbb{R})$. He was also the first to revive ideas of Ehresmann from the mid 1930s, highlighting the holonomy as a map from the deformation space of geometric structures into the representation variety, making this a general guiding principle for the classification of locally homogeneous structures. We refer the reader to Goldman's article [88] in which he talks about an *Ehresmann–Weil–Thurston holonomy principle*. Labourie and McShane use the expression "Higher Teichmüller–Thurston theory" for the study of a specific component of the representation space of a surface group of genus in $\mathrm{PSL}(n, \mathbb{R})$. In their paper [123], they extend Thurston's shear coordinates to the setting of Hitchin representations of fundamental groups of surfaces and they prove a McShane–Mirzakhani identity in that setting. Vlamis and Yarmola use the same expression in the paper where they prove a Basmajian identity in higher Teichmüller–Thurston theory [237]. Among the large number of results in higher Teichmüller theory that are inspired by Thurston's work on surfaces, we mention Labourie's work on representations of surface fundamental groups into $\mathrm{PSL}(n, \mathbb{R})$, and in particular his discovery of a curve which is the limit set of the quasi-Fuchsian representation in this setting [122]. We also mention the generalization of Thurston's shear coordinates to the context of decorated representations into split real Lie groups by Fock and Goncharov [81], and the generalization of Thurston's compactification of Teichmüller space to compactifications of spaces of various sorts of representations of finitely generated groups (see e.g. [178] for representations into reductive Lie groups). One should also mention the recasting of Thurston's compactification of Teichmüller space from the point of view of the character variety, in terms of group actions on Λ-trees, by Morgan and Shalen, see [162]. Finally, we mention the work done on the pressure metric on higher Teichmüller spaces (in particular for Anosov representation) a higher-generalization of Thurston's version of the Weil–Petersson metric on Teichmüller space, see [41, 42].

1.3.10 The Grothendieck–Thurston Theory

Alexander Grothendieck, at several places of his manuscript *Esquisse d'un programme* [92] (released in 1984), in which he introduced the theory of *dessins* *d'enfants* and where he set out the basis of the theory that later on became known as Grothendieck–Teichmüller theory, mentions Thurston's work as a source of inspiration. On p. 12 of his manuscript, Grothendieck writes: "The lego-Teichmüller toy which I am trying to describe, arising from motivations and reflections of absolute algebraic geometry on the field $\mathbb{Q}$, is very close to Thurston's hyperbolic geodesic surgery." Grothendieck drew a parallel between his own algebraic constructions in the field $\mathbb{Q}$ of rational numbers and what he calls Thurston's "hyperbolic geodesic surgery" of a surface by pairs of pants decompositions. He outlined in this paper a principle which today bears the name "Grothendieck reconstruction principle," or the "two-level principle." In broad terms, the principle says that some important geometric, algebraic and topological objects that are associated with a surface S (e.g. the Teichmüller space, the mapping class group, the space of measured foliations, spaces of representations of its fundamental group, etc.) can be reconstructed from the "small" corresponding spaces associated with the (generally infinite) set of level-zero, level-one and level-two essential subsurfaces of S. Here, the "level" of a surface is the number of simple closed curves that are needed to decompose it into pairs of pants. Thus, level-zero surfaces are pairs of pants, level-one surfaces are tori with one hole or spheres with four holes, level-two surfaces are the 2-holed tori and 5-holed spheres etc. The geometric structures on the level-zero spaces are the building blocks of the general structures, and the structures on the level-one and the level-two spaces are the objects that encode the gluing. There is a group-theoretic flavor where the level-one surfaces play the role of generators and the level-two surfaces are the corresponding relators. Paraphrasing Grothendieck from his Esquisse d'un programme, "the Teichmüller tower can be reconstructed from level zero to level two, and in this reconstruction, level-one gives a complete set of generators and level-two gives a complete set of relations." Grothendieck made a comparison with analogous situations in algebraic geometry, in particular in reductive group theory, where the semi-simple rank of a reductive group plays the role of "level."

The reconstruction principle was used (without the name) in the paper by Hatcher and Thurston, *A presentation for the mapping class group of a closed orientable surface* [101], published in 1980. In this paper, the authors find a presentation of the mapping class group in which the generators and the relations, which correspond to moves in the pants decomposition complex, are all supported on the level-two surfaces of the given topological surface. The reconstruction principle appears in the same paper at the level of functions: the authors use Cerf theory (the study singularities in the space of smooth functions on the surface) in a construction which is also limited to the level-one and level-two subsurfaces. The analogy between Grothendieck's and Thurston's theories is expanded in Feng Luo's paper *Grothendieck's reconstruction principle and 2-dimensional topology and geometry* [130].

On p. 41 of his manuscript [92], Grothendieck formulates and discusses a conjecture concerning the canonical realization of conformal structures on surfaces by complex algebraic curves. He then declares: "An elementary familiarization with Thurston's beautiful ideas on the construction of Teichmüller space in terms of a very simple game of Riemannian hyperbolic surgery reinforces my premonition."

Grothendieck also used ideas of Thurston in his works on the actions of the absolute Galois group and in profinite constructions in Teichmüller's theory. The author may refer to the surveys [4, 236]. The same ideas are also developed in his manuscript *Longue marche à travers la théorie de Galois* [91], written around the same period.

At the University of Montpellier, where he worked for the last 15 years of his academical life, Grothendieck conducted a seminar on Thurston's theory on surfaces.

Grothendieck again mentions Thurston's work on surfaces in his mathematical autobiography, *Récoltes et semailles* [93, §6.1]. In that manuscript he singles out twelve themes that dominate his work and which he describes as "great ideas" (grandes idées). Among the two themes he considers as being the most important is what he calls the "Galois–Teichmüller yoga", which is precisely the topic that now bears the name Grothendieck–Teichmüller theory [93, §2.8, Note 23].

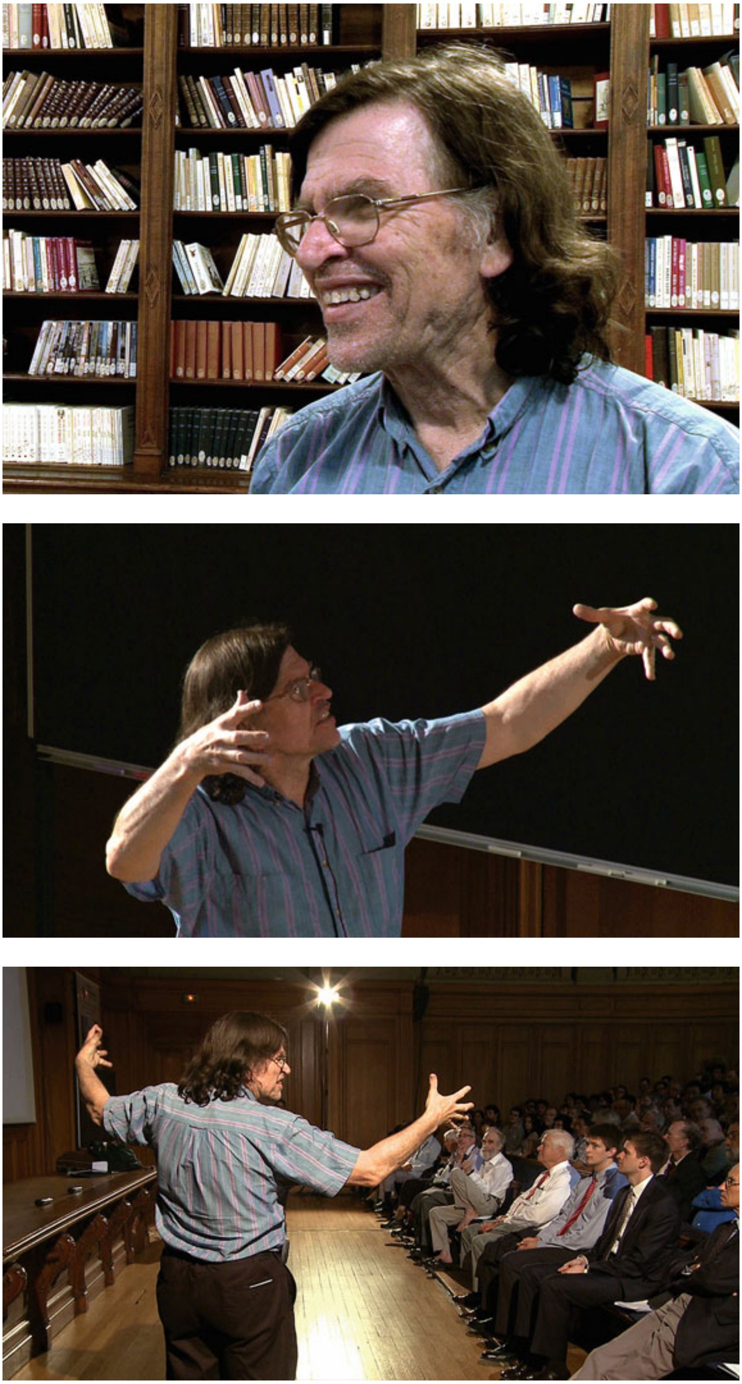
Grothendieck and Thurston had different approaches to Teichmüller space, because the motivations were different (algebraic geometry and low-dimensional topology), but reuniting the two approaches is still now a challenging field of research. Mapping class groups of surfaces occur in the Grothendieck setting in the form of the so-called Grothendieck–Teichmüller group and in the Teichmüller tower, built out of finite type surfaces. The curve complexes and other simplicial complexes from low-dimensional topology have their analogues in this theory. In fact, some of the tools in Grothendick's theory are profinite versions of notions discovered by Thurston. The interested reader may refer to the surveys [83, 236]. Conversely, Grothendieck's dessins d'enfants were studied by several authors in the setting of Thurston's theory; we refer to the surveys [100, 103].

To close this section, let us recall that both Grothendieck and Thurston campaigned against military funding of mathematics. In France, Grothendieck resigned abruptly from his position at IHÉS after he learned that the institute was run partially by military funds. Ten years later, in the US, Thurston was thoroughly involved in a campaign against military funding of mathematics. He wrote several letters to the editors of the *Notices of the AMS*, see e.g. his article *Military funding in mathematics* [220].

1.4 In Guise of a Conclusion

A description of Thurston's work would be incomplete without a few words on his personality.

Thurston valued the notion of mathematical community, and he was pleased to see that he could share his ideas with more and more people. Beyond mathematics, his militancy for a good educational system, for the protection of nature and for a clean environment, his search for beauty, his gentleness, his humbleness, his honesty, and his care for people around him and for humanity in general were exceptionally high. He was a rebel in every sense of the word.



Clay conference in Paris, Oceanographic Institute, June 2010. @ Atelier EcoutezVoir

Acknowledgment We would like to thank Vlad Sergiescu for his comments on the section on foliations.

References

1. W. Abikoff, *The Real Analytic Theory of Teichmüller Space*. Lecture Notes in Mathematics, vol. 820 (Springer, Berlin, 1980)
2. N. A'Campo, A. Papadopoulos, Notes on hyperbolic geometry, in *Strasbourg Master-Class in Geometry* (European Mathematical Society Publishing House, Zürich, 2012), pp. 1–183
3. N. A'Campo, A. Papadopoulos, On transitional geometries, in *Sophus Lie and Felix Klein: The Erlangen Program and its Impact in Mathematics and in Physics*, vol. 23 (European Mathematical Society Publishing House, Zürich, 2015), pp. 217–235
4. N. A'Campo, L. Ji, A. Papadopoulos, Actions of the absolute Galois group, in *Handbook of Teichmüller Theory. Vol. VI*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 27 (European Mathematical Society, Zürich, 2016), pp. 397–435
5. I. Agol, Tameness of hyperbolic 3-manifolds (2004). arXiv.org, May 2004
6. I. Agol, Criteria for virtual fibering. J. Topol. **1**(2), 269–284 (2008)
7. L.V. Ahlfors, Finitely generated Kleinian groups. Amer. J. Math. **86**, 413–429 (1964)
8. D. Alessandrini, V. Disarlo, Generalized stretch lines for surfaces with boundary (2019). Preprint
9. J. W. Alexander, A lemma on systems of knotted curves. Proc. Nat. Acad. Sci. U. S. A. **9**, 93–95 (1923)
10. A.D. Alexandrov, Existence of a convex polyhedron and of a convex surface with a given metric. Rec. Math. (Mat. Sbornik) N.S. **11**(53), 15–65 (1942)
11. L. Andersson, T. Barbot, R. Benedetti, F. Bonsante, W.M. Goldman, F. Labourie, K. Scannell, J.-M. Schlenker, Notes on: "Lorentz spacetimes of constant curvature" by G. Mess. Geom. Dedicata **126**, 47–70 (2007)
12. E.M. Andreev, Convex polyhedra in Lobacevskii spaces. Mat. Sb. (N.S.) **81**(123), 445–478 (1970)
13. E.M. Andreev, Convex polyhedra of finite volume in Lobacevskii space. Mat. Sb. (N.S.) **83**(125), 256–260 (1970)
14. Apollonius de Perge, Coniques. Tome 2.2: Livre IV. Greek and Arabic text, translated into French and annotated under the direction of Roshdi Rashed, in *Historical and Mathematical Commentary*. Scientia Graeco-Arabica, 1/2.2 (Walter de Gruyter, Berlin, 2009)
15. S. Baba On Thurston's parameterization of $\mathbb{C}P^1$-structures, in *In the Tradition of Thurston: Geometry and Topology*, ed. by K. Ohshika, A. Papadopoulos (Springer, Cham, 2020), pp. 241–254
16. H. Baik, K. Kim, Laminar groups and 3-manifolds, in *In the Tradition of Thurston: Geometry and Topology*, ed. by K. Ohshika, A. Papadopoulos (Springer, Cham, 2020), pp. 365–421
17. A. Banyaga, Sur la structure du groupe des difféomorphismes qui préservent une forme symplectique. Comment. Math. Helv. **53**, 174–227 (1978)
18. A. Banyaga, *The Structure of Classical Diffeomorphism Groups* (Kluwer Academic, Dordrecht, 1997)
19. T. Barbot, Lorentzian Kleinian groups, in *Handbook of Group Actions*, vol. III, ed. J. Ji, A. Papadopoulos, S.-T. Yau. Advanced Lecture in Mathematics, vol. 40 (International Press, Somerville, 2018), pp. 311–358
20. T. Barbot, F. Fillastre, Quasi-Fuchsian co-Minkowski manifolds, in *In the Tradition of Thurston: Geometry and Topology*, ed. by K. Ohshika, A. Papadopoulos (Springer, Cham, 2020), pp. 645–703
21. T. Barbot, Q. Mérigot, Anosov AdS representations are quasi-Fuchsian. Groups Geom. Dyn. **6**(3), 441–483 (2012)

22. T. Barbot, F. Bonsante, J. Danciger, W.M. Goldman, F. Guéritaud, F. Kassel, K. Krasnov, J.-M. Schlenker, A. Zeghib. Some open questions in anti-de Sitter geometry (2012). arXiv:1205.6103
23. A. Belkhirat, A. Papadopoulos, M. Troyanov, Thurston's weak metric on the Teichmüller space of the torus. Trans. Amer. Math. Soc. **357**(8), 3311–3324 (2005)
24. R. Benedetti, F. Bonsante, (2+1) Einstein spacetimes of finite type, in *Handbook of Teichmüller Theory, Vol. II*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 13 (European Mathematical Society, Zürich, 2009), pp. 533–609
25. N. Bergeron, D.T. Wise, A boundary criterion for cubulation. Amer. J. Math. **134**(3), 843–859 (2012)
26. L. Bers, On boundaries of Teichm üller spaces and on Kleinian groups. I. Ann. Math. (2) **91**, 570–600 (1970)
27. L. Bers, An extremal problem for quasiconformal mappings and a theorem by Thurston. Acta Math. **141**(1–2), 73–98 (1978)
28. J.S. Birman, Nielsen's investigations of surface mapping class groups, in *Collected Works*, ed. by J. Nielsen (Birkhäuser, Basel, 1986), pp. 407–416
29. I. Biswas, S. Nag, Weil–Petersson geometry and determinant bundles on inductive limits of moduli spaces, in *Lipa's Legacy (New York, 1995)*. Contemporary in Mathematics, vol. 211 (American Mathematical Society, Providence, 1997), pp. 51–80
30. M. Boileau, J. Porti, Geometrization of 3-orbifolds of cyclic type. With an appendix: Limit of hyperbolicity for spherical 3-orbifolds by Michael Heusener and Joan Porti. Paris: Société Mathématique de France. Astérisque **272** (2001)
31. F. Bonahon, Bouts des variétés hyperboliques de dimension 3. Ann. of Math. (2) **124**(1), 71–158 (1986)
32. F. Bonsante, Flat spacetimes with compact hyperbolic Cauchy surfaces. J. Differential Geom. **69**(3), 441–521 (2005)
33. F. Bonsante, J.-M. Schlenker, AdS manifolds with particles and earthquakes on singular surfaces. Geom. Funct. Anal. **19**(1), 41–82 (2009)
34. F. Bonsante, J.-M. Schlenker, Fixed points of compositions of earthquakes. Duke Math. J. **161**(6), 1011–1054 (2012)
35. F. Bonsante, A. Seppi, Anti-de Sitter geometry and Teichmüller theory, in *In the Tradition of Thurston: Geometry and Topology*, ed. by K. Ohshika, A. Papadopoulos (Springer, Cham, 2020), pp. 545–643
36. F. Bonsante, G. Mondello J.-M. Schlenker, A cyclic extension of the earthquake flow I. Geom. Topol. **17**(1), 157–234 (2013)
37. F. Bonsante, G. Mondello, J.-M. Schlenker, A cyclic extension of the earthquake flow II. Ann. Sci. Éc. Norm. Supér. (4) **48**(4), 811–859 (2015)
38. B. Bowditch, The Cannon–Thurston map for punctured-surface groups. Math. Z. **255**(1), 35–76 (2007)
39. P.L. Bowers, Introduction to circle packing: the theory of discrete analytic functions (Book review). Bull. Amer. Math. Soc. (N.S.) **46**(3), 511–525 (2009)
40. P.L. Bowers, Combinatorics encoding geometry: the legacy of Bill Thurston in the story of one theorem, in *In the Tradition of Thurston: Geometry and Topology*, ed. by K. Ohshika, A. Papadopoulos (Springer, Cham, 2020), pp. 173–239
41. M. Bridgeman, R. Canary, F. Labourie, A. Sambarino, The pressure metric for Anosov representations. Geom. Funct. Anal. **25**(4), 1089–1179 (2015)
42. M. Bridgeman, R. Canary, A. Sambarino, An introduction to pressure metrics for higher Teichmüller spaces. Ergodic Theory Dyn. Syst. **38**(6), 2001–2035 (2018)
43. J.F. Brock, K.W. Bromberg, On the density of geometrically finite Kleinian groups. Acta Math. **192**(1), 33–93 (2004)
44. J.F. Brock, K. Bromberg, R. Evans, J. Souto, Tameness on the boundary and Ahlfors' measure conjecture. Publ. Math. Inst. Hautes Études Sci. **98**, 145–166 (2003)
45. J.F. Brock, R.D. Canary, Y.N. Minsky, Yair The classification of Kleinian surface groups, II: The ending lamination conjecture. Ann. of Math. (2) **176**(1), 1–149 (2012)

46. K. Bromberg, Projective structures with degenerate holonomy and the Bers density conjecture. Ann. Math. (2) **166**(1), 77–93 (2007)
47. X. Buff, G. Cui, L. Tan, Teichmüller spaces and holomorphic dynamics. *Handbook of Teichmüller Theory. Vol. IV*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 19 (European Mathematical Society, Zürich, 2014), pp. 717–756
48. D. Calegari, *Foliations and the Geometry of 3-manifolds*. Oxford Mathematical Monographs (Oxford University Press, Oxford, 2007)
49. D. Calegari, N. Dunfield, Laminations and groups of homeomorphisms of the circle. Invent. Math. **152**, 149–207 (2003)
50. D. Calegari, D. Gabai, Shrinkwrapping and the taming of hyperbolic 3-manifolds. J. Amer. Math. Soc. **19**(2), 385–446 (2006)
51. R.C. Canary, Ends of hyperbolic 3-manifolds. J. Amer. Math. Soc. **6**(1), 1–35 (1993)
52. R.D. Canary, Y.N. Minsky, On limits of tame hyperbolic 3-manifolds. J. Differential Geom. **43**(1), 1–41 (1996)
53. J.W. Cannon, The combinatorial structure of cocompact discrete hyperbolic groups. Geom. Dedicata **16**, 123–148 (1984)
54. J.W. Cannon, W.P. Thurston, Group invariant Peano curves. Geom. Topol. **11**, 1315–1355 (2007)
55. J.W. Cannon, W.J. Floyd, M.A. Grayson, W.P. Thurston, Solvgroups are not almost convex. Geom. Dedicata **31**(3), 291–300 (1989)
56. A.J. Casson, S. Bleiler, *Automorphisms of Surfaces after Nielsen and Thurston*. London Mathematical Society Student Texts, vol. 9 (Cambridge University Press, Cambridge, 1988)
57. P.L. Chebyshev, Sur la coupe des vêtements, in *Assoc. Française pour l'Avancement des Sciences, 7ème session à Paris, 28 Août* (1878), pp. 154–155. Reprinted in: P. L. Tchebycheff, Œuvres, Vol. 2, p. 708 (excerpt). Reprint, Chelsea, NY
58. J.H. Conway, J.C. Lagarias, Tiling with polyominoes and combinatorial group theory. J. Combin. Theory Ser. A **53**(2), 183–208 (1990)
59. J.H. Conway, O. Delgado Friedrichs, D.H. Huson, W.P. Thurston, On three-dimensional space groups. Beitr. Algebra Geom. **42**(2), 475–507 (2001)
60. D. Cooper, D.D. Long, A.W. Reid, Essential closed surfaces in bounded 3-manifolds. J. Amer. Math. Soc. **10**(3), 553–563 (1997)
61. D. Cooper, C. Hodgson, S. Kerckhoff, *Three-dimensional Orbifolds and Cone-manifolds, With a postface by S. Kojima*. MSJ Memoirs, vol. 5 (Mathematical Society of Japan, Tokyo, 2000)
62. D. Cooper, J. Danciger, A. Wienhard, Trans. Amer. Math. Soc. **370**, 6585–6627 (2018)
63. E.M. Coven, W. Geller, S. Silberger, W.P. Thurston, The symbolic dynamics of tiling the integers. Isr. J. Math. **130**, 21–27 (2002)
64. M. Culler, P.B. Shalen, Varieties of group representations and splittings of 3-manifolds. Ann. Math. (2) **117**(1), 109–146 (1983)
65. J. Danciger, A geometric transition from hyperbolic to anti-de Sitter geometry. Geom. Topol. **17**(5), 3077–3134 (2013)
66. J. Danciger, Ideal triangulations and geometric transitions. J. Topol. **7**(4), 1118–1154 (2014)
67. G. Darboux, Sur le problème de Pfaff. Bull. Sci. Math. Astron. Sér. 2 **6**(1), 14–36 (1882)
68. A. Douady, J.H. Hubbard, A proof of Thurston's topological characterization of rational functions. Acta Math. **171**(2), 263–297 (1993)
69. K. Delp, W.P. Thurston, Playing with surfaces: Spheres, monkey pants, and zippergons, in *Bridges 2011. Mathematics, Music, Art, Architecture, Culture. 14th Annual Bridges Conference in the University of Coimbra, Portugal* (2011), pp. 1–8
70. D. Dumas, Schwarzian and measured foliations. Duke Math. J. **140**(2), 203–243 (2007)
71. D. Dumas, Complex projective structures, in *Handbook of Teichmüller Theory. Vol. II*. IRMA Lectures in Mathematics and Theoretical Physics, vol. 13 (European Mathematical Society, Zürich, 2009), pp. 455–508
72. C. Ehresmann, Sur les espaces localement homogenes. Enseign. Math. **35**, 317–333 (1936)

73. Y.M. Eliashberg, W.P. Thurston, Contact structures and foliations on 3-manifolds. Turkish J. Math. 20(1), 19–35 (1996)
74. Y.M. Eliashberg, W.P. Thurston, *Confoliations*. University Lecture Series, vol. 13 (American Mathematical Society, Providence, 1998)
75. D.B.A. Epstein, J. Cannon, D. Holt, S. Levy, M.S. Paterson, W.P. Thurston, *Word Processing in Groups* (Jones and Bartlett Publishers, Boston, 1992)
76. M. Farber, J.-C. Hausmann, D. Schütz, On the conjecture of Kevin Walker. J. Topol. Anal. **1**(1), 65–86 (2009)
77. A. Fathi, F. Laudenbach, V. Poénaru, T*ravaux de Thurston sur les surfaces (Séminaire Orsay)*. Astérisque (Société Mathématique de France, Paris, 1979), pp. 66–67. English translation by D. M. Kim and D. Margalit. Mathematical Notes, 48. Princeton University Press, Princeton, NJ, 2012.
78. W. Fenchel, *Elementary Geometry in Hyperbolic Space*. With an editorial by Heinz Bauer. De Gruyter Studies in Mathematics, vol. 11 (Walter de Gruyter, Berlin, 1989)
79. W. Fenchel, J. Nielsen, Discontinuous groups of isometries in the hyperbolic plane. Edited and with a preface by Asmus L. Schmidt. *Biography of the authors by Bent Fuglede*. De Gruyter Studies in Mathematics, vol. 29 (Walter de Gruyter, Berlin, 2003)
80. F. Fillastre, Fuchsian polyhedra in Lorentzian space-forms. Math. Ann. **350**(2), 417–453 (2011)
81. V. Fock, A. Goncharov, Moduli spaces of local systems and higher Teichmüller theory. Publ. Math. Inst. Hautes Études Sci. **103**, 1–211 (2006)
82. S. Francaviglia, A. Martino, Metric properties of outer space. Publ. Mat. **55**(2), 433–473 (2011)
83. L. Funar, Ch. Kapoudjian, V. Sergiescu, Asymptotically rigid mapping class groups and Thompson's groups. *Handbook of Teichmüller Theory. Volume III*, ed. by A. Papadopoulos, IRMA Lectures in Mathematics and Theoretical Physics, vol. 17 (European Mathematical Society, Zürich, 2012), pp. 595–664
84. D. Gabai, Foliations and the topology of 3-manifolds. J. Differential Geom. **18**(3), 445–503 (1983)
85. D. Gabai, S. Kerckhoff (coordinating editors), William P. Thurston. Notices of the AMS **62**(11), 1318–1332 (2015)
86. J. Gilman, On the Nielsen type and the classification for the mapping class group. Adv. Math. **40**(1), 68–96 (1981)
87. W.M. Goldman, Projective structures with Fuchsian holonomy. J. Diff. Geom. **25**(3), 297–326 (1987)
88. W.M. Goldman, Flat affine, projective and conformal structures on manifolds: A historical perspective, in *Geometry in History*, ed. by S. G. Dani, A. Papadopoulos (Springer, Cham, 2019), pp. 515–552
89. M. Gromov, Groups of polynomial growth and expanding maps. Inst. Hautes Études Sci. Publ. Math. **53**, 53–73 (1981)
90. M. Gromov, Volume and bounded cohomology. Inst. Hautes Études Sci. Publ. Math. **56**, 5–99 (1982)
91. A. Grothendieck, *La longue marche à travers la théorie de Galois*, ed. by J. Malgoire. Université Montpellier II, 1995 (1981), 1600 pp.
92. A. Grothendieck, Esquisse d'un programme, mimeographed notes (1984). Available on the internet
93. A. Grothendieck, Récoles et semailles: Réflexions et témoignage sur un passé de mathématicien, manuscript, 1983–1986 (book to appear)
94. F. Guéritaud, F. Kassel, Maximally stretched laminations on geometrically finite hyperbolic manifolds. Geom. Topol. **21**(2), 693–840 (2017)
95. A. Haefliger, Structures feuilletées et cohomologie à valeur dans un faisceau de groupoïdes. Comment. Math. Helv. **32**, 248–329 (1958)
96. F. Haglund, D.T. Wise, Special cube complexes. Geom. Funct. Anal. **17**(5), 1551–1620 (2008)

97. W. Haken, *Some Results on Surfaces in 3-manifolds*. Studies in Modern Topology (Mathematical Association of America, Washington, 1968), pp. 39–98
98. R.S. Hamilton, Three-manifolds with positive Ricci curvature. J. Differential Geom. **17**(2), 255–306 (1982)
99. M. Handel, W.P. Thurston, New proofs of some results of Nielsen. Adv. Math. **56**(2), 173191 (1985)
100. W.J. Harvey, Teichmüller spaces, triangle groups and Grothendieck dessins, in *Handbook of Teichmüller Theory. Vol. I*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 11 (European Mathematical Society, , Zürich, 2007), pp. 249–292
101. A. Hatcher, W.P. Thurston, A presentation for the mapping class group of a closed orientable surface. Topology **19**, 221–237 (1980)
102. D.A. Hejhal, Monodromy groups and linearly polymorphic functions, in *Discontinuous Groups and Riemann Surfaces (Proc. Conf., Univ. Maryland, College Park, MD, 1973).* Annals of Mathematical Studies, No. 79 (Princeton University Press, Princeton, 1974), pp. 247–261
103. F. Herrlich, G. Schmithüsen, Dessins d'enfants and origami curves, in *Handbook of Teichmüller Theory. Vol. II*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 13 (European Mathematical Society, Zürich, 2009), pp. 767–809
104. C.D. Hodgson, Degeneration and regeneration of geometric structures on 3-manifolds. Ph.D. thesis, Princeton University, 1986
105. Y. Huang, A. Papadopoulos, Optimal Lipschitz maps on one-holed tori and the Thurston metric theory of Teichmüller space (2019). Preprint
106. W.H. Jaco, P.B. Shalen, *Seifert Fibered Spaces in 3-manifolds*. Mem. Amer. Math. Soc. 21(220) (1979)
107. K. Johannson, *Homotopy Equivalences of 3-manifolds with Boundaries*. Lecture Notes in Mathematics, vol. 761 (Springer, Berlin, 1979)
108. J. Kahn, V. Markovic, Immersing almost geodesic surfaces in a closed hyperbolic three manifold. Ann. Math. (2) **175**(3), 1127–1190 (2012)
109. J. Kahn, V. Markovic, The surface subgroup and the Ehrenpreis conjectures. [Corrected title: The surface subgroup and the Ehrenepreis conjectures], in *Proceedings of the International Congress of Mathematicians, Seoul 2014*, vol. II (Kyung Moon Sa, Seoul, 2014), pp. 897–909
110. J. Kahn, V. Markovic, The good pants homology and the Ehrenpreis conjecture. Ann. Math. (2) **182**(1), 1–72 (2015)
111. Y. Kamishima, S. Tan, Deformation spaces on geometric structures, in *Aspects of Low-dimensional Manifolds*. Advanced Studies in Pure Mathematics, vol. 20 (Kinokuniya, Tokyo, 1992), pp. 263–299
112. M. Kapovich, J.J. Millson, Universality theorems for configuration spaces of planar linkages. Topology **41**(6), 1051–1107 (2002)
113. A.B. Kempe, On a general method of describing plane curves of the n-th degree by linkwork. Proc. London Math. Soc. **7**, 213–216 (1876)
114. S.P. Kerckhoff, The Nielsen realization problem. Bull. Amer. Math. Soc. (N.S.) **2**(3), 452–454 (1980)
115. S.P. Kerckhoff, The Nielsen realization problem. Ann. Math. (2) **117**(2), 235–265 (1983)
116. S.P. Kerckhoff, P.A. Storm, From the hyperbolic 24-cell to the cuboctahedron. Geom. Topol. **14**(3), 1383–1477 (2010)
117. H.C. King, Planar linkages and algebraic sets. Turkish J. Math. **23**(1), 33–56 (1999). Proceedings of 6th Gökova Geometry-Topology Conference
118. R. Kirby, Problems in low-dimensional topology, in *Geometric Topology, Athens, GA, 1993* (American Mathematical Society, Providence, 1997), pp. 35–473
119. E. Klarreich, Semiconjugacies between Kleinian group actions on the Riemann sphere. Amer. J. Math. **121**(5), 1031–1078 (1999)
120. P. Koebe, Kontaktprobleme der Konformen Abbildung. Ber. Sächs. Akad. Wiss. Leipzig, Math.-Phys. Kl. **88**, 141–164 (1936)

121. S. Kojima, Circle packing and Teichmüller space, in *Handbook of Teichmüller Theory. Vol. II*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 13 (European Mathematical Society, , Zürich, Zürich, 2009), pp. 509–531
122. F. Labourie, Anosov flows, surface groups and curves in projective space. Invent. Math. **165**(1), 51–114 (2006)
123. F. Labourie, G. McShane, Cross ratios and identities for higher Teichmüller–Thurston theory. Duke Math. J. **149**(2), 279–345 (2009)
124. F. Labourie, J.-M. Schlenker, Surfaces convexes fuchsiennes dans les espaces lorentziens à courbure constante. Math. Ann. **316**(3), 465–483 (2000)
125. M. Lackenby, Surface subgroups of Kleinian groups with torsion. Invent. Math. **179**(1), 175–190 (2010)
126. F. Laudenbach, A. Papadopoulos, W.P. Thurston, French mathematics. EMS Surv. Math. Sci. **6**(1), 33–81 (2019)
127. C. Lecuire, The double limit theorem and its legacy, in *In the Tradition of Thurston: Geometry and Topology*, ed. by K. Ohshika, A. Papadopoulos (Springer, Cham, 2020), pp. 263–290
128. W.B.R. Lickorish, A representation of orientable combinatorial 3-manifolds. Ann. Math. (2) **76**, 531–540 (1962)
129. L. Liu, A. Papadopoulos, W. Su, G. Théret, On length spectrum metrics and weak metrics on Teichmüller spaces of surfaces with boundary. Ann. Acad. Sci. Fenn. Math. **35**(1), 255–274 (2010)
130. F. Luo, Grothendieck's reconstruction principle and 2-dimensional topology and geometry, in *Handbook of Teichmüller Theory. Vol. II* (European Mathematical Society, Zürich, 2009), pp. 733–765
131. F. Luo, The Riemann mapping theorem and its discrete counterparts, in *From Riemann to Differential Geometry and Relativity*, ed. by L. Ji, A. Papadopoulos, S. Yamada (Springer, Cham, 2017), pp. 367–388
132. R. Lutz, Sur quelques propriétés des formes différentielles en dimension trois, Thèse Doct. Sci. Math., Université de Strasbourg, Centre Document. C.N.R.S., No. 5851, 90 pp. (1971)
133. R. Mañé, P. Sad, D. Sullivan, On the dynamics of rational maps. Ann. Sci. École Norm. Sup. (4) **16**(2), 193–217 (1983)
134. A. Marden, The geometry of finitely generated Kleinian groups. Ann. Math. (2) **99**, 383–462 (1974)
135. A. Marden, B. Rodin, On Thurston's formulation and proof of Andreev's theorem, in *Computational Methods and Function Theory (Valparaiso, Chile, 1989)*. Lecture Notes in Mathematics, vol. 1435 (Springer, Berlin, 1990), pp. 103–115
136. V. Markovic, Realization of the mapping class group by homeomorphisms. Invent. Math. **168**(3), 523–566 (2007)
137. V. Markovic, D. Šarić, The mapping class group cannot be realized by homeomorphisms (2008). arXiv
138. J. Martinet, Formes de contact sur les variétés de dimension 3, in *Proceedings of Liverpool Singularities Symposium II*. Lecture Notes in Mathematics, vol. 209 (Springer, Berlin, 1971), pp. 142–163
139. B. Maskit, On a class of Kleinian groups. Ann. Acad. Sci. Fenn. Ser. A I No. 442, 8 p. (1969)
140. H.A. Masur, Y.N. Minsky, Geometry of the complex of curves. I. Hyperbolicity. Invent. Math. **138**(1), 103–149 (1999)
141. H.A. Masur, Y.N. Minsky, Geometry of the complex of curves. II. Hierarchical structure. Geom. Funct. Anal. **10**(4), 902–974 (2000)
142. C.T. McMullen, Local connectivity, Kleinian groups and geodesics on the blowup of the torus. Invent. Math. **146**(1), 35–91 (2001)
143. G. Mess, Lorentz spacetimes of constant curvature. Geom. Dedicata **126**, 3–45 (2007)
144. R.T. Miller, Geodesic laminations from Nielsen's viewpoint. Adv. Math. **45**(2), 189–212 (1982)

145. G.L. Miller, S.-H. Teng, W.P. Thurston, S.A. Vavasis, Automatic mesh partitioning, in *Graph Theory and Sparse Matrix Computation*. The IMA Volumes in Mathematics and its Applications, vol. 56 (Springer, New York, 1993), pp. 57–84
146. G.L. Miller, S.-H. Teng, W.P. Thurston, S.A. Vavasis, Separators for sphere-packings and nearest neighbor graphs. J. ACM **44**(1), 1–29 (1997)
147. G.L. Miller, S.-H. Teng, W.P. Thurston, S.A. Vavasis, Geometric separators for finite-element meshes. SIAM J. Sci. Comput. **19**(2), 364–386 (1998)
148. J. Milnor, A note on curvature and fundamental group. J. Differential Geom. **2**, 1–7 (1968)
149. J. Milnor, Collected papers: VI, in *Dynamical Systems (1953–2000)*, ed. by A. Bonifant (American Mathematical Society, Providence, 2013)
150. J. Milnor, W.P. Thurston, On iterated maps of the interval, in *Dynamical Systems, Proc. Spec. Year, College Park/Maryland*. Lecture Notes in Mathematics, vol. 1342 (1988), pp. 465–563
151. Y.N. Minsky, Teichmüller geodesics and ends of hyperbolic 3-manifolds. Topology **32**(3), 625–647 (1993)
152. Y.N. Minsky, The classification of punctured-torus groups. Ann. Math. (2) **149**(2), 559–626 (1999)
153. Y.N. Minsky, The classification of Kleinian surface groups. I. Models and bounds. Ann. Math. (2) **171**(1), 1–107 (2010)
154. M. Mitra, Cannon-Thurston maps for trees of hyperbolic metric spaces. J. Differential Geom. **48**(1), 135–164 (1998)
155. M. Mj, Cannon-Thurston maps for pared manifolds of bounded geometry. Geom. Topol. **13**(1), 189–245 (2009)
156. M. Mj, Cannon-Thurston maps for surface groups. Ann. Math. (2) **179**(1), 1–80 (2014)
157. M. Mj, Cannon-Thurston maps for Kleinian groups, in *Forum of Mathematics*, Pi 5 (Cambridge University Press, Cambridge, 2017), pp. 105–149
158. M. Mj, K. Ohshika, Discontinuous motions of limit sets (2017). arXiv:1704.00269
159. M. Mj, C. Series, Limits of limit sets I. Geom. Dedicata **167**, 35–67 (2013)
160. M. Mj, C. Series, Limits of limit sets II: Geometrically infinite groups. Geom. Topol. **21**(2), 647–692 (2017)
161. J.W. Morgan, H. Bass (eds.), *The Smith Conjecture*. Pure and Applied Mathematics, vol. 112 (Academic Press, Orlando, 1984). Papers presented at the symposium held at Columbia University, New York, 1979
162. J.W. Morgan, P.B. Shalen, Valuations, trees, and degenerations of hyperbolic structures. I. Ann. Math. (2) **120**(3), 401–476 (1984)
163. J.W. Morgan, P.B. Shalen, Degenerations of hyperbolic structures. II. Measured laminations in 3-manifolds. Ann. Math. (2) **127**(2), 403–456 (1988)
164. J.W. Morgan, P.B. Shalen, Degenerations of hyperbolic structures. III. Actions of 3-manifold groups on trees and Thurston's compactness theorem. Ann. Math. (2) **127**(3), 457–519 (1988)
165. S. Morita, Characteristic classes of surface bundles. Invent. Math. **90**, 551–577 (1987)
166. R. Myers, Simple knots in compact, orientable 3-manifolds. Trans. Amer. Math. Soc. **273**(1), 75–91 (1982)
167. H. Namazi, J. Souto, Non-realizability and ending laminations: proof of the density conjecture. Acta Math. **209**(2), 323–395 (2012)
168. J. Nielsen, Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen. I. Acta Math. **50**, 189–358 (1927). English translation: Investigations in the topology of closed orientable surfaces, I, In Jakob Nielsen's Collected Mathematical papers, Vol. I, Birkhäuser, 1986
169. J. Nielsen, Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen. II. Acta Math. **53**, 1–76 (1929). English translation: Investigations in the topology of closed orientable surfaces II, In Jakob Nielsen's Collected Mathematical papers, Vol. I, Birkhäuser, 1986
170. J. Nielsen, Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen. III. Acta Math. **58**, 87–167 (1932). English translation: Investigations in the topology of closed orientable surfaces III, In Jakob Nielsen's Collected Mathematical papers, Vol. I, Birkhäuser, 1986

171. J. Nielsen, Surface transformation classes of algebraically finite type. Danske Vid. Selsk. Mat.-Fys. Medd. **21**(2), 89 p. (1944)
172. K. Ohshika, Rigidity and topological conjugates of topologically tame Kleinian groups. Trans. Amer. Math. Soc. **350**(10), 3989–4022 (1998)
173. K. Ohshika, *Kleinian Groups which are Limits of Geometrically Finite Groups*. Memoirs of the American Mathematical Society, vol. 177(834) (American Mathematical Society, Providence, 2005)
174. K. Ohshika, Realising end invariants by limits of minimally parabolic, geometrically finite groups. Geom. Topol. **15**(2), 827–890 (2011)
175. A. Papadopoulos, Euler and Chebyshev: From the sphere to the plane and backwards. Proc. Cybern. (A volume dedicated to the jubilee of Academician Vladimir Betelin) **2**, 55–69 (2016)
176. A. Papadopoulos, *Ideal Triangles, Hyperbolic Surfaces and the Thurston Metric on Teichmüller Space* (International Press and Higher Education Press, To appear)
177. A. Papadopoulos, G. Théret, On Teichmüller's metric and Thurston's asymmetric metric on Teichmüller space, in *Handbook of Teichmüller Theory, Vol. I*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 11 (European Mathematical Society, Zürich, Zürich, 2007), pp. 111–204
178. A. Parreau, Compactification d'espaces de représentations de groupes de type fini. Math. Z. **272**(1–2), 51–86 (2012)
179. R.C. Penner, Surfaces, circles, and solenoids, in *Handbook of Teichmüller Theory. Vol. I*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 11 (European Mathematical Society, Zürich, 2007), pp. 205–221
180. R.C. Penner, *Decorated Teichmüller Theory. With a Foreword by Yuri I. Manin*. QGM Master Class Series (European Mathematical Society (EMS), Zürich, 2012)
181. R.C. Penner, J.L. Harer, *Combinatorics of Train Tracks*. Annals of Mathematics Studies, vol. 125 (Princeton University Press, Princeton, 1992)
182. G. Perelman, The entropy formula for the Ricci flow and its geometric applications (2002)
183. G. Perelman, Finite extinction time for the solutions to the Ricci flow on certain three-manifolds (2003)
184. G. Perelman, Ricci flow with surgery on three-manifolds (2003)
185. J.F. Plante, W.P. Thurston, Polynomial growth in holonomy groups of foliations. Comment. Math. Helv. **51**(4), 567–584 (1976)
186. J. Porti, Regenerating hyperbolic cone structures from Nil. Geom. Topol. **6**, 815–852 (2002)
187. J. Porti, Regenerating hyperbolic cone 3-manifolds from dimension 2. Ann. Inst. Fourier **63**(5), 1971–2015 (2013)
188. J. Porti, H. Weiss, Deforming Euclidean cone 3-manifolds. Geom. Topol. **11**, 1507–1538 (2007)
189. G. Reeb, Propriétés topologiques des variétés feuilletées, thèse de doctorat, université de Strasbourg,1943, published under the title Sur certaines propriétés topologiques des variétés feuilletées, Actualités Sci. Ind., n^o 1183, Paris, Hermann et Cie, 1952
190. B. Riemann, Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse, (Göttingen, 1851), in *Gesammelte mathematische Werke* (Teubner Verlagsgesellschaft, Leipzig, 1862); new edition: (Springer, Berlin, 1990), pp. 3–48
191. B. Riemann, Theorie der Abel'schen Functionen. J. Reine Angew. Math. **54**, 115–155 (1857). Reprinted in Riemann's Gesammelte mathematische Werke, Teubner Verlagsgesellschaft, Leipzig, 1862; new edition: Springer-Verlag, Berlin (1990) pp. 88–144
192. T.R. Riley, W.P. Thurston, The absence of efficient dual pairs of spanning trees in planar graphs. Electron. J. Combin. **13**(1), 7 p. (2006)
193. I. Rivin, C.D. Hodgson, A characterization of compact convex polyhedra in hyperbolic 3-space. Invent. Math. **111**(1), 77–111 (1993)
194. B. Rodin, D. Sullivan, The convergence of circle packings to the Riemann mapping. J. Differential Geom. **26**(2), 349–360 (1987)
195. M. Sageev, Ends of group pairs and non-positively curved cube complexes. Proc. Lond. Math. Soc. (3) **71**(3), 585–617 (1995)

196. D. Šarić, The Teichmüller theory of the solenoid, in *Handbook of Teichmüller Theory. Vol. II*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 13 (European Mathematical Society, Zürich, 2009), pp. 811–857
197. D. Sleator, R.E. Tarjan, W.P. Thurston, Rotation distance, triangulations, and hyperbolic geometry. J. Amer. Math. Soc. **1**(3), 647–681 (1988)
198. S. Slodkowski, Holomorphic motions and polynomial hulls. Proc. Amer. Math. Soc. **111**, 347–355 (1991)
199. A. Sossinsky, Configuration spaces of planar linkages, in *Handbook of Teichmüller Theory. Vol. VI*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 27 (European Mathematical Society, Zürich, 2016), pp. 335–373
200. W. Su, Problems on the Thurston metric, in *Handbook of Teichmüller Theory. Vol. V*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics (European Mathematical Society, Zürich, 2015), pp. 55–72
201. D. Sullivan, A homological characterization of foliations consisting of minimal surfaces. Comment. Math. Helv. **54**(2), 218–223 (1979)
202. D. Sullivan, Linking the universalities of Milnor–Thurston, Feigenbaum and Ahlfors-Bers, in *Topological Methods in Modern Mathematics (Stony Brook, NY, 1991)*, ed. by L.R. Goldberg, A.V. Phillips (Publish or Perish, Inc., Houston, 1993), pp. 543–564
203. D. Sullivan, W.P. Thurston, Manifolds with canonical coordinate charts: some examples. Enseign. Math. (2) **29**(1–2), 15–25 (1983)
204. D.P. Sullivan, W.P. Thurston, Extending holomorphic motions. Acta Math. **157**(3–4), 243–257 (1986)
205. H. Tanigawa, Grafting, harmonic maps and projective structures on surfaces. J. Differential Geom. **47**(3), 399–419 (1997)
206. W.P. Thurston, Noncobordant foliations of S^3. Bull. Amer. Math. Soc. **78**, 511–514 (1972)
207. W.P. Thurston, On the structure of the group of volume preserving diffeomorphisms (1972). Preprint
208. W.P. Thurston, A generalization of the Reeb stability theorem. Topology **13**, 347–352 (1974)
209. W.P. Thurston, Foliations and groups of diffeomorphisms. Bull. Amer. Math. Soc. **80**(2), 304–307 (1974)
210. W.P. Thurston, The theory of foliations of codimension greater than one. Comment. Math. Helv. **49**, 214–231 (1974)
211. W.P. Thurston, On the construction and classification of foliations, in *Proceedings of the International Congress of Mathematicians (Vancouver, B.C., 1974)*, vol. 1 (Canadian Mathematical Congress, Montreal, 1975), pp. 547–549
212. W.P. Thurston, H.E. Winkelnkemper, On the existence of contact forms. Proc. Am. Math. Soc. **52**, 345–347 (1975)
213. W.P. Thurston, Existence of codimension-one foliations. Ann. Math. (2) **104**(2), 249–268 (1976)
214. W.P. Thurston, Some simple examples of symplectic manifolds. Proc. Am. Math. Soc. **55**(2), 467–468 (1976)
215. W.P. Thurston, *The Geometry and Topology of Three-manifolds*. Lecture Notes (Princeton University Press, Princeton, 1979)
216. W.P. Thurston, Three-dimensional manifolds, Kleinian groups and hyperbolic geometry. Bull. Amer. Math. Soc. (N.S.) **6**(3), 357–381 (1982)
217. W.P. Thurston, A list of questions distributed at the Geometric Topology course, Princeton University, Spring semester, 1983
218. W.P. Thurston, The combinatorics of iterated rational maps (1985). Preprint
219. W.P. Thurston, A norm for the homology of 3-manifolds. Mem. Amer. Math. Soc. **59**(339), 99 (1986)
220. W.P. Thurston, Military funding in mathematics. Not. AMS **34**(1), 39–44 (1987)
221. W.P. Thurston, On the geometry and dynamics of diffeomorphisms of surfaces. Bull. Amer. Math. Soc. (N.S.) **19**(2), 417–431 (1988)
222. W.P. Thurston, Mathematical education. Not. AMS **37**, 844–850 (1990)

223. W.P. Thurston, Conway's tiling groups. Am. Math. Mon. **97**(8), 757–773 (1990)
224. W.P. Thurston, *Three-Dimensional Geometry and Topology*, vol. 1 (Princeton University Press, Princeton, 1997)
225. W.P. Thurston, Zippers and univalent functions, in *The Bieberbach Conjecture (West Lafayette, Ind., 1985)*. Mathematical Surveys and Monographs, vol. 21 (American Mathematical Society, Providence, 1986), pp. 185–197
226. W.P. Thurston, Hyperbolic structures on 3-manifolds. I. Deformation of acylindrical manifolds. Ann. Math. **124**(2), 203–246 (1986)
227. W.P. Thurston, Shapes of polyhedra and triangulations of the sphere, in *The Epstein Birthday Schrift*. Geometry and Topology Monographs, vol. 1 (Geom. Topol. Publ., Coventry, 1998), pp. 511–549
228. W.P. Thurston, Entropy in dimension one, in *Proceedings of a Conference in Celebration of John Milnor's 80th Birthday. Banff, February 2011. Frontiers in Complex Dynamics*. Princeton Mathematical Series, vol. 51 (Princeton University Press, Princeton, 2014), pp. 339–384
229. W. Thurston, Minimal stretch maps between hyperbolic surfaces (1986), arXiv:9801039 [math.GT]
230. W.P. Thurston, Groups, tilings and finite state automata, A series of lectures at the summer AMS colloquium (1989). Preprint
231. W.P. Thurston, Three-manifolds, foliations and circles, I (1997). arXiv:math/9712268v1
232. W.P. Thurston, Hyperbolic structures on 3-manifolds, II: Surface groups and 3-manifolds which fiber over the circle (1998). arXiv:math/9801045
233. W.P. Thurston, Hyperbolic Structures on 3-manifolds, III: Deformations of 3-manifolds with incompressible boundary (1998). arXiv:math/9801058
234. W.P. Thurston, Earthquakes in 2-dimensional hyperbolic geometry, in *Fundamentals of Hyperbolic Geometry: Selected Expositions*. London Mathematical Society Lecture Note Series, vol. 328 (Cambridge University Press, Cambridge, 2006), pp. 267–289
235. W.P. Thurston, J.R. Weeks, The mathematics of three-dimensional manifolds. Sci. Am. **251**, 108–120 (1984). Extended French version: Les variétés à trois dimensions, Pour la Science, n° 83, Sept. 1984, p. 90
236. A.M. Uludağ, I. Sağlam, Hypergeometric Galois actions, in *Handbook of Teichmüller Theory. Vol. VI*, ed. by A. Papadopoulos. IRMA Lectures in Mathematics and Theoretical Physics, vol. 27 (European Mathematical Society, Zürich, 2016), pp. 467–500
237. N. Vlamis, A. Yarmola, Basmajian's identity in higher Teichmüller–Thurston theory. J. Topol. **10**(3), 744–764 (2017)
238. F. Waldhausen, On irreducible 3-manifolds which are sufficiently large. Ann. Math. (2) **87**, 56–88 (1968)
239. F. Waldhausen, Some problems on 3-manifolds. Algebraic and geometric topology, in *Proceedings of the Symposium in Pure Mathematics* (Stanford University, Stanford, 1976), Part 2, pp. 313–322. Proceedings of the Symposium in Pure Mathematics, XXXII (American Mathematical Society, Providence, 1978)
240. S. Wolpert, The Fenchel–Nielsen deformation. Ann. Math (2) **115**(3), 501–528 (1982)
241. S. Wolpert, Thurston's Riemannian metric for Teichmüller space. J. Differential Geom. **23**(2), 143–174 (1986)
242. H. Zieschang, *Finite Groups of Mapping Classes of Surfaces*. Lecture Notes in Mathematics, vol. 875 (Springer, Berlin, 1981)

Chapter 2
Thurston's Influence on Japanese Topologists up to the 1980s

Ken'ichi Ohshika

Abstract In this chapter, I describe how Thurston's work influenced Japanese topologists, focusing on the period 1970s–1980s.

Keywords Foliation · Hyperbolic manifold · Japan · 01A27 · 01A60 · 57K32 · 57R30

2.1 Introduction

Although mathematics is an international discipline by its nature, the way for mathematicians to communicate with each other has enormously changed even during recent 50 years. Electronic mailing system appeared in the community of mathematicians as an indispensable tool only in the mid 1980s. I remember that when I was a young research associate at Tokyo Metropolitan University, the only way to send an electronic mail was to get access to a big computer called Tansei installed in the University of Tokyo. This was the only machine in Japanese academia working also as a mail server. Mathematicians got to know the WWW only around 1993–1994. During the years which we are talking about in this essay, the only way to communicate with foreign mathematicians was to go to meet them physically, which was possible for Japanese mathematicians only during their sabbatical years, or by airmail.

Still, Thurston's work was widely known to Japanese topologists soon after he gave talks or wrote preprints through letters written by the Japanese audience who happened to be present there, or by photocopies of preprints sent by airmails. (To be more precise, photocopies were rather expensive in Japan back in the 1970s. Preprints were mostly disseminated in the form of blueprints at that time.) This fact shows how important Thurston's work was to Japanese topologists those

K. Ohshika (✉)
Department of Mathematics, Gakushuin University, Tokyo, Japan
e-mail: ohshika@math.gakushuin.ac.jp

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_2

days, and was thought of as something very important which you should learn with great effort. In this chapter, I shall describe how Thurston's work influenced Japanese topologists, focusing on the period between the 1970s and the 1980s, the period when Thurston worked mainly on foliations and hyperbolic structures on 3-manifolds.

I am greatly indebted to Takashi Tsuboi for the description of Sect. 2.2 and Sadayoshi Kojima for that of Sect. 2.3, both of whom brought me very precious information.

2.2 Foliations

The study of foliations within Japan started around 1970 by Itiro Tamura, who was a professor in Department of Mathematics of University of Tokyo at that time, and who is the founder of the Japanese school of foliations. The years 1968–1970 are known for the upheaval of the student movement all over the world, which raised many issues ranging from the anti-Vietnam-war campaign to a reform of the old-fashioned university system. In the case of the University of Tokyo (or more broadly in Japanese universities in general), the student movement in this period is often epitomised by two events: one is the occupation of the Yasuda Hall by radical (or far-left depending on one's political standpoint) students from July 1968 to January 1969, and the other is the cancellation of the entrance examination in 1969, for the first time since the foundation the University of Tokyo. Still, behind these scenes which were known to the public, there was a pain-staking effort of both faculties and students to make a concrete and fruitful reform plan for the university. The meeting (a kind of collective bargaining) between delegates of professors and those of students of the Faculty of Science took place in January, as is recorded in the first issue of the monthly report of the Faculty of Science [7]. We can see in the report that Itiro Tamura was a delegate representing the professors of Department of Mathematics. It can be readily imagined that this made him very busy, and it is astonishing that his research in topology could be continued in this situation. Over the Pacific Ocean, there was also a strong student anti-war movement in the USA, which was in the midst of the long Vietnam war. Berkeley, where Thurston was a graduate student, was its epicentre, and anti-war demonstrations took place everyday there. Thurston, who was a convinced antimilitarist for life, was naturally involved in this movement. His anti-military standing did not change all though his life, as his opposition to military funding within mathematics departments, expressed in [19], eloquently tells.

Tamura's first paper on foliation [12] was published in 1972, but apparently was written in 1970–1971. He shows there, generalising the result of Lawson [5], that every odd-dimensional homotopy sphere admits a codimension-one foliation. This paper was the first one on the topic of foliation written by a Japanese mathematician. Reading this paper now, we can see that Tamura's argument relies on his familiarity with the techniques of higher-dimensional differential topology.

In fact, Tamura's work up to this point had mainly dealt with topics in higher-dimensional differentiable manifolds.

Around the same time, Thurston proved a striking result on foliations on the 3-sphere while he was still a graduate student of Berkeley, which appeared in the Bulletin of the AMS [13]. In this paper he shows that there are uncountably many cobordism classes of codimension-1 foliations of the 3-sphere, making use of the Godbillon-Vey invariant, which is a cobordism invariant of codimension-one foliations. Several Japanese topologists remember that they were really impressed with this early work of Thurston. The paper uses arguments and techniques which were quite unfamiliar to them at that time. In 1976, Tamura wrote a book on foliations entitled "Topology of Foliations" [15]. This was the first book on foliations ever published in the world (excepting Reeb's thesis published in the form of a book [10]). The book appeared in the series of mathematical monographs dealing with selected topics in several fields of mathematics which were published by Iwanami Bookstore, regarded as one of the most prestigious publishers for academic books in Japan. The original was written in Japanese, but it was later translated into Russian and English. The last section of this book is devoted to Thurston's result on codimension-one foliations on the 3-sphere, where Tamura described the theorem with a detailed proof. In the postscript of the book, we can see that Thurston's paper was closely studied in Tamura's seminar which included his graduate students at that time, Tadayoshi Mizutani and Toshiyuki Nishimori. It is noteworthy that throughout this book, we can feel much more flavours of low-dimensional topology than Tamura's previous papers, which we can also regard as an influence of Thurston's work. This book served as a textbook for younger Japanese topologists interested in foliations such as Takashi Tsuboi, who read the book when he was the first year student of the graduate school. As his first paper [23] shows, we can find an influence of Thurston's low-dimensional approach on Tsuboi's work both directly or indirectly.

Copies of Thurston's PhD thesis "Foliations of three-manifolds which are circle bundles" [14], which was never published, were widely distributed to Japanese topologists studying foliations, above all by Tamura's students, and the paper was fastidiously studied in seminars. In 1973, the first international conference in Japan in the field of topology/geometry, entitled "Manifolds Tokyo 1973", took place in Tokyo (see [3]). Thurston did not participate in the conference, but his impressive work in Haefliger's classifying space was presented in Mather's talk "Loops and foliations" [6]. Tamura travelled to the USA after the conference, and met Thurston for the first time there. Thurston explained to Tamura his idea to prove the existence of codimension-1 foliations on any closed manifold with zero Euler characteristic, which was later published in 1976 [16]. Tamura later told us that Thurston's explanation was very intuitive and that it was very hard to follow its details.

In the following year, 1974, Thurston went to Princeton as a professor. Shigeyuki Morita stayed at Princeton in the academic year 1974–1975 and attended Thurston's inaugural lecture there, whose topic was foliations. Morita found that the style of Thurston's lecture was quite different from the others and was greatly impressed with it. Thurston drew figures on the blackboard and continued to explain his ideas with a quiet voice for a long time without writing down statements of his theorems. Thurston's interest would move to hyperbolic manifolds in the following years, although he went on writing very original papers in the field of foliations, including those related to what is now called the Thurston norm.

2.3 Hyperbolic Manifolds

Thurston gave a course on hyperbolic manifolds for several years since 1976. Yukio Matsumoto, who stayed at the IAS from 1976 to 1978, attended his lectures. The notes taken by Thurston's students, with the title of "The geometry and topology of 3-manifolds", were widely disseminated, and they inspired many topologists all over the world. Matsumoto sent photocopies of the notes back to the University of Tokyo immediately after they became available, chapter by chapter. Sadayoshi Kojima remembers that he read the notes with great interest but also with some difficulties in the seminar directed by his advisor, Mistuyoshi Kato. The notes contain so many ideas in a very condensed way that some could write a paper, even a thesis, based on what is written in a few pages in the notes. The first five sections of the lecture notes were revised substantially and appeared as a book later [20]. Still the parts which remain unpublished, especially Sections 9 and 10, contain much more inspiring material, in particular for those studying Kleinian groups.

The main purpose of these lectures was to present a proof of his uniformisation theorem for Haken manifolds. The Japanese topologists got to know this great theorem for the first time through an announcement of the resolution of the Smith conjecture (see [8]). This conjecture, which says that every orientation-preserving cyclic group action on the three-sphere with one-dimensional fixed point set is topologically conjugate to an orthogonal action, was very famous among knot theorists. In fact, it was regarded as one of the most important unsolved problems in knot theory along with the Property P conjecture and the Property R conjecture. The news that the Smith conjecture was solved essentially making use of Thurston's uniformisation theorem, which was still under a veil at that time for most of Japanese topologists, was received with awe and amazement. Although Thurston presented his idea of proof in several occasions, including his lectures at Princeton, most people started to understand the overall argument only after John Morgan's seminar which took place at Columbia University in 1980–1981, and whose summary was later published in [8]. Sadayoshi Kojima, who was writing a PhD thesis under Morgan at Columbia then, participated in this seminar, and wrote an expository paper in Japanese on Thurston's uniformisation theorem, making use of what he learnt from this seminar [4]. This was before the publication of [8] or [18], and Kojima's exposition was the only published source from which younger Japanese topologists learnt the uniformisation theorem.

Thurston announced in [17] that he would give a complete proof of his uniformisation theorem in a series of seven papers. In the early 1980s he wrote parts I and II [18, 21], and part III [22] a bit later. Preprints of these papers reached the University of Tokyo, and were widely distributed among Japanese topologists. There are seminars in Japan where a proof of Thurston's uniformisation was studied based on these preprints combined with Kojima's survey. There were not many people in Japan who really understood the proof. Still techniques and tools which Thurston invented there turned out to be very strong and useful for studying hyperbolic 3-manifolds and Kleinian groups, especially if the papers were read with Sections 9 and 10 of the lecture notes. For instance, Soma and myself, who were graduate students at that time took much benefit of them.

In 1982 and 1984, Thurston gave courses on deformations of hyperbolic cone structures, which led to a proof of the geometrisation for 3-manifolds with symmetries (with one-dimensional fixed points), or equivalently the geometrisation of 3-orbifolds. Teruhiko Soma, who was staying at Princeton in 1983–1984, attended these courses, took notes, and brought them back to Japan. In Tokyo, based on his notes, a seminar on deformation of hyperbolic cone structures was held in 1984–1985, whose contents was published in RIMS Kokyuroku [11]. Although it was impossible to recover all the arguments that Thurston had given, in particular the last part where spherical structures are obtained making use of Hamilton's theorem, we could at least present the overall logic and the proofs for the cases where no degeneration occurs and where a Seifert fibration is obtained by degeneration.

In July 1998 an international conference entitled "Cone-Manifolds and Hyperbolic Geometry" was held at the Tokyo Institute of Technology, which Steven Kerckhoff, Sadayoshi Kojima, Tomoyoshi Yoshida and myself organised. A series of talks on Thurston's geometrisation theorem for 3-orbifolds was given there by Cooper, Hodgson and Kerckhoff, in which a detailed proof of the theorem in the case when the singularity is a link (i.e. does not have a vertex) was given. Thurston himself participated in the conference, and this was his first and last visit to Japan for an academic purpose. This talk was later published as a monograph from the Mathematical Society of Japan [2]. Since Thurston himself has not published the proof, this constituted the first piece of literature containing a proof of Thurston's geometrisation theorem of 3-orbifolds, although with an extra hypothesis. A complete proof containing also the case when the singularity contains vertices was finally given by Boileau–Leeb–Porti [1].

2.4 Conclusion

As an another chapter in this volume [9] shows, Thurston's work extends to much wider subjects other than the two main topics which I talked about in this chapter. Also, from around the 1990s, it became easier for Japanese mathematicians to participate in international conferences or to stay abroad for a short time. Up to the 1980s, for Japanese mathematicians, visiting a foreign country meant staying there for a fairly long period, such as during a sabbatical year. This change of the situation made the interaction of Japanese mathematicians and foreign mathematicians more frequent, and international collaborations became quite common. Thurston's influence on Japanese mathematicians after this period has been very diversified. Also, there was a growing number of ex-students of Thurston who in turn had great influence on Japanese topologists and with whom some Japanese topologists worked. Anyway, this is out of the range of the period I am talking about in this chapter.

I finish this chapter with my personal memory. I met Thurston for the first time in 1996 in Berkeley. Yet before that, I had once a correspondence with him. I found some difficulty in understanding one part of his uniformisation theorem of Haken manifolds in the late 1980s, which seemed to me to be a kind of gap in his argument. I sent him a letter explaining the point, and got his reply by airmail. He explained to me how the argument works by elaborating the point over two pages. To be honest, I could not understand his argument, which was rather intuitive, quite well at that time, but reading the letter again now, I realise that the question which I asked him was completely answered without ambiguity. This is parallel to my experience when I read his lecture notes. Sometimes, just reading a few paragraphs was very demanding and it seemed that the argument might have a gap, but after we learnt more and got familiar with his writing, we realised that the essential points were all written down in the notes. This kind of difficulty might have been easily cleared if we could talk with him face to face. Still I have somehow come to believe that this tantalising process might have helped topologists living in Japan to grow and to become independent researchers.

References

1. M. Boileau, B. Leeb, J. Porti, Geometrization of 3-dimensional orbifolds. Ann. Math. (2) **162**(1), 195–290 (2005)
2. D. Cooper, C.D. Hodgson, S.P. Kerckhoff, *Three-dimensional Orbifolds and Cone-manifolds*. MSJ Memoirs, vol. 5 (Mathematical Society of Japan, Tokyo, 2000). With a postface by Sadayoshi Kojima
3. A. Hattori (ed.), Manifolds—Tokyo 1973. Published for the Mathematical Society of Japan by University of Tokyo Press, Tokyo, 1975
4. S. Kojima, On the "monster theorem" of Thurston. Sūgaku **34**(4), 301–316 (1982)
5. H.B. Lawson Jr., Codimension-one foliations of spheres. Ann. Math. (2) **94**, 494–503 (1971)
6. J.N. Mather, Loops and foliations, in *Manifolds—Tokyo 1973 (Proceedings of the International Conference, Tokyo, 1973)* (1975), pp. 175–180
7. Monthly Report of the Faculty of Science, the University of Tokyo 1, 1 (Jan 1968)
8. J.W. Morgan, H. Bass (eds.), *The Smith Conjecture*. Pure and Applied Mathematics, vol. 112 (Academic Press, Orlando, 1984). Papers presented at the symposium held at Columbia University, New York, 1979
9. K. Ohshika, A. Papadopoulos, A glimpse into Thurston's work, in *In the Tradition of Thurston* (Springer, Cham, 2020), pp. 1–58
10. G. Reeb, *Sur certaines propriétés topologiques des variétés feuilletées*. Actualités Sci. Ind., no. 1183 (Hermann & Cie., Paris, 1952). Publ. Inst. Math. Univ. Strasbourg 11, pp. 5–89, 155–156.
11. T. Soma, K. Ohshika, S. Kojima, Towards a proof of Thurston's geometrization theorem for orbifolds, in *Hyperbolic Geometry and* 3*-manifolds (Japanese) (Kyoto, 1985)*, no. 568 (1985), pp. 1–72

12. I. Tamura, Every odd dimensional homotopy sphere has a foliation of codimension one. Comment. Math. Helv. **47**(1), 164–170 (1972)
13. W. Thurston, Noncobordant foliations of S^3. Bull. Amer. Math. Soc. **78**, 511–514 (1972)
14. W.P. Thurston, Foliations of three-manifolds which are circle bundles. Thesis (Ph.D.)–University of California, Berkeley (1972)
15. I. Tamura, *Yosono toporojī (Topology of Foliations)*. Sūgaku Sensho [Mathematical Monographs] (Iwanami Shoten, Tokyo, 1976)
16. W.P. Thurston, Existence of codimension-one foliations. Ann. Math. (2) **104**(2), 249–268 (1976)
17. W.P. Thurston, Three-dimensional manifolds, Kleinian groups and hyperbolic geometry. Bull. Amer. Math. Soc. (N.S.) **6**(3), 357–381 (1982)
18. W.P. Thurston, Hyperbolic structures on 3-manifolds. I. Deformation of acylindrical manifolds. Ann. of Math. (2) **124**(2), 203–246 (1986)
19. W. Thurston, Military funding in mathematics. Notices Amer. Math. Soc. **34**(1), 39–44 (1987)
20. W.P. Thurston, *Three-dimensional Geometry and Topology. Vol. 1*. Princeton Mathematical Series, vol. 35 (Princeton University Press, Princeton, 1997). Edited by Silvio Levy
21. W.P. Thurston, Hyperbolic structures on 3-manifolds, II: Surface groups and 3-manifolds which fiber over the circle. arXiv:math/9801045
22. W.P. Thurston, Hyperbolic structures on 3-manifolds, III: Deformations of 3-manifolds with incompressible boundary. arXiv:math/9801058
23. T. Tsuboi, Foliations with trivial $\mathcal{F}$-subgroups. Topology **18**(3), 223–233 (1979)

Chapter 3
A Survey of the Impact of Thurston's Work on Knot Theory

Makoto Sakuma

Abstract This is a survey of the impact of Thurston's work on knot theory, laying emphasis on the two characteristic features, rigidity and flexibility, of 3-dimensional hyperbolic structures. We also lay emphasis on the role of the classical invariants, the Alexander polynomial and the homology of finite branched/unbranched coverings.

Mathematics Subject Classification (2010) Primary 57M25; Secondary 57M50

3.1 Introduction

Knot theory is the analysis of pairs (S^3, K), where K is a knot (i.e., an embedded circle) in the 3-sphere S^3, and classification of knots has been one of the main problems in knot theory. The Alexander polynomial is an excellent invariant of knots, and it had been a dominating tool and theme in knot theory, until knot theory was influenced by Thurston's work and the Jones polynomial was discovered. In fact, the classical textbook by Crowell and Fox [71] is devoted to the calculation of the knot group and the definition of the Alexander polynomial by using the free differential calculus. The influential textbook by Rolfsen [275] lies emphasis on geometric understanding of the Alexander polynomial through surgery description of the infinite cyclic cover (cf. [134]). However, the Alexander polynomial is far from being complete: there are infinitely many nontrivial knots with trivial Alexander polynomial. The famous Kinoshita–Terasaka knot and the Conway knot are related by mutation, and therefore no skein polynomial, including the Alexander polynomial, can distinguish between them. The first proof of their inequivalence was given by Riley [270] by studying parabolic representations of the knot groups

M. Sakuma (✉)
Osaka City University Advanced Mathematical Institute, Osaka, Japan

Department of Mathematics, Hiroshima University, Higashi-Hiroshima, Japan
e-mail: sakuma@hiroshima-u.ac.jp

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_3

into the finite simple group $\mathrm{PSL}(2, \mathbb{Z}/7\mathbb{Z})$ and the homology of corresponding finite branched/unbranched coverings. (This work led him to the discovery of the hyperbolic structure of the figure-eight knot complement, which inspired Thurston.) Riley called this a *universal method* for obtaining algebraic invariants of knots. The method turned out to be, at least experimentally, a very powerful tool in knot theory, due to the development of computer technology. However theoretical background of the universal method has not been given yet.

In 1976, around the time Rolfsen's book was published, William Thurston started a series of lectures on "The geometry and topology of 3-manifolds". His lecture notes [300] begin with the following words.

> The theme I intend to develop is that topology and geometry, in dimensions up through 3, are intricately related. Because of this relation, many questions which seem utterly hopeless from a purely topological point of view can be fruitfully studied. It is not totally unreasonable to hope that eventually all 3-manifolds will be understood in a systematic way.

This prophecy turned out to be true. Thurston's work has revolutionized 3-dimensional topology, and it has had tremendous impact on knot theory. The first major impact was the proof of the Smith conjecture [219], a result of the efforts by Thurston, Meeks and Yau, Bass, Shalen, Gordon and Litherland, and Morgan. As Morgan predicted in [219, p.6], this was just the beginning of the saga.

In this chapter, we give a survey of the impact of Thurston's work on knot theory. However, the impact is huge, whereas both my ability and knowledge are poor. Moreover, there already exist excellent surveys, including Callahan–Reid [54], Adams [4] and Futer–Kalfagianni–Purcell [96]. So, I decided to lay emphasis on the two characteristic features, *rigidity* and *flexibility*, of hyperbolic 3-manifolds.

As the title of Section 5 of Thurston's lecture notes [300] represents, hyperbolic structures on 3-manifolds have two different features, rigidity and flexibility.

The Mostow–Prasad rigidity theorem says that a complete hyperbolic structure of finite volume on an n-manifold with $n \geq 3$ is rigid: it does not admit local deformation, and moreover, such a structure is unique. Thus any geometric invariant determined by the complete hyperbolic structure of an n-manifold M with $n \geq 3$ is automatically a topological invariant of M. Thurston's uniformization theorem for Haken manifolds implies that almost every knot K in S^3 is hyperbolic, namely the complement $S^3 - K$ admits a complete hyperbolic structure of finite volume. Thus we obtain plenty of topological invariants of hyperbolic knots, including the volume, the maximal cusp volume, the Euclidean modulus of the cusp torus, the length spectrum, the lengths of geodesic paths joining the cusp to itself, the invariant trace field, the invariant quaternion algebra, etc. In particular, the *canonical decomposition* (see Sect. 3.6.1) gives a complete combinatorial invariant for hyperbolic knots, by virtue of the Gordon–Luecke knot complement theorem. The computer program, SnapPea, developed by J. Weeks enables us to calculate the canonical decomposition of hyperbolic knot complements. For example, we can easily detect the inequivalence of the Kinoshita–Terasaka knot and the Conway knot, by checking with SnapPea that the number of 3-cells in the canonical decompositions of the knot complements are 12 and 14, respectively. The rigidity theorem provides us a number of powerful invariants, and it has enriched knot theory by opening new directions of research, namely the study of the behavior of the geometric invariants. An enormous amount of deep research have been made in these new directions. (See Sects. 3.6, 3.8 and 3.9.)

There are two kinds of flexibility of hyperbolic structures on 3-manifolds. One of them is that of cusped hyperbolic manifolds: the complete hyperbolic structure admits nontrivial continuous deformations into incomplete hyperbolic structures. By considering the metric completions of incomplete hyperbolic structures, Thurston established the hyperbolic Dehn filling (surgery) theorem, which says that "almost all" Dehn fillings of an orientable cusped hyperbolic 3-manifold produce complete hyperbolic manifolds. Since every closed orientable 3-manifold is obtained by Dehn surgery of a hyperbolic link, the theorem implies that "almost all closed orientable 3-manifolds" are hyperbolic. This gave strong evidence for Thurston's geometrization conjecture, which was eventually proved by Perelman. The natural and important problem of the study of the exceptional surgeries of hyperbolic knot complements attracted the attention of many mathematicians and numerous research was made on this problem. Due to the development of Heegaard–Floer homology, this problem now attracts renewed interest.

The other flexibility of the 3-dimensional hyperbolic structure is that of complete hyperbolic structures of infinite volume, in other words, the flexibility of complete hyperbolic structures on the interior of a compact orientable 3-manifold whose boundary contains a component with negative Euler characteristic. The deformation theory of such structures is the heart of Kleinian group theory, and it is this flexibility that enabled Thurston to prove the hyperbolization theorem of atoroidal Haken 3-manifolds. In particular, the complete hyperbolic structure of a surface bundle over S^1 (with pseudo-Anosov mondromy) was constructed by developing the deformation theory of the complete hyperbolic structures on $\Sigma \times \mathbb{R}$, where Σ is the fiber surface. The idea of the Cannon–Thurston map, a $\pi_1(\Sigma)$-equivariant sphere filling curve, naturally arose from this construction. Thurston produced various astonishing pictures of (approximations of) Cannon–Thurston maps. (See [302, Figures 8 and 10], [308, Figure 1] and the beautiful book [227] by Mumford–Series–Wright.) It was indeed a shocking event for the author of this survey (who was ignorant of deformation theory and had no idea that it has something to do with knot theory) to learn that a simple topological object, such as the figure-eight knot, carries such mysterious mathematics under cover.

In conclusion, the flexibility of 3-dimensional hyperbolic structure has enriched knot theory by bringing the concept of deformation into knot theory. (See Sects. 3.7 and 3.10.)

In this review, we also consider the role of the classical knot invariants, the Alexander polynomial and the homology of finite branched/unbranched coverings. After the appearance of Thurston's work and the Jones polynomials, the role of these invariants in knot theory might have decreased. However, they continue to be important themes in knot theory. For the Alexander polynomial, its twisted version was defined by Lin [183] for classical knots and by Wada [313] in a general setting. For a hyperbolic knot, we can consider the *hyperbolic torsion polynomial* (see [81]) as the most natural twisted Alexander polynomial, and a beautiful *Thurstonian* *connection* (cf. [10, Section 1.2]) between the topology and geometry of knots is found (see Sect. 3.12.2). For the homology of finite branched/unbranched coverings of a knot, Thang Le [180] proved a mysterious relation between the asymptotic growth of the order of the torsion part and the Gromov norm of the knot (see Sect. 3.13.4). This result is particularly surprising to the author of this survey, for whom homology of finite coverings is a favorite invariant, but who had never imagined that the whole family of the familiar invariant could contain such deep geometric information.

3.2 Knot Theory Before Thurston

In this section, we recall basic definitions and the classical results in knot theory, mostly obtained before knot theory was influenced by Thurston's work: (i) genera of knots, (ii) Schubert's unique prime decomposition theorem, (iii) knot groups, consequences of Waldhausen's work on Haken manifolds, and the Gordon–Lueke knot complement theorem, (iv) fibered knots and open book decompositions, (v) the definition of the Alexander polynomial and its effectiveness and weakness, and (vi) representations of knot groups in finite groups.

The book of Adams [2] is a wonderful introduction to knot theory. For classical results in knot theory, see the textbooks Crowell–Fox [71], Rolfsen [275], Kauffman [159], Burde–Zieschang [52], Kawauchi [161], Murasugi [235], Lickorish [182], Livingston [184], Prasolov–Sossinsky [258], Cromwell [69] and Burde–Zieschang–Heusener [53]. See also the special issue edited by Adams [3] and the handbook Menasco–Thistlethwaite [207].

3.2.1 The Fundamental Problem in Knot Theory

A *knot* K is a smoothly (or piecewise-linearly) embedded circle in the 3-sphere $S^3 = \{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1|^2 + |z_2|^2 = 1\}$. Two knots K and K' are said to be *equivalent*, denoted by $K \cong K'$, if there is a self-homeomorphism f of S^3 such that $f(K) = K'$, i.e., the pair (S^3, K) is homeomorphic to the pair (S^3, K'). If the homeomorphism f preserves the orientation of S^3 and hence is isotopic to the identity homeomorphism 1_{S^3}, then K and K' are said to be *isotopic*. We do not distinguish between a knot K and the equivalence/isotopy class represented by K. A knot is *trivial* if it is isotopic to a standard circle $O := \{(z_1, 0) \in S^3 \mid |z_1| = 1\}$.

Every knot is represented by a *knot diagram*, a 4-valent plannar graph whose vertices are endowed with over/under information. A vertex of a knot diagram with over/under information is called a *crossing*.

For a knot K, we denote by K^* the image of K by an orientation-reversing homeomorphism of S^3, and call it the *mirror image* of K. K^* is represented by the knot diagram which is obtained from that of K by reversing the over/under

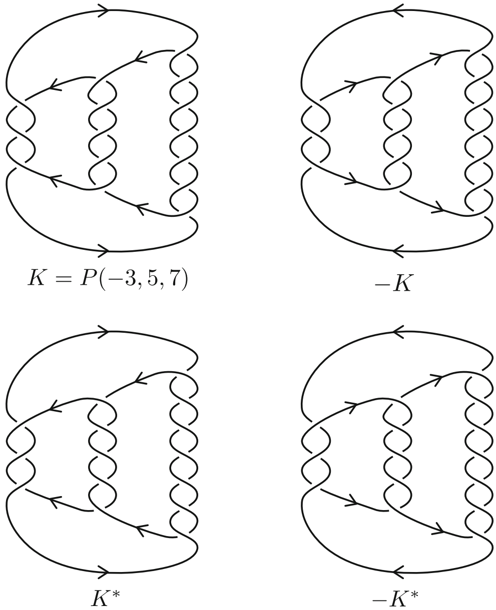


Fig. 3.1 The pretzel knot $P(-3, 5, 7)$ with four different orientations. These oriented knots are not non-isotopic to each other

information at every crossing. A knot K is *achiral* (or *amphicheiral*)[1] if K^* is isotopic to K, otherwise it is *chiral*.

An *oriented knot* is a knot K where the circle K is also endowed with an orientation. (We assume that S^3 is endowed with the standard orientation.) Two oriented knots K and K' are said to be *isotopic* if there is an orientation-preserving self-homeomorphism f of S^3 with $f(K) = K'$ such that $f|_K : K \to K'$ is also orientation-preserving. This is equivalent to the condition that there is an isotopy of S^3 which carries the oriented circle K to the oriented circle K'. For a given oriented knot K, we obtain the following three (possibly isotopic) oriented knots, by reversing one or both of the orientations of S^3 and the circle K (see Fig. 3.1):

$$-K := (S^3, -K), \quad K^* := (-S^3, K) \cong (S^3, K^*), \quad -K^* := (-S^3, -K) \cong (S^3, -K^*)$$

A knot K is *invertible*, *positive-amphicheiral*, or *negative-amphicheiral*, respectively, if K is isotopic to $-K$, K^*, or $-K^*$. If the symmetry can realize by an involution, then we say that K is *strongly invertible*, *strongly positive-amphicheiral*, or *strongly negative-amphicheiral*, respectively (see Fig. 3.2).

It is one of the most fundamental problems in knot theory to detect whether two given knots K and K' are equivalent or not, in particular if a given knot K is trivial or not. The problem of detecting whether a given knot is chiral (or invertible) is a special case of a refinement for oriented knots of this fundamental problem.

[1]This follows [71], though the terminology "amphichiral" seems to be more popular.

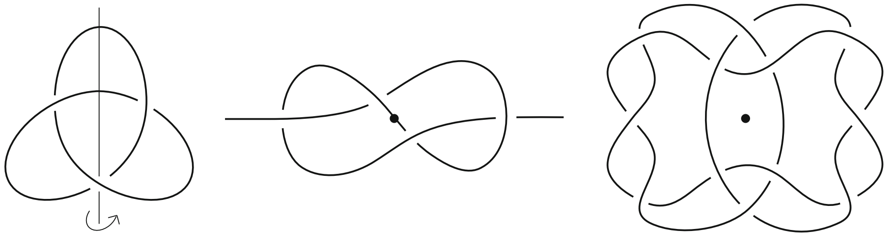


Fig. 3.2 Symmetries of knots realized by involutions

To end this section, we note that the first proof of the existence of non-invertible knots due to Trotter [309] essentially uses 2-dimensional hyperbolic geometry (see the paragraph after Theorem 3.4.2).

3.2.2 Seifert Surface

A *Seifert surface* of a knot K in S^3 is a connected compact orientable surface Σ in S^3 with $\partial\Sigma = K$. The existence of a Seifert surface was first proved by Frankel and Pontryagin [90], through a smooth map $f : S^3 - K \to S^1$ which represents a generator of $H^1(S^3 - K; \mathbb{Z}) \cong \mathbb{Z}$ as the closure of the inverse image $f^{-1}(b)$ of a regular point $b \in S^1$. Later, Seifert gave a simple effective method, called the *Seifert algorithm*, for constructing a Seifert surface from an oriented knot diagram (see Fig. 3.3). The *genus* $g(K)$ of a knot K is the minimum of the genera of Seifert surfaces for K. This is one of the most fundamental invariants of a knot, generalized by Thurston to the concept of Thurston norm. The trivial knot O is characterized by the property $g(O) = 0$.

3.2.3 The Unique Prime Decomposition of a Knot

We recall Shubert's unique prime decomposition theorem, which reduces the classification problem of knots to that of prime knots. Given two oriented knots K_1 and K_2, we can define the *composition* $K_1\#K_2$ as the pairwise connected sum $(S^3, K_1)\#(S^3, K_2)$ of oriented manifold pairs, as in Fig. 3.4. With respect to the connected sum, the set of all oriented knots up to isotopy becomes a *commutative* semi-group having the trivial knot O as the unit.

A knot K is *prime* if $K \cong K_1\#K_2$ implies $K_1 \cong O$ or $K_2 \cong O$. It is a classical theorem due to Schubert [283] that every oriented knot has a unique prime decomposition.

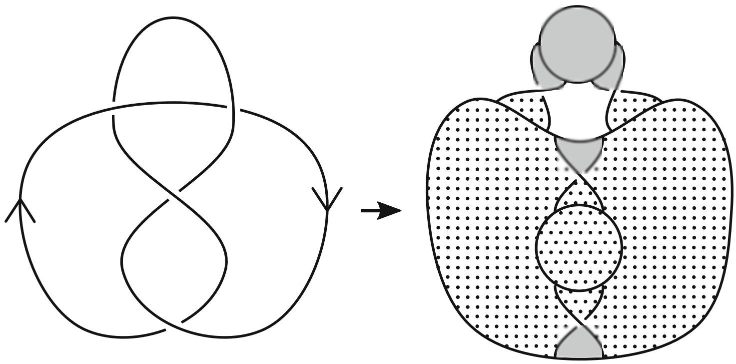


Fig. 3.3 Seifert algorithm: By smoothing all crossings of a knot diagram, we obtain *Seifert circles* (mutually disjoint circles in the plane). Construct mutually disjoint disks in $\mathbb{R}^3$ bounded by the Seifert circles, and join them by bands. The resulting surface is a Seifert surface

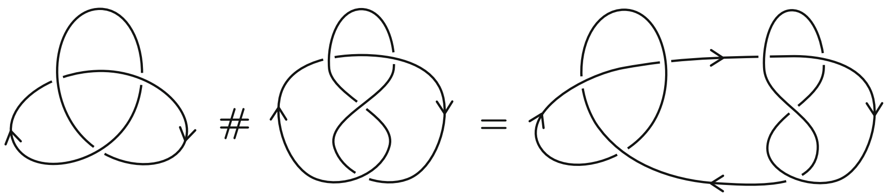


Fig. 3.4 Connected sum of knots

Theorem 3.2.1 (The Unique Prime Decomposition of Knots) *Every nontrivial oriented knot* (S^3, K) *can be decomposed as the sum of finitely many nontrivial prime oriented knots. Moreover if* $K \cong K_1 \# K_2 \# \cdots \# K_n$ *and* $K \cong J_1 \# J_2 \# \cdots \# J_m$ *with each* K_i *and* J_i *nontrivial prime knots, then* $m = n$*, and after reordering,* $K_i \cong J_i$ *as oriented knots.*

The existence of a prime decomposition is guaranteed by the additivity of genus with respect to connected sum, i.e.,

$$g(K_1 \# K_2) = g(K_1) + g(K_2) \quad \text{for any oriented knots } K_1 \text{ and } K_2.$$

The uniqueness of the prime decomposition is proved by a simple cut and paste argument.

3.2.4 Knot Complements and Knot Groups

The *exterior* of a knot K is defined by $E(K) := S^3 - \operatorname{int} N(K)$, where $N(K)$ is a regular neighborhood of K. The *knot complement* $S^3 - K$ is homeomorphic to the interior of $E(K)$, and the fundamental group $\pi_1(S^3 - K) \cong \pi_1(E(K))$ is called the *knot group*, and denoted by $G(K)$. By using the sphere theorem [252] (cf. [127, Chapter 4]), we can see that $E(K)$ is aspherical, and hence the homotopy type of $E(K)$ is completely determined by the knot group $G(K)$. A group presentation, called the *Wirtinger presentation*, of $G(K)$ can be obtained from a knot diagram of K (see [71, 275]).

The *peripheral subgroup* $P(K)$ of the knot group $G(K)$ is defined as (the conjugacy class of) the image of the homomorphism $j_* : \pi_1(\partial E(K)) \to \pi_1(E(K))$ induced by the inclusion map $j : \partial E(K) \to E(K)$. For the trivial knot O, $E(O)$ is homeomorphic to the solid torus, and so $G(O) = P(O) \cong \mathbb{Z}$. Dehn's lemma, established by Papakyriakopoulos [252], gives the following characterization of the trivial knot.

Theorem 3.2.2 *A knot K is trivial if and only if the following mutually equivalent conditions hold.*

(1) $\mathrm{Ker}[j_* : \pi_1(\partial E(K)) \to \pi_1(E(K))]$ *is nontrivial.*
(2) The peripheral subgroup $P(K)$ is isomorphic to $\mathbb{Z}$.
(3) The knot group $G(K)$ is isomorphic to $\mathbb{Z}$.

If K is a nontrivial knot, then the peripheral subgroup $P(K) \cong \pi_1(\partial E(K)) \cong \mathbb{Z} \oplus \mathbb{Z}$ is generated by two special elements, a *meridian* μ and a *longitude* λ, represented by the simple loops $\mu := \partial D^2 \times \{*\}$ and $\lambda := \{*\} \times S^1$ respectively, in $\partial E(K) = \partial N(K) = \partial(D^2 \times S^1)$. Here the framing $N(K) \cong D^2 \times S^1$ is chosen so that the linking number $\mathrm{lk}(K, \lambda) = 0$. When K is oriented, the orientations of μ and λ are chosen so that $\mathrm{lk}(K, \mu) = +1$ and that K and λ are homologous in $N(K)$.

The classical work of Waldhausen [317] on Haken manifolds implies the following theorem which reduces the equivalence problem for knots to a problem of knot groups.

Theorem 3.2.3 *For two knots K and K', the following hold.*

(1) *$E(K)$ and $E(K')$ are homeomorphic if and only if $(G(K), P(K))$ and $(G(K'), P(K'))$ are isomorphic, i.e., there is an isomorphism $\varphi : G(K) \to G(K')$ such that $\varphi(P(K)) = P(K')$ up to conjugacy.*
(2) *K and K' are equivalent if and only if $(G(K), P(K), \mu)$ and $(G(K'), P(K'), \mu'^{\pm 1})$ are isomorphic, i.e., there is an isomorphism $\varphi : (G(K), P(K)) \to (G(K'), P(K'))$ such that $\varphi(\mu) = \mu'^{\pm 1}$ up to conjugacy.*

For nontrivial oriented knots K_1 and K_2, the knot groups $G(K_1 \# K_2)$ and $G(K_1 \# (-K_2^*))$ are isomorphic. In fact, both $E(K_1 \# K_2)$ and $E(K_1 \# (-K_2^*))$ are obtained from $E(K_1)$ and $E(K_2)$ by gluing annuli in their boundaries, and so homotopy equivalent to the space obtained from $E(K_1)$ and $E(K_2)$ by identifying the meridians μ_1 and μ_2. On the other hand, by the unique prime decomposition Theorem 3.2.1, the oriented knots $K_1 \# K_2$ and $K_1 \# (-K_2^*)$ are isotopic if and only if K_2 is negative amphicheiral (i.e., $-K_2^*$ is isotopic to K_2). Thus, in general, the knot group alone is not a complete invariant for knots.

Building on the cyclic surgery theorem (Theorem 3.11.2) by Culler, Gordon, Luecke and Shalen [76], Whitten [324] proved that prime knots with isomorphic knot groups have homeomorphic exteriors. On the other hand, we have the following celebrated theorem of Gordon and Luecke [114].

Theorem 3.2.4 (Knot Complement Theorem) *Two knots are equivalent if and only if they have homeomorphic complements.*

Thus we have the following theorem.

Theorem 3.2.5 *Two prime knots are equivalent if and only if they have isomorphic knot groups.*

3.2.5 Fibered Knots

A knot K is *fibered* if $E(K)$ has the structure of a bundle over the circle, namely, there is a connected compact orientable surface Σ and an orientation-preserving homeomorphism $\varphi : \Sigma \to \Sigma$, such that

$$E(K) \cong \Sigma \times [0, 1]/(x, 0) \sim (\varphi(x), 1).$$

The homeomorphism φ is called the *monodromy* of the fiber structure. Each fiber Σ of the bundle structure is a compact orientable surface in $E(K)$ such that $\Sigma \cap \partial E(K) = \partial\Sigma$ is a longitude of K. The union of Σ and an annulus in $N(K)$ cobounded by $\partial\Sigma$ and K is a minimal genus Seifert surface for K. This is the unique minimal genus Seifert surface for K up to isotopy fixing K.

We may choose φ so that its restriction to $\partial\Sigma$ is the identity map and thus the image of $y \times [0, 1]$ in $E(K)$ is a meridian of K for every $y \in \partial\Sigma$. Then

$$(S^3, K) \cong (\Sigma, \partial\Sigma) \times [0, 1]/[(x, 0) \sim (\varphi(x), 1);\ \ y \times [0, 1] \sim y \text{ (for } y \in \partial\Sigma)].$$

This structure is called an *open book decomposition* with *binding* K, and the homeomorphism φ is called the *monodromy* of the fibered knot K. It was proved by Alexander [17] that every connected closed orientable 3-manifold admits an open book decomposition. Later, Giroux [107] found a very important correspondence between the open book decompositions (up to positive stabilization) of a given closed oriented 3-manifold M and oriented contact structures on M up to isotopy (see [83] for details). The following characterization of fibered knots in terms of knot groups was proved by Stallings [295], and attracted the attention of researchers at the time.

Theorem 3.2.6 *A knot K in S^3 is a fibered knot if and only if the commutator subgroup $G(K)' = [G(K), G(K)]$ is finitely generated.*

The only if part follows from the fact that the infinite cyclic covering $E_\infty(K)$ of $E(K)$, introduced in the subsection below, is identified with $\Sigma \times \mathbb{R}$ and so $G(K)' \cong \pi_1(E_\infty(K)) \cong \pi_1(\Sigma)$ is a free group of rank $2g(K)$. The heart of the theorem is that the converse also holds.

3.2.6 Alexander Invariants

Though the knot group is a complete invariant for prime knots, it is, in general, not easy to distinguish two given knot groups. The Alexander polynomial serves as a convenient and tractable tool for this problem, even though it is not almighty.

Let K be an oriented knot. Then the first integral homology group $H_1(E(K);\mathbb{Z})$ is the infinite cyclic group generated by the image, t, of the meridian μ. Thus there is a unique infinite cyclic covering $p_\infty : E_\infty(K) \to E(K)$, and the covering transformation group is identified with the infinite cyclic group $\langle t \rangle$ generated by t. $H_1(E_\infty(K);\mathbb{Z})$ has the structure of a module over the integral group ring $\mathbb{Z}\langle t \rangle$. This module is called the *knot module*. As an abelian group, $H_1(E_\infty(K);\mathbb{Z})$ is identified with $G(K)'/G(K)''$ where $G(K)'$ and $G(K)''$, respectively, are the first and second commutator subgroups of $G(K)$. Moreover the action of the generator t is given by $t[\alpha] = [\mu\alpha\mu^{-1}]$ for $\alpha \in G(K)'$, where μ is a meridian. Thus the knot module is determined by $G(K)$. In fact, a presentation matrix is obtained from a presentation of the knot group, via Fox's free differential calculus (see [71], [161, Chapter 7]). The *Alexander polynomial* $\Delta_K(t)$ of K is defined as the generator of the first elementary ideal of the knot module.

A more conceptual definition can be given by using the $\mathbb{Q}\langle t \rangle$-module $H_1(E_\infty(K);\mathbb{Q})$ as follows. Since the rational group ring $\mathbb{Q}\langle t \rangle$ is a principal ideal domain and since $H_1(E_\infty(K);\mathbb{Q})$ is a finitely generated torsion module over $\mathbb{Q}\langle t \rangle$, we have

$$H_1(E_\infty(K);\mathbb{Q}) \cong \frac{\mathbb{Q}\langle t \rangle}{(f_1(t))} \oplus \cdots \oplus \frac{\mathbb{Q}\langle t \rangle}{(f_r(t))},$$

where $f_i(t)$ are elements of $\mathbb{Z}\langle t \rangle$ whose coefficients are relatively prime. Then $\Delta_K(t) \doteq f_1(t) \cdots f_r(t)$, where $\doteq$ means equality up to multiplication by a unit $\pm t^i$ of the integral Laurent polynomial ring $\mathbb{Z}\langle t \rangle$. The Alexander polynomial $\Delta_K(t)$ is an integral Laurent polynomial in the variable t, defined up to multiplication by a unit. For the trivial knot O, we have $\Delta_O(t) \doteq 1$. We summarize basic properties of the Alexander polynomial.

Theorem 3.2.7

(1) *For any knot K, its Alexander polynomial $\Delta_K(t)$ satisfies the following condition.*

$$\Delta_K(1) = \pm 1, \quad \Delta_K(t^{-1}) \doteq \Delta_K(t)$$

Conversely, for any Laurent polynomial $\Delta(t)$ satisfying the above condition, there is a knot K whose Alexander polynomial is equal to $\Delta(t)$.

(2) *For every knot K in S^3, we have the following estimate of the genus:*

$$g(K) \geq \deg \Delta_K(t).$$

(3) *For any fibered knot K, the Alexander polynomial $\Delta_K(t)$ is monic, and the equality hold in the estimate (2).*

Proof of Theorem 3.2.7. The proof relies on an analysis of the manifold $M := E(K)\backslash\Sigma$, the manifold obtained from $E(K)$ by cutting along a Seifert surface Σ; in other words, M is the complement of an open regular neighborhood of Σ in $E(K)$. Let Σ_+ and Σ_- be copies of Σ on ∂M_Σ, and consider the annulus $A := M \cap \partial E(K)$. Then $(M, \Sigma_+, \Sigma_-, A)$ is a *sutured manifold* (see [98, 100], [161, Chapter 5]), and this together with the natural homeomorphism $\Sigma_+ \to \Sigma_-$ recovers $E(K)$. The infinite cyclic covering $E_\infty(K)$ is obtained from the set of copies $\{M_n\}$ of M indexed with $n \in \mathbb{Z}$, by gluing the copy of Σ_- in M_n with the copy of F_+ in M_{n+1}. The homological glueing information is given by the *Seifert matrix* $V = (\mathrm{lk}(\alpha_i, \alpha_j^+))_{1\leq i,j\leq 2g}$, where $\{\alpha_i\}_{1\leq i,j\leq 2g}$ with $g = 2g(\Sigma)$ is a set of oriented simple loops on Σ which forms a basis of $H_1(\Sigma)$, α_j^+ is a copy of α_j on the $+$-side of Σ, and $\mathrm{lk}(\cdot,\cdot)$ denotes the linking number. The matrix $tV - V^T$ gives a presentation matrix of $H_1(E_\infty(K))$ as a $\mathbb{Z}\langle t\rangle$-module, and hence $\Delta_K(t) = \det(tV - V^T)$. Using this formula we can prove Theorem 3.2.7.

For knots with small crossing numbers, the Alexander polynomial is quite efficient. For any prime knot K up to 10 crossings, equality holds in the estimate of the genus in Theorem 3.2.7(2). Moreover, such a knot K is fibered if and only if $\Delta_K(t)$ is monic (see Kanenobu [155]).

The Alexander polynomial is also very efficient for alternating knots. A knot K is said to be *alternating* if it is represented by an *alternating diagram*, namely a diagram in which the crossings alternate under and over as one travels along the diagram. A knot diagram is said to be *reduced* if there is no circle in the plane which intersects the diagram only at a single crossing. Any alternating diagram can be deformed into a (possible trivial) reduced alternating diagram. When K is an alternating knot and Σ is a Seifert surface obtained by the Seifert algorithm from a reduced alternating diagram of K, the complementary sutured manifold $(M, \Sigma_+, \Sigma_-, A)$ has a nice structure, which in particular implies $\det(V) \neq 0$. This shows that the estimate Theorem 3.2.7(2) is sharp for alternating knots (see Crowell

[70] and Murasugi [233]). Moreover, Murasugi [234] proved that the converse to Theorem 3.2.7(3) also holds for alternating knots.

Theorem 3.2.8 *For any alternating knot K, the following hold.*

(1) $g(K) = \deg \Delta_K(t)$.
(2) *K is fibered if and only if $\Delta_K(t)$ is monic.*

In order to prove the above results, Murasugi introduced the concept of a *Murasugi sum* of two Seifert surfaces. The simplest case corresponds to the connected sum of knots and the second simplest case corresponds to *plumbing* introduced by Stallings [296]. It was later shown by Gabai [99] that the Murasugi sum is a natural geometric operation in the following sense: If Σ is a Murasugi sum of Σ_1 and Σ_2, then the following hold.

(1) Σ is of minimal genus if and only if Σ_1 and Σ_2 are of minimal genus.
(2) Σ is a fiber surface if and only if Σ_1 and Σ_2 are fiber surfaces.

In addition to Theorems 3.2.7 and 3.2.8, various applications of the Alexander polynomials were found. Among them, we explain a theorem by Kinoshita [165], which gives a condition on the Alexander polynomial that a counter-example to the Smith Conjecture must satisfy. As described in Sect. 3.4.4, the Smith conjecture (Theorem 3.4.6) was later proved using Thurston's geometrization theorem for Haken manifolds.

Theorem 3.2.9 *If K is a fixed point of an orientation-preserving periodic diffeomorphism of period n, then there is an integral Laurent polynomial $f(t)$ such that*

(1) $\Delta_K(t^n) = \Pi_{i=0}^{n-1} f(\xi^i t)$ *where ξ is a primitive n-th root of unity, and*
(2) $f(1) = \pm 1, \quad f(t^{-1}) \doteq f(t)$

See [161, Chapter 10] for other applications of the Alexander polynomial to the study of symmetry of knots, including the first proof of the non-invertibility of the knot 8_{17} by Kawauchi [160], answering to a question of Fox (cf. [88, Problem 10]). (Another proof of the non-invertibility of 8_{17} was announced almost at the same time by Bonahon and Siebenmann, based on their characteristic splitting theory (see Sect. 3.4).) It should be noted that though the definition of the Alexander polynomial depends on the orientation of K, by Theorem 3.2.7(1) the resulting $\Delta_K(t)$ does not depend on the orientation. It is interesting that, despite this fact, the Alexander polynomial can be used for the study of invertibility and chirality of knots. Finally, we point out that the Alexander module $H_1(E_\infty(K))$ does depend on the orientation of K, though it is not easy to detect the dependence (see [89], [133]).

Though we have observed the effectiveness of the Alexander polynomial, there are a lot of knots for which the Alexander polynomial is useless. In fact, H. Seifert [288], J.H.C. Whitehead [323], and Kinoshita–Terasaka [166] gave systematic construction of nontrivial knots with trivial Alexander polynomial. For example, the pretzel knot $K(-3, 5, 7)$ in Fig. 3.1, the Whitehead double of any nontrivial knot (cf. Fig. 3.6), the *Kinoshita–Terasaka knot* and the *Conway knot* in Fig. 3.5 have trivial Alexander polynomial.

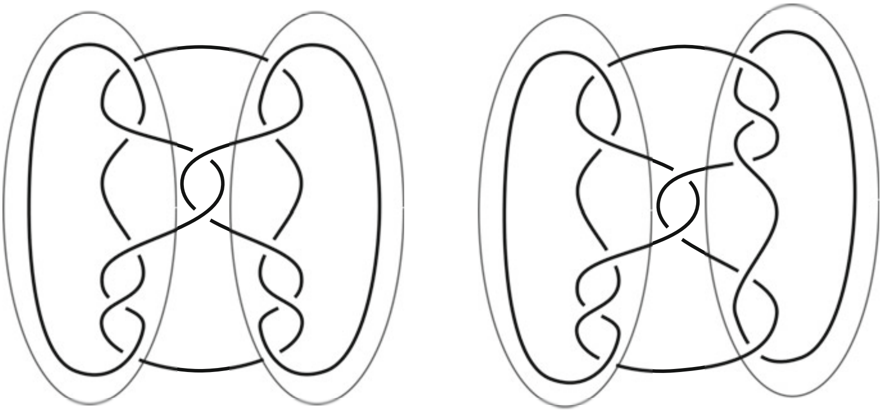


Fig. 3.5 The Kinoshita–Terasaka knot and the Conway knot: The circles in the figure represent the Conway spheres which determine the Bonahon–Siebenmann decompositions, described in Sect. 3.4.1. From this picture, we can see that the Conway knot is a mutant of the Kinoshita Terasaka knot

The Conway knot is a *mutant* of the Kinoshita–Terasaka knot (see Construction 4 in Sect. 3.4.5 for the precise definition). It is known that various invariants coincide for a knot and its mutant, including the Alexander polynomial, the Jones polynomial, the Homflypt polynomial, the double branched covering, and Gromov norm. So it is not easy to distinguish a knot from its mutant.

3.2.7 Representations of Knot Groups onto Finite Groups

The definition of the Alexander polynomial is based on the fact that the knot group $G(K)$ of an oriented knot K admits a unique preferred epimorphism onto the infinite cyclic group $\langle t\rangle$. By replacing $\mathbb{Z}$ with an arbitrary group Γ, we obtain the following family of invariants of knots. Let $R(G(K),\Gamma)$ be the set of homomorphisms from $G(K)$ to Γ, up to conjugacy (i.e., modulo post composition of inner-automorphisms of Γ), is an invariant of $G(K)$. Then its cardinality $|R(G(K),\Gamma)|$ is an invariant of the knot group $G(K)$. We may also consider the quotient of $R(G(K),\Gamma)$ by the action of the automorphism group of Γ.

Fix a conjugacy class γ of an element of Γ, and let $R(G(K),\Gamma,\gamma)$ be the subset of $R(G(K),\Gamma)$ consisting of the homomorphisms which map the meridian to an element in the conjugacy class γ. Then the cardinality $|R(G(K),\Gamma,\gamma)|$ is again an invariant of the oriented knot K. If Γ is the dihedral group $D_{2p} = \langle a,t \mid a^p, t^2, tat^{-1} = a^{-1}\rangle$ of order $2p$ and if γ is the conjugacy class of the element t, the *Fox p-coloring number* [87] is essentially equal to $|R(G(K), D_{2p}, t)|$.

Fix a transitive representation of Γ into the symmetric group S_n of degree n, where n is possibly infinite. Then for each $\phi \in R(G(K),\Gamma)$, we have a (possibly disconnected) n-fold covering $E_\phi(K)$ of $E(K)$. Then the family of homology groups $\{H_1(E_\phi(K);\mathbb{Z})\}_{\phi\in R(G(K),\Gamma)}$ forms an invariant of the knot K. Furthermore, if the image of the element γ in S_n is of finite order, then we obtain a branched covering $M_\phi(K)$ of S^3 branched over K. The family of homology groups $\{H_1(M_\phi(K);\mathbb{Z})\}_{\phi\in R(G(K),\Gamma,\gamma)}$ forms another invariant of K. We can also consider the torsion linking numbers among the components of the inverse image of K.

Riley [270] applied this method by choosing Γ to be the simplest finite simple group PSL(2, p), with p a prime ≥ 5, and setting γ to be the parabolic transformation $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. This enabled him to prove that the Kinoshita–Terasaka knot and the Conway knot are different. His proof also showed that none of them is amphicheiral. He then considered parabolic representations of knot groups to PSL(2, $\mathbb{C}$), and this led him to the discovery of the complete hyperbolic structure of the Figure-eight knot complement in [271].

Hartley [123] realized that one can apply this method to the problem of identifying noninvertible knots, as follows. Suppose no automorphism of Γ maps γ to γ^{-1}. Then the set $R(G(K),\Gamma,\gamma)$ is possibly different from the set $R(G(K),\Gamma,\gamma^{-1})$, and there is a chance to show noninvertibility of K by comparing the homology invariants associated with $\phi \in R(G(K),\Gamma,\gamma)$ with those associated with $\phi' \in R(G(K),\Gamma,\gamma^{-1})$. Hartley showed that this method is quite effective: he completely determined the 36 non-invertible knots up to 10 crossings claimed by Conway to be noninvertible.

A variation of the method is to consider the subset $R_t(G(K), S_n, \gamma)$ of $R(G(K), S_n, \gamma)$ consisting of the transitive representations, where n is a finite positive integer. By virtue of the development of computers, this turns out to be an extremely efficient method for distinguishing knots. In fact, Thistlethwaite [299] succeeded in distinguishing knots up to 11 crossings, and later the same method was applied successfully to knots up to 16 crossings in [144].

3.3 The Geometric Decomposition of Knot Exteriors

The purpose of this section is to explain the geometric decompositions of knot exteriors into Seifert pieces and hyperbolic pieces, obtained as a special case of the Thurston's geometrization theorem of Haken manifolds.

We recall (i) the prime decomposition theorem of general compact orientable 3-manifolds, (ii) the torus decomposition theorem of compact irreducible orientable 3-manifolds, (iii) the eight homogeneous 3-dimensional geometries, and (iv) the geometrization conjecture, which was finally established by Perelman. In the final subsection, we give detailed exposition of the geometric decompositions of knot exteriors.

For standard facts in 3-manifold theory, see the short note Hatcher [125] and the textbooks Hempel [127], Jaco [148], Jaco–Shalen [149], Johannson [150] and

Schultens [286]. For an introduction to geometric structures, see the surveys Scott [287] and Bonahon [40] and the textbook Martelli [194].

3.3.1 Prime Decomposition of 3-Manifolds

In this subsection, we recall the canonical decomposition of compact orientable 3-manifolds by 2-spheres. Let M be a compact connected orientable 3-manifold. A 2-sphere in M is *essential* if it does not bound a 3-ball in M. M is *irreducible* if it contains no essential 2-sphere. Suppose M is not irreducible, and let S be an essential 2-sphere in M. If S is *separating* (i.e., $M - S$ consists of two components), then M is the *connected sum* $M_1 \# M_2$ of the two compact orientable 3-manifolds M_1 and M_2, which are obtained from the closures of the components by capping off the resulting sphere boundaries by adding 3-balls. If S is *non-separating* (i.e., $M - S$ is connected), then M is expressed as the connected sum $(S^2 \times S^1)\# M'$ of $S^2 \times S^1$ with some compact orientable 3-manifold M' (possibly S^3).

M is *prime* if whenever $M = M_1 \# M_2$ we have $M_i \cong S^3$ for $i = 1$ or 2. Then we have the following Kneser–Milnor unique prime decomposition theorem [170, 211] (cf. [127]).

Theorem 3.3.1 (Unique Prime Decomposition of Compact Orientable 3-Manifolds) *Any compact orientable 3-manifold admits a decomposition $M = P_1 \# \cdots \# P_n$ into prime manifolds $\{P_i\}$. Moreover, the prime factors $\{P_i\}$ are uniquely determined by M, up to change of the indices.*

3.3.2 Torus Decomposition of Irreducible 3-Manifolds

In this subsection, we explain the torus decomposition theorem for compact orientable irreducible 3-manifolds. Torus decomposition is a simple version of more intricate JSJ (Jaco–Shalen–Johannson) decomposition, in which decompositions along annuli are also involved. The JSJ decomposition theory grew out of the study to understand homotopy equivalences among 3-manifolds, and its simplified version, the torus decomposition, turned out to be a complete obstruction for the hyperbolization of a 3-manifold.

Throughout this subsection, Σ denotes a compact orientable surface in M which is *properly embedded* in M, i.e., $\Sigma \cap M = \partial\Sigma$. We also assume that $\Sigma \not\cong S^2$. Then Σ is *incompressible* in M if for any disk D in M such that $D \cap \Sigma = \partial\Sigma$, the simple loop ∂D bounds a disk in Σ. By the loop theorem (see [127]), this is equivalent to the algebraic condition that the homomorphism $j_* : \pi_1(\Sigma) \to \pi_1(M)$ induced by the inclusion is injective.

M is *Haken* if it is irreducible and contains a properly embedded compact orientable surface which is incompressible.

A surface Σ in M is *essential* if it is incompressible and is not ∂*-parallel*, i.e., Σ does not cut off a 3-manifold in M homeomorphic to $\Sigma \times I$. M is *atoroidal* if it does not contain an essential torus.

M is a *Seifert fibered space* if it is expressed as a union of disjoint circles, in a particular way. The quotient of M obtained by collapsing each fiber into a point has the structure of a 2-dimensional orbifold, and is called the *base orbifold*. If M admits a smooth S^1 action without a fixed point (i.e., the stabilizer of any point is not the whole group S^1), then M is a Seifert fibered space whose base orbifold is the orbit space M/S^1. Seifert fibered spaces are regarded as S^1-bundles over 2-dimensional orbifolds, and are completely described by the *Seifert invariants*. See [127, 287] for details.

Now we state the torus decomposition theorem, which is a simplified version of the JSJ decomposition theorem due to Jaco and Shalen [149] and Johannson [150]. (See [241] and [66] for an alternative proof, and see [125] for a simple proof of the torus decomposition theorem.)

Theorem 3.3.2 (Torus Decomposition Theorem) *For a compact orientable irreducible* 3*-manifold M, there is a unique (up to isotopy) family $\mathcal{T}$ of disjoint essential tori, satisfying the following properties.*

(a) Each closed up component of $M - \mathcal{T}$ is either a Seifert fibered space or atoroidal.
(b) If any component of $\mathcal{T}$ is deleted, Property (a) fails.

In the above theorem, by a *closed up component* $M - \mathcal{T}$, we mean the closure of a component of $M - N(\mathcal{T})$, where $N(\mathcal{T})$ is a regular neighborhood of $\mathcal{T}$. The subsurface $\mathcal{T}$ is called the *characteristic toric family* of M, and each closed up component of $M - \mathcal{T}$ is called a *JSJ piece* of M.

It should be noted that the family $\mathcal{T}$ is not only unique up to homeomorphism but also unique up to *isotopy*. This forms a sharp contrast to the fact that in the prime decomposition theorem, the family of the splitting spheres is not unique even up to homeomorphisms. (It only says that the resulting prime manifolds are unique up to homeomorphisms.)

3.3.3 The Geometrization Conjecture of Thurston

Thurston's geometrization conjecture says that any compact orientable irreducible 3-manifold has a canonical splitting, by tori, into pieces which admit one of the following eight homogeneous geometries.

- The spaces of constant curvature, $\mathbb{S}^3$, $\mathbb{E}^3$ and $\mathbb{H}^3$;
- The product spaces $\mathbb{S}^2 \times \mathbb{E}^1$ and $\mathbb{H}^2 \times \mathbb{E}^1$; and
- The 3-dimensional Lie groups Nil, Sol, and $\widetilde{\mathrm{SL}}_2(\mathbb{R})$.

Here a compact connected orientable 3-manifold M is *geometric* if either it is a 3-ball or its interior can be presented as the quotient $\operatorname{int} M = X/\Gamma$ of one of the above homogeneous spaces, X, by a discrete group Γ of isometries acting freely and discontinuously on X. If $X = \mathrm{Sol}$ then M is a bundle over S^1 or the 1-dimensional orbifold $S^1/(z \sim \bar{z})$ with torus fiber; if X is neither Sol nor $\mathbb{H}^3$, then M is a Seifert fibered space and it is completely described by Seifert invariants. Conversely any Seifert fibered space admits one of the 6 remaining geometric. See the nice expositions [40, 287] for details.

Thus we have a complete topological classification of the geometric manifolds with X geometry for $X \neq \mathbb{H}^3$, and the study of hyperbolic manifolds forms the crucial part in 3-manifold theory and knot theory.

Thurston proposed the following geometrization theorem as a conjecture, which says that the torus decomposition gives a complete obstruction for a compact orientable 3-manifold to be hyperbolic. Here M is said to be *hyperbolic* if its interior can be presented as the quotient $\mathrm{int}\, M = \mathbb{H}^3/\Gamma$ by a discrete torsion-free group Γ of isometries of $\mathbb{H}^3$ (cf. Sect. 3.5). Thurston proved the conjecture for various cases, including the case when M is Haken, and the whole conjecture was finally proved by Perelman [253–255]. (See [201, 202, 218] for a survey, and see [23, 59, 169, 223, 224] for detailed expositions.)

Theorem 3.3.3 (Geometrization Theorem) *Let M be a connected irreducible atoroidal compact orientable 3-manifold. Then M is either a Seifert fibered space or a hyperbolic manifold.*

By combining Theorems 3.3.2 and 3.3.3, we obtain the following geometric decomposition theorem.

Theorem 3.3.4 (Geometric Decomposition Theorem) *For a compact orientable irreducible 3-manifold M, there is a unique (up to isotopy) family $\mathcal{T}$ of disjoint essential tori satisfying the following properties.*

(a) Each closed up component of $M - \mathcal{T}$ is either a Seifert fibered space or a hyperbolic manifold.
(b) If any component of $\mathcal{T}$ is deleted, Property (a) fails.

In the above theorem, a closed up component of $M - \mathcal{T}$ is called a *Seifert piece* or a *hyperbolic piece* according to whether it is a Seifert fibered space or a hyperbolic manifold.

3.3.4 Geometric Decompositions of Knot Exteriors

We describe a consequence for knot exteriors of the torus decomposition theorem and the geometrization theorem described in the previous subsection. Let K be a knot and consider its exterior $E(K)$. Then $E(K)$ is irreducible by the Schönflies theorem. Moreover, $E(K)$ is Haken, because a minimal genus Seifert surface is an incompressible surface in $E(K)$.

Theorem 3.3.5 (The Geometric Decomposition of Knot Exteriors) *Given a knot K in S^3, there is a unique (up to isotopy) compact subsurface $\mathcal{T}$ in the interior $E(K)$ satisfying the following properties.*

(a) Each component of $\mathcal{T}$ is an essential torus.
(b) Each closed up component of $E(K) - \mathcal{T}$ is either a Seifert fibered space or a hyperbolic manifold of finite volume.
(c) If any component of $\mathcal{T}$ is deleted, Property (b) fails.

We call the JSJ piece of $E(K)$ containing $\partial E(K)$ the *root JSJ piece* . The JSJ-decomposition is intimately related to Schubert's satellite operation [284]. To see this, assume that $\mathcal{T} \neq \emptyset$ and pick a component T of $\mathcal{T}$. Then T bounds a solid torus $V = S^1 \times D^2$ in S^3, which satisfies the following conditions.

(1) The core $k := S^1 \times 0$ of V forms a nontrivial knot in S^3.
(2) K is contained in V geometrically essentially, i.e., there is no 3-ball B such that $K \subset B \subset V$. Moreover, K is not isotopic in V to the core k of V.

Thus K is a *satellite knot* of the *companion knot* $k \subset S^3$ with *pattern* (V, K) (cf. Fig. 3.6 and [275, Section 4.D]).

It should be noted that composition of knots is a special case of the satellite operation. In fact, the composite knot $K_1 \# K_2$ is a satellite of K_1 with pattern $(S^3 - \operatorname{int} N(\mu_2), K_2)$ where μ_2 is a meridian of K_2. It is also a satellite of K_2 with pattern $(S^3 - \operatorname{int} N(\mu_1), K_1)$.

It turns out that JSJ pieces of knot exteriors are expressed as link exteriors (Theorem 3.3.7). A *link* L is a smoothly (or piecewise-linearly) embedded disjoint union of circles in S^3, namely L is a closed 1-submanifold of S^3. Thus a knot is a link of 1 component. A link of μ-components is called a *μ-component trivial link* if it bounds μ disjoint disks in S^3, and we denote it by O_μ. The *exterior* of a link L is defined by $E(L) := S^3 - \operatorname{int} N(L)$, where $N(L)$ is a regular neighborhood of L. The links in the example below play a key role in torus decompositions of knot exteriors.

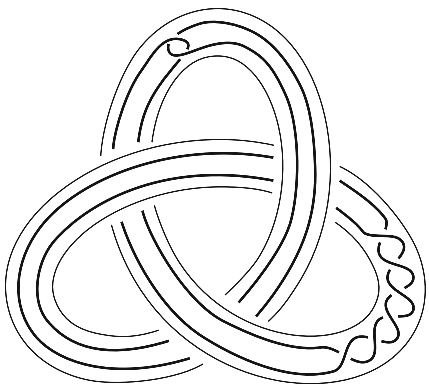


Fig. 3.6 A *Whitehead double* of a trefoil knot is a satellite knot whose companion knot is a trefoil knot and whose pattern knot is represented by the Whitehead link

Example 3.3.6

(1) The $\mu + 1$-component *key chain link* $H_{\mu+1} = K_0 \cup O_\mu$ is a union of the μ-component trivial link O_μ and the trivial knot K_0 which intersects each member of μ disjoint links bounded by O_μ. Then $E(H_{\mu+1})$ is homeomorphic to (μ holed disk) $\times S^1$, and is called a *composing space*. If $E(H_{\mu+1})$ is the root JSJ piece of a knot exterior $E(K)$, then K is a connected sum of μ prime knots.

(2) For a pair (p, q) of relatively prime integers, the (p, q)*-torus knot* $K_{p,q}$ is defined by

$$K_{p,q} := \{(z_1, z_2) \in S^3 \mid z_1^p + z_2^q = 0\}.$$

$K_{p,q}$ is a regular orbit of the circle action on S^3 given by

$$\omega \cdot (z_1, z_2) = (\omega^q z_1, \omega^p z_2) \quad (\omega \in S^1 \subset \mathbb{C}).$$

Thus $E(K_{p,q})$ is a Seifert fibered space. $K_{p,q}$ is contained in the torus

$$T := \{(z_1, z_2) \in S^3 \mid |z_1|^p = |z_2|^q\},$$

and it wraps q times in the z_1 direction and p times in the z_2 direction. The annulus $A := T \cap E(K)$ divides $E(K_{p,q})$ into two solid tori. By van-Kampen's theorem in this setting, we see that

$$G(K_{p,q}) = \langle a, b \mid a^p = b^q \rangle.$$

The cyclic subgroup generated by $a^p = b^q$ forms the infinite cyclic center of $G(K_{p,q})$. Moreover, a knot K is a torus knot if and only if $G(K)$ has a nontrivial center. $K_{p,q}$ is nontrivial if and only if both p and q have absolute value ≥ 2. If $E(K_{p,q})$ is a JSJ piece of a knot exterior $E(K)$, then K is a satellite of the torus knot $K_{p,q}$.

(3) For a pair (p, q) of relatively prime integers with $p \geq 2$, the (p, q)*-Seifert link* $C_{p,q}$ is defined by

$$C_{p,q} := K_0 \cup K_{p,q} \quad \text{with } K_0 = \{(z_1, z_2) \in S^3 \mid z_2 = 0\}.$$

If $C_{p,q}$ is the root JSJ piece of a knot exterior $E(K)$, then K is the (p, q)*-cable* of some nontrivial knot.

We have the following characterization of the torus decompositions of knot exteriors.

Theorem 3.3.7

(1) *A compact orientable* 3*-manifold M is a JSJ piece of $E(K)$ for some nontrivial knot K in S^3 if and only if $M \cong E(L)$ for some link L in S^3, which is the union*

of a knot K_0 *and a trivial link* O_μ *(with* μ *possibly* 0*), such that* $E(L)$ *is either (i) hyperbolic or (ii) a Seifert fibered space homeomorphic to a composing space, a nontrivial torus knot exterior, or a cable space.*

(2) *Let* K *be a nontrivial knot in* S^3*, and let* $\mathcal{T}$ *be a union of disjoint essential tori in* $E(K)$*, satisfying the following conditions.*
   (i) *Each closed up component of* $E(K) - \mathcal{T}$ *is homeomorphic to a link exterior* $E(L)$ *which satisfies the condition in (1).*
   (ii) *There does not exist a pair of adjacent closed up components of* $E(K) - \mathcal{T}$*, both of which are composing spaces.*

   Then $\mathcal{T}$ *is the characteristic toric family of* $E(K)$*.*

The way JSJ pieces fit together in $E(K)$ is recorded by the *companionship tree*, defined as follows: The vertices correspond to the JSJ pieces, and the edges correspond to the components of $\mathcal{T}$, where if an edge corresponds to a component T of $\mathcal{T}$, it joins the vertices corresponding to the two JSJ pieces containing T as a boundary component. Since $H_1(S^3) = 0$, this graph is a tree. For a more detailed description of torus decompositions, see [42, Chapter 2] and [51].

3.4 The Orbifold Theorem and the Bonahon–Siebenmann Decomposition of Links

In [300, Chapter 13], Thurston initiated the systematic study of orbifolds, namely quotients of spaces by properly discontinuous group actions which are not necessarily free. In 1978, he announced the orbifold theorem, the geometrization theorem of 3-orbifolds which have non-empty 1-dimensional singular set. Every link $L = \cup_j K_j$ determines an infinite family of orbifolds, by regarding each component K_j as the singular locus of cone angle $2\pi/n_j$ for some $n_j \geq 2$. The case when $n_j = 2$ for every j is particularly important, and the Bonahon–Siebenmann decomposition theory of links is essentially the decomposition theory of such orbifolds. Their theory is intimately related with Conway's ingenious analysis of link diagrams, and gives us a nice method for understanding links directly from their diagrams. In particular, it gives a complete classification of the "algebraic links", which implies, for example, that the Kinoshita–Terasaka knot and the Conway knot are different and that they admit no symmetry.

The purpose of this section is to recall the orbifold theorem and its impact on knot theory. To be precise, we will give surveys of (i) the Bonahon–Siebenmann decomposition theory, (ii) the classification of 2-bridge links, (iii) the orbifold theorem, and (iv) application of the orbifold theorem to the study of branched cyclic coverings.

3.4.1 The Bonahon–Siebenmann Decompositions for Simple Links

By the geometric decomposition Theorems 3.3.5 and 3.3.7 of knot exteriors, the classification of knots is reduced to that of the links whose exteriors have trivial torus decompositions. Deriving from Montesinos' work [214, 215] on double branched coverings of links and Thurston's work on 3-dimensional orbifolds, Bonahon and Siebenmann established a new decomposition theorem for such links. This is essentially a $\mathbb{Z}/2\mathbb{Z}$-equivariant JSJ decomposition theory, applied to the double branched coverings of links.

To explain their results, we introduce a few definitions. A link L in S^3 is *splittable* if there is a 2-sphere S in S^3 which separates the components of L. L is *unsplittable* if it is not splittable. This is equivalent to the condition that $E(L)$ is irreducible. L is *simple for Schubert* if $E(L)$ is irreducible and atoroidal. If L is simple for Schubert, then the JSJ decomposition of $E(L)$ is trivial. The converse also holds for knots, but not for links. For example, the key-chain link $H_{\mu+1}$ is not simple for Schubert, but the torus decomposition of $E(H_{\mu+1})$ is trivial.

Let (M, L) be a pair consisting of a compact orientable 3-manifold and a proper 1-submanifold L in M. A *Conway sphere* in (M, L) is a 2-sphere Σ in int M or ∂M which meets L transversely in 4 points. A Conway sphere Σ is said to be *pairwise compressible* if there is a disk D in $M - L$ such that $D \cap \Sigma = \partial D$ does not bound a disk in $\Sigma - L$. Otherwise, Σ is said to be *pairwise incompressible*. Two Conway spheres Σ and Σ' in M are said to be *pairwise parallel*, if there is a closed up component N of $M - (\Sigma \cup \Sigma')$ bounded by Σ and Σ' such that $(N, N \cap L) \cong (\Sigma, \Sigma \cap L) \times [0, 1]$. A Conway sphere is *essential* if it is pairwise incompressible and is not pairwise parallel to a boundary component. (M, L) is *simple for Conway* if it does not contain an essential Conway sphere.

A *trivial tangle* is a pair (B^3, t), where t is a union of two arcs properly embedded arcs in B^3 which is parallel to a pair of disjoint arc in ∂B^3. A *rational tangle* is a trivial tangle (B^3, t) which is endowed with an identification of $\partial(B^3, t)$ with the Conway sphere standardly embedded in $\mathbb{R}^3 \subset S^3$. A rational tangle, up to the natural equivalence relation, is determined uniquely by its *slope* as illustrated in Fig. 3.7. (See [63] for the original definition, and see [42, Chapter 8.1] or [69, Section 8.6] for detailed exposition.)

A *Montesinos pair* is a pair (M, L) which is built from a *hollow Montesinos pair* or a *hollow Montesinos pair with a ring* in Fig. 3.8 by plugging some of the holes with rational tangles of finite slope.

Bonahon and Siebenmann established the following decomposition theorem [42, Theorem 3.4].

Theorem 3.4.1 *For a link L in S^3 that is simple for Schubert, there is a unique (up to isotopy respecting L) compact subsurface $\mathcal{G} \subset S^3$ satisfying the following property.*

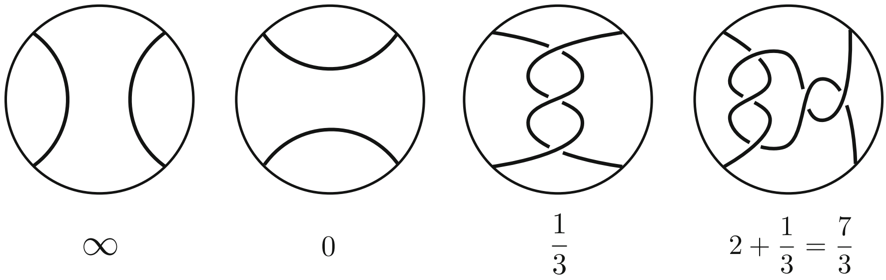


Fig. 3.7 The pair of arcs forming a trivial tangle is parallel to a pair of arcs on the boundary Conway sphere. If we identify the Conway sphere with the quotient of $\mathbb{R}^2$ by the group generated by π-rotations around the lattice points, then the inverse image of the pair of arcs in $\mathbb{R}^2$ forms a family of mutually disjoint lines of rational slope passing through the lattice points. This slope is the *slope* of the rational tangle

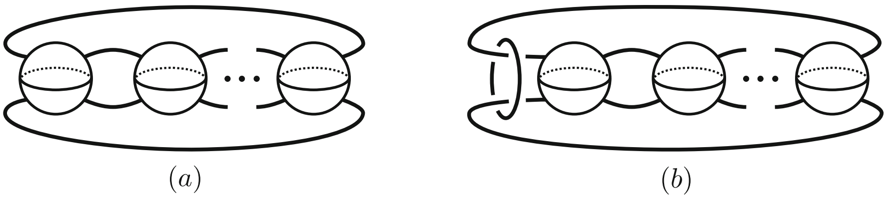


Fig. 3.8 (**a**) a hollow Montesinos pair, (**b**) a hollow Montesinos pair with a ring

(a) Each component of $\mathcal{G}$ is a pairwise incompressible Conway sphere.
(b) Each closed up component N of $S^3 - \mathcal{G}$ gives a pair $(N, N \cap L)$ that either is simple for Conway, or else is a Montesinos pair.
(c) If any component of $\mathcal{G}$ is deleted, Property (b) fails.

The above decomposition is called the *characteristic decomposition* (or the *Bonahon–Siebenmann decomposition*) of (S^3, L). The union of the Montesinos pairs is called the *algebraic part* or *arborescent part* of (S^3, L). The link (S^3, L) is called an *arborescent link* if its arborescent part is equal to the whole pair (S^3, L). This terminology comes from the fact that arborescent parts can be represented by weight planar trees. The classification the arborescent parts and links is given by [42, Part V].

Example 3.4.2 The Bonahon–Siebenmann decomposition of the Kinoshita–Terasaka knot and the Conway knot are given by the spheres in Fig. 3.5 (cf. [215, Sections 4.1 and 4.2]). This fact gives an intuitive proof of the inequivalence of these two knots. It also shows that both knots are arborescent.

For a link L in S^3, let $p : M_2(L) \to S^3$ be the double branched covering of S^3 branched over L, and let τ be the covering involution. Then the Bonahon–Siebenmann decomposition of L can be regarded as a $\mathbb{Z}_2$-equivariant version of the torus decomposition of $M_2(L)$ for the following reasons:

- The inverse image of an essential Conway sphere of (S^3, L) is an essential torus in $M_2(L)$.
- Let $(N, N \cap L)$ be a piece of the Bonahon–Siebenmann decomposition of (S^3, L) which is a Montesinos pair. Then the inverse image $p^{-1}(N)$ is a Seifert fibered space, where the base orbifold is orientable or non-orientable according to whether $(N, N \cap L)$ is obtained from a hollow Montesinos pair or that with a ring (see [214, 216]). Moreover, the covering involution τ preserves the Seifert fibration of $p^{-1}(N)$. The image of its fiber in S^3 is either a circle disjoint from L or an interval with endpoints in L.
- The above fact implies that the inverse image in $M_2(L)$ of the arborescent part of (S^3, L) is a graph manifold (cf. Waldhausen [316]). In particular, if (S^3, L) is an arborescent link then $M_2(L)$ is a graph manifold.
- For each piece $(N, N \cap L)$ of the Bonahon–Siebenmann decomposition of (S^3, L) which is not a Montesinos pair, the inverse image $p^{-1}(N)$ is irreducible and atoroidal. Moreover, by the orbifold theorem (Theorem 3.4.4) explained later in this section, $p^{-1}(N)$ admits a complete hyperbolic structure of finite volume, with respect to which $\tau|_{p^{-1}(N)}$ is an isometry.

We note that the Bonahan-Siebenmann decomposition is intimately related with Conway's ingenious analysis of knot diagrams, which in turn is based on Kirkmann's idea from the nineteenth century (see [144] for the history). In fact, it reveals that Conway's notation for a link diagram is not merely a convenient tool for describing diagrams but also contains geometric information of the link represented by a diagram. This is certainly the case for algebraic parts of the link. As shown in [42, Theorems 1.4 and 6.11], Conway's notation for non-algebraic parts also has geometric information under certain conditions.

3.4.2 2-Bridge Links

In this subsection, we introduce 2-bridge links, which form a very special but important class of links. For a rational number $r \in \mathbb{Q} \cup \{1/0\}$, the *2-bridge link, $K(r)$, of slope r* is defined as the "sum" of the rational tangles of slopes r and $1/0$. To be precise, it is obtained from the rational tangles, $(B^3, t(r))$ and $(B^3, t(1/0))$, of slopes r and $1/0$, respectively, by glueing $(B^3, t(r))$ and $(-B^3, t(1/0))$ along the boundaries via the identity map. (Note that the boundaries of rational tangles are identified with the Conway sphere standardly embedded in $\mathbb{R}^3$.) If $r = q/p$ where $p \geq 0$ and q are relatively prime integers, then $K(r)$ is a knot or a two-component link according to whether p is odd or even. The following classification theorem was proved by Schubert [285], by establishing the uniqueness up to isotopy of *2-bridge spheres* (2-spheres which divide $K(r)$ into two trivial tangles).

- Two 2-bridge links $K(q/p)$ and $K(q'/p')$ are isotopic if and only if $p = p'$ and either $q \equiv q' \pmod p$ or $qq' \equiv 1 \pmod p$. They are homeomorphic if and only if $p = p'$ and either $q \equiv \pm q' \pmod p$ or $qq' \equiv \pm 1 \pmod p$.

The double branched covering, $M_2(K(q/p))$, of S^3 branched over $K(q/p)$ is the lens space $L(p,q)$, and the above classification of 2-bridge links can be also deduced from the classification of lens spaces, which in turn was established by Reidemeister [265], using the Reidemeister torsion. Moreover, the following characterization of 2-bridge link was obtained by Hodgson and Rubinstein [139], by classifying involutions on lens spaces with 1-dimensional fixed point sets.

- A link L in S^3 is a 2-bridge link if and only if the double branched covering $M_2(L)$ is a lens space.

The result of [139] is a special but important case of the orbifold theorem (Theorem 3.4.4) explained later in this section. They also proved the uniqueness up to isotopy of genus 1 Heegaard surfaces of lens spaces, which in turn gives a purely topological proof of the classification of lens spaces.

Thurston's uniformization theorem for Haken manifold (cf. Theorems 3.3.3), together with an analysis of incompressible surfaces in the exterior of 2-bridge links, imply the following (cf. [272, p.102], [171, Lemmas 4.4], [126]).

- The 2-bridge link $K(q/p)$ is hyperbolic if and only if $q \not\equiv \pm 1 \pmod p$.

3.4.3 Bonahon–Siebenmann Decompositions and π-Orbifolds

For a link L in S^3, the pair (S^3, L) is homeomorphic to the quotient $(M_2(L), \mathrm{Fix}(\tau))/\tau$, where $M_2(L)$ is the double branched covering of S^3 branched over L and τ is the covering involution. This means that the good 3-orbifold $\mathcal{O}(L) := M_2(L)/\tau$ has S^3 as underlying space and L as singular set, and each component of the singular set L has cone angle π. The Bonahon–Siebenmann decomposition is regarded as the torus decomposition of this orbifold.

Recall that an *n-orbifold* is a metrizable topological space $\mathcal{O}$ locally modeled on the quotient of $\mathbb{R}^n$ by a finite subgroup G of the orthogonal group $\mathrm{O}(n)$. If a point $x \in \mathcal{O}$ corresponds to the image of the origin of $\mathbb{R}^n$, then the finite group G is called the *local group* at x, and is denoted by G_x. If G_x is trivial, x is *regular*, otherwise x is *singular*. The *singular locus* is the subset, $\Sigma_{\mathcal{O}}$, of $\mathcal{O}$ consisting of the singular points. When G_x is the cyclic group generated by a $2\pi/m$-rotation around the codimension 2 subspace $\mathbb{R}^{n-2} \times \{0\}$, we say that the point x (and the stratum of the singular set containing x) has *cone angle* $2\pi/m$ or *index* m.

A quotient space $\mathcal{O} := X/\Gamma$, where X is a smooth n-manifold and Γ is a smooth properly discontinuous action, is an n-orbifold, and its singular locus is the image of the subspace of X consisting of points with nontrivial stabilizer. If Γ is a finite group, such an orbifold is called a *good orbifold*. The *orbifold fundamental group*

$\pi_1^{\text{orb}}(\mathcal{O})$ of $\mathcal{O}$ is defined as the group consisting of all lifts of Γ to the universal covering space $\tilde{X}$ of X. Thus we have the following exact sequence.

$$1 \to \pi_1(X) \to \pi_1^{\text{orb}}(\mathcal{O}) \to \Gamma \to 1$$

For a link (S^3, L), the orbifold $\mathcal{O}(L) := (M_2(L), \text{Fix}(\tau))/\tau$ is called the *π-orbifold* associated with L. The orbifold fundamental group $\pi_1^{\text{orb}}(\mathcal{O}(L))$ is called the *π-orbifold group* of L. It is calculated from the link group $G(L) = \pi_1(S^3 - L)$ and a set of meridians $\{\mu_1, \ldots, \mu_m\}$ as follows:

$$\pi_1^{\text{orb}}(\mathcal{O}(L)) = G(L)/\langle\langle \mu_1^2, \ldots, \mu_m^2 \rangle\rangle.$$

Here m is the number of components of L, and μ_j is a meridian of the j-th component of L. By using the orbifold theorem explained in the next subsection, Boileau and Zimmermann [34] proved the following theorem which shows that $\pi_1^{\text{orb}}(\mathcal{O}(L))$ is a very strong invariant for links.

Theorem 3.4.2 *Let L be a prime unsplittable link in S^3 such that $\pi_1^{\text{orb}}(\mathcal{O}(L))$ is infinite. Then the following hold.*

(1) For any link L' in S^3, the pairs (S^3, L) and (S^3, L') are homeomorphic if and only if their π-orbifold groups $\pi_1^{\text{orb}}(\mathcal{O}(L))$ and $\pi_1^{\text{orb}}(\mathcal{O}(L'))$ are isomorphic.

(2) The natural homomorphism from the symmetry group $\text{Sym}(S^3, L)$ *to the outer-automorphism group* $\text{Out}(\pi_1^{\text{orb}}(\mathcal{O}(L)))$ *is an isomorphism.*

Here the *symmetry group* $\text{Sym}(S^3, L)$ is the group of diffeomorphisms of the pair (S^3, L) up to isotopy. It should be noted that the problem of determining the symmetry group of a knot is a refinement of the fundamental problem of determining whether the knot is chiral/invertible.

Using the above theorem, Boileau and Zimmermann [33] determined the symmetry groups of all non-elliptic Montesinos links, i.e., the Montesinos links with infinite π-orbifold groups. (The symmetry groups of elliptic Montesinos links were determined by [279] using the orbifold theorem.) This result may be regarded as a broad extension of Trotter's proof [309] of non-invertibility of the pretzel knot $P(p, q, r)$ with $|p|, |q|, |r|$ distinct odd integers ≥ 3. Trotter's proof is based on the fact that $\pi_1^{\text{orb}}(\mathcal{O}(P(p, q, r)))$ is an extension of the hyperbolic triangular reflection group

$$[p, q, r] = \langle x, y, z \mid x^2, y^2, z^2, (xy)^p, (yz)^q, (zx)^r \rangle$$

by the infinite cyclic group, which in turn is a consequence of the fact that the π-orbifold $\mathcal{O}(P(p, q, r))$ is a *Seifert fibered orbifold* over the 2-dimensional hyperbolic orbifold $\mathbb{H}^2/[p, q, r]$.

The symmetry groups of the arborescent links are completely determined by Bonahon and Siebenmann in [42]. In particular, this implies that the symmetry groups of the Kinoshita–Terasaka knot and the Conway knot are trivial, and so they are chiral and noninvertible. The knot 8_{17} is also arborescent, and its symmetry group is the order 2 cyclic group, generated by an orientation-reversing involution representing the negative-amphicheirality of the knot. This is the Bonahon-Siebenmann's proof of the non-invertibility of 8_{17} (cf. Sect. 3.2.6).

3.4.4 The Orbifold Theorem and the Smith Conjecture

Many of the concepts for 3-manifolds, such as irreducibility, atoroidality and Seifert fibrations, have natural generalization for 3-orbifolds, and a characteristic splitting (torus decomposition) theorem was established by Bonahon and Siebenmann [41] (cf. [35, 40]). The characteristic splitting Theorem 3.4.1 for links is a special case of the general splitting theorem, though the detailed analysis for the algebraic parts and the application to knot theory in [42] cannot be found in [41].

Bonahon-Siebenmann's theory forms the first step towards the proof of the following geometrization theorem for orbifolds, which was announced by Thurston [301], and finally proved by Boileau, Leeb and Porti [36] (see also Cooper–Hodgson–Kerckhoff [65] and Boileau–Porti [32] for an earlier account, and Dinkelbach–Leeb [79] for the generalization to non-orientable orbifolds using the equivariant Ricci flow).

Theorem 3.4.4 (Orbifold Theorem) *Every compact orientable good* 3*-orbifold with nonempty singular set has a canonical splitting by spherical* 2*-dimensional suborbifolds and toric* 2*-dimensional suborbifolds into geometric* 3*-orbifolds.*

Here a 3-orbifold $\mathcal{O}$ is *geometric* if either it is the quotient of a ball by an orthogonal action, or its interior has one of the eight Thurston geometries, namely $\mathcal{O} = X/\Gamma$, where X is one of the eight Thurston's geometries and Γ is a discrete subgroup of $\mathrm{Isom}(X)$. (If X is different from the constant curvature spaces $\mathbb{H}^3$, $\mathbb{E}^3$ and $\mathbb{S}^3$, then there is no canonical metric on X, however, it admits a family of natural metrics for which $\mathrm{Isom}(X)$ are identical. See the beautiful surveys [40, 287].)

The orbifold theorem was first announced as the following symmetry theorem concerning finite group actions on 3-manifolds.

Theorem 3.4.5 (Symmetry Theorem) *Let M be a compact orientable irreducible* 3*-manifold. Suppose M admits an action by a finite group G of orientation-preserving diffeomorphisms such that some non-trivial element has a fixed point set of dimension one. Then M admits a geometric decomposition preserved by the group action.*

This theorem poses a very strong restriction on finite group actions on knots (see [28, 187, 278]). In particular, it includes, as a special case, the following positive answer to the Smith conjecture.

Theorem 3.4.6 (The Smith Conjecture) *If $h : S^3 \to S^3$ is an orientation-preserving periodic diffeomorphism with non-empty fixed point set, then h is smoothly conjugate to an orthogonal diffeomorphism. In particular,* $\mathrm{Fix}(h)$ *is the trivial knot.*

The proof of this conjecture recorded in [219] may be regarded as the first major impact of Thurston's uniformization theorem for Haken manifolds, and it was established using the uniformization theorem, the equivariant loop theorem by Meeks–Yau [205], and a refinement of Bass–Serre theory [289].

In Theorems 3.4.5 and 3.4.6, the smoothness of the action is essential. In fact there is an orientation-preserving periodic homeomorphism h of S^3 which has a wild knot as the fixed point set; in particular, the cyclic action generated by h is not topologically conjugate to an orthogonal action. It is this phenomena that lead Shin'ichi Kinoshita and Hidetaka Terasaka, the founders of knot theory in Japan, into knot theory. It is an amazing coincidence that Terasaka published an introductory book [298] to non-Euclidean geometry for the general public in 1977, around the time Thurston started the series of lectures on the geometry and topology of 3-manifolds.

3.4.5 Branched Cyclic Coverings of Knots

In Sect. 3.4.3, we explained the important role of the double branched coverings of knots and links. Not only the double branched covering but also the cyclic branched covering has attracted keen attention of various mathematicians, because the latter gives a bridge between knot theory and 3-manifold theory and because of its special beauty. In this subsection, we review the impact of Thurston's work, in particular the orbifold theorem, on the study of branched cyclic coverings of knots.

For a knot K in S^3, let $M_n(K)$ be the n-fold cyclic branched covering of S^3 branched over K. We also call $M_n(K)$ the *n-fold cyclic branched covering of K*. Then we have the following natural question.

Problem 3.4.7 To what extent does the topological type of $M_n(K)$ determine K?

It should be noted that $M_n(K)$ inherits the orientation of the ambient space S^3, but it is independent of the orientation of the circle K. Namely $M_n(K) \cong M_n(-K)$ as oriented manifolds. Thus the precise meaning of the above question is as follows. To what extent does the topological type of the *oriented manifold* $M_n(K)$ determine the *isotopy type* of the *unoriented* knot K?

The positive solution of the Smith conjecture is essentially equivalent to the following partial answer to the above problem (see [219]).

Theorem 3.4.8 (Branched Covering Theorem) *A knot K in S^3 is trivial if and only if $M_n(K) \cong S^3$ for some $n \geq 2$.*

The orbifold theorem gives a very strong tool for the study of Problem 3.4.7. Before describing its influence, let us recall two classical constructions of pairs of knots sharing the same cyclic branched covering.

Construction 1 Let K be a non-invertible prime oriented knot. Then, by the unique prime factorization theorem, the knots $K\#K$ and $K\#(-K)$ are not isotopic as unoriented knots. However, they share the same n-fold cyclic branched covering for all $n \geq 2$, because:

$$M_n(K\#K) \cong M_n(K)\#M_n(K) \cong M_n(K)\#M_n(-K) \cong M_n(K\#(-K))$$

Construction 2 Let $L = K_1 \cup K_2$ be a 2-component link consisting of two trivial knots. For integers $n_1, n_2 \geq 2$ which are relatively prime to the linking number $lk(K_1, K_2)$, the inverse image $\tilde{K}_1$ of K_1 in $M_{n_1}(K_2) \cong S^3$ is a knot, and so is the inverse image $\tilde{K}_2$ of K_2 in $M_{n_2}(K_1) \cong S^3$. Moreover, both $M_{n_2}(\tilde{K}_1)$ and $M_{n_1}(\tilde{K}_2)$ are homeomorphic to the $(\mathbb{Z}/n_1\mathbb{Z}) \oplus (\mathbb{Z}/n_2\mathbb{Z})$-covering of S^3 branched over L, and hence they are homeomorphic. (There is an analogous construction by using a three-component link such that any 2-component sublink is a Hopf link (see [267, 0.2]).)

Now, we state an important consequence of the orbifold theorem (see [65]).

Theorem 3.4.9 *Let K be a hyperbolic knot in S^3, i.e., K is a knot which is neither a torus knot nor a satellite knot. Then $M_n(K)$ is hyperbolic for all $n \geq 3$, except for the 3-fold covering of the figure eight knot (which is a Euclidean manifold). Moreover, the covering transformation group acts on $M_n(K)$ by isometries.*

Remark 3.4.10 In the above theorem the assumption $n \geq 3$ is essential. In fact, if a hyperbolic knot contains an essential Conway sphere, Σ, then the inverse image, $\tilde{\Sigma}$, of Σ in $M_2(K)$ is an essential torus and hence $M_2(K)$ is non-hyperbolic even though K itself is hyperbolic. Moreover, every arborescent link has a graph manifold as double branched covering.

The hyperbolic Dehn surgery theorem implies that if n is sufficiently large, then the branch line forms the unique shortest closed geodesic in $M_n(K)$ (cf. Sect. 3.7.3). Using this fact, we can see that $M_n(K)$ for sufficiently large n determines the knot K. More generally, Kojima [172] proved the following theorem, which gives a positive answer to a question of Goldsmith [167, Problem 1.27].

Theorem 3.4.11 *For each prime knot K there exists a constant n_K, such that two prime knots K and K' are equivalent if their n-fold cyclic branched covers are homeomorphic for some $n > \max(n_K, n_{K'})$.*

We can reformulate Problem 3.4.7 as follows: For a given connected closed orientable 3-manifold M, in how many different ways can M occur as a cyclic branched covering of a knot in S^3? There are two basic cases: the case when M is a Seifert fibered space and the case when M is a hyperbolic manifold. (The general case can be treated by using the equivariant sphere theorem and torus decomposition [204] into Seifert fibered space and hyperbolic manifolds.)

When M is a Seifert fibered space, the covering transformation group, H, is fiber-preserving by [204] (when $\pi_1(M)$ is infinite) and by the orbifold theorem (when $\pi_1(M)$ is finite). If H reverses the fiber-orientation, then the quotient knot is a Montesinos knot whereas if H preserves the fiber-orientation then the quotient knot is a torus knot.

In the case where M is hyperbolic, we may assume, by the orbifold theorem, that H is a cyclic subgroup of the finite group $\mathrm{Isom}^+(M)$. The group H must be a *hyper-elliptic group*, namely H is a finite cyclic group such that $\mathrm{Fix}\, h$ is a circle for every non-trivial element $h \in H$, and $M/H \cong S^3$ (cf. [39, Definition 1]). Thus there is a one-to-one correspondence

$$\{\text{knots } K \text{ such that } M_n(K) \cong M \text{ for some } n \geq 2\}/\text{isotopy}$$

$$\leftrightarrow \{\text{hyper-elliptic subgroups of } \mathrm{Isom}^+(M)\}/\text{conjugacy}.$$

By Kojima's theorem [173], any finite group can be the full isometry group of a closed orientable hyperbolic 3-manifold. However, the geometric condition for a hyper-elliptic group, H, implies purely group theoretical conditions on H. For example, we can see by using the Smith conjecture (Theorem 3.4.6) that the normalizer of H in $\mathrm{Isom}^+(M)$ is a finite subgroup of the semi-direct product $(\mathbb{Z}/2\mathbb{Z}) \ltimes (\mathbb{Q}/\mathbb{Z} \oplus \mathbb{Q}/\mathbb{Z})$, where $\mathbb{Z}/2\mathbb{Z}$ acts on $\mathbb{Q}/\mathbb{Z} \oplus \mathbb{Q}/\mathbb{Z}$ as multiplication by -1 (see [39, Remark 3]). Thus we have a chance to apply finite group theory to the study of cyclic branched coverings. For example, if we are interested in the case when the degree n is a prime number p, then by Sylow's theorem, every hyper-elliptic subgroup of order p is conjugate to a cyclic subgroup of a single Sylow p-subgroup S_p of $\mathrm{Isom}^+(M)$. This interplay between the study of cyclic branched coverings and finite group theory was initiated by Reni and Zimmermann, and various interesting results were obtained, including the following.

- Reni–Zimmermann [267]: Let K and K' be two hyperbolic knots such that $M_n(K) \cong M_{n'}(K')$ for some $n, n' \geq 3$. Suppose further that n and n' have a common prime divisor $p > 2$. Then K and K' are related by Construction 2. In particular, if $n = n'$ is not a power of 2, then the same conclusion holds (cf. [328]).
- Paoluzzi [251]: A hyperbolic knot is determined by any three of its cyclic branched coverings of order ≥ 2. Indeed, two coverings suffice if their orders are not coprime.
- Boileau–Franchi–Mecchia–Paoluzzi–Zimmermann [39]: A closed hyperbolic 3-manifold is a cyclic branched covering of at most fifteen inequivalent knots in S^3.

A noteworthy aspect of the proof of the last result is the substantial use of finite group theory, in particular of the classification of finite simple groups.

For the double branched coverings, we have the following additional construction.

Construction 3 Let θ be a θ-curve in S^3, namely a spatial graph consisting of two vertices and three edges α_1, α_2 and α_3, each of which connects the two vertices. For $\{i, j, k\} = \{1, 2, 3\}$, suppose $A_k := \alpha_i \cup \alpha_j$ forms a trivial knot. Then the inverse image, K_k, of the arc α_k in $M_2(A_k) \cong S^3$ forms a (strongly invertible) knot, and $M_2(K_k)$ is identified with the $(\mathbb{Z}/2\mathbb{Z})^2$-covering of S^3 branched over θ. If A_k is a trivial knot for more than one $k \in \{1, 2, 3\}$, we obtain knots in S^3 sharing the same double branched coverings. (A similar construction is applied to embeddings of the 1-skeleton of the tetrahedron and the Kuratwski graph in S^3, which produce potentially distinct 4 and 9 knots, respectively, sharing the same double branched coverings (see [203]).)

A link L is said to be π*-hyperbolic* if $M_2(L)$ is hyperbolic. For double coverings of π-hyperbolic knots, the following results were obtained.

- Boileau–Flapan [29]: If K is a π-hyperbolic knot, then every knot K' which shares the same double branched covering with K is constructed by repeatedly applying Constructions 2 and 3.
- Reni [266]: There are at most nine different π-hyperbolic knots with the same double branched coverings. Mecchia–Reni [203] gave a more geometric proof to this estimate and proved that the same estimate holds for π-hyperbolic links.
- Kawauchi [162]: Reni's estimate is the best possible, i.e., there are nine mutually inequivalent π-hyperbolic knots K_i $(i = 1, \cdots, 9)$, in S^3 with the same double branched coverings.

In the proof of the second result, the study of the Sylow 2-subgroup of $\mathrm{Isom}^+(M)$ of a closed orientable hyperbolic 3-manifold holds a key. The third result was obtained by using Kawauchi's imitation theory, which yields, for a given (3, 1)-manifold pair (M, L), a family of (3, 1)-manifold pairs (M^*, L^*) which is "topologically similar" to (M, L). A key example in the theory is the Kinoshita–Terasaka knot, which is an *imitation* of the trivial knot. (This fact was first found by Nakanishi [236] and a beautiful generalization of this fact was given by Kanenobu [156].)

For the double branched covering of links which are not π-hyperbolic, we have the following additional construction. (See Paoluzzi [250] for further construction.)

Construction 4 (Mutation) Let Σ be an essential Conway sphere of a link L in S^3. Cut (S^3, L) along Σ and reglue by an orientation-preserving involution of $(\Sigma, \Sigma \cap L)$ whose fixed point set is disjoint from $\Sigma \cap L$. This process, called a *mutation*, results in a new link L' in S^3, called a *mutant* of L. A pair of links are called *mutants* if they are related by a sequence of mutations. It was proved by Viro [312, Theorem 1] that if L and L' are mutants then they share the same double branched coverings (cf. [161, Proposition 3.8.2]).

In [115], Greene studied the Heegaard Floer homology of the double branched coverings of alternating links, and proved that a reduced alternating link diagram is determined up to mutation by the Heegaard Floer homology of the double branched covering of the link. In particular, the following result follows.

- Two reduced alternating links L and L' share the same double branched covering if and only if L and L' are mutants.

He also proposes the mysterious conjecture: *if a pair of links have homeomorphic double branched coverings, then either both are alternating or both are non-alternating*.

3.5 Hyperbolic Manifolds and the Rigidity Theorem

In this section, we recall basic facts concerning hyperbolic manifolds and the Mostow–Prasad rigidity theorem for hyperbolic manifolds of finite volume and of dimension ≥ 3. The rigidity theorem has had tremendous influence on knot theory, because it guarantees that any geometric invariant of the hyperbolic structure of a hyperbolic knot complement is automatically a topological invariant of the knot complement.

For further information on hyperbolic geometry, see the textbooks Benedetti–Petronio [19], Ratcliffe [260], Matsuzaki–Taniguchi [197], Anderson [18] and Marden [191].

3.5.1 Hyperbolic Space

Let $\mathbb{H}^n$ be the hyperbolic n-space, i.e., the upper-half space

$$\mathbb{H}^n := \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_n > 0\}$$

in $\mathbb{R}^n$ equipped with the Riemannian metric

$$ds^2 = \frac{1}{x_n^2}(dx_1^2 + \cdots + dx_n^2).$$

$\mathbb{H}^n$ is the unique connected, simply connected, complete Riemannian manifold of constant sectional curvature -1. The isometry group $\mathrm{Isom}(\mathbb{H}^n)$ is a real Lie group and acts transitively on $\mathbb{H}^n$, and the stablizer of each point is identified with the orthogonal group $\mathrm{O}(n)$. If $n \geq 3$, the ideal boundary $\partial\mathbb{H}^n = (\mathbb{R}^{n-1} \times \{0\}) \cup \{\infty\}$ has a natural conformal structure, and the orientation-preserving isometry group $\mathrm{Isom}^+(\mathbb{H}^n)$ is identified with the group of conformal maps of $\partial\mathbb{H}^n$.

Let γ be a nontrivial element of $\mathrm{Isom}^+(\mathbb{H}^n)$. Then precisely one of the following holds.

(1) γ is *elliptic*, i.e., γ has a fixed point in $\mathbb{H}^n$.
(2) γ is *parabolic*, i.e., γ has a unique fixed point, x, in $\partial\mathbb{H}^n$, called the *parabolic fixed point*. Then γ preserves every *horoball*, H_x, centered at x. Here, if $x \neq \infty$,

then H_x is the intersection of a (closed) Euclidean ball with $\mathbb{H}^n$ which touches $\partial\mathbb{H}^n$ at x, and if $x = \infty$ then H_x is the closed upper-half space

$$H_{\infty,c} := \{(x_1, \dots, x_n) \in \mathbb{H}^n \mid x_n \geq c\} \quad \text{for some } c > 0,$$

called the horoball centered at ∞ with hight c. The *horosphere* ∂H_x inherits a Euclidean metric from the hyperbolic metric, which is invariant by γ.

(3) γ is *hyperbolic*, i.e., γ has precisely two fixed points in $\partial\mathbb{H}^n$, one of which is repelling and the other is attracting. The geodesic in $\mathbb{H}^n$ joining the two fixed points is the unique geodesic which is preserved by γ; it is called the *axis* of γ, and denoted by axis γ.

For low dimensions $n = 2$ and 3, we have:

$$\mathrm{Isom}^+(\mathbb{H}^2) \cong \mathrm{PSL}(2, \mathbb{R}), \quad \mathrm{Isom}^+(\mathbb{H}^3) \cong \mathrm{PSL}(2, \mathbb{C}).$$

We identify the upper-half space $\mathbb{H}^3 = \mathbb{R}^2 \times \mathbb{R}_+$ with $\mathbb{C} \times \mathbb{R}_+$ and identify the ideal boundary $\partial\mathbb{H}^3$ with the Riemann sphere $\mathbb{C} \cup \{\infty\}$. Then the action of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in$ $\mathrm{PSL}(2, \mathbb{C})$ on $\partial\mathbb{H}^3 = \mathbb{C} \cup \{\infty\}$ is given by the linear fractional transformation

$$A(z) = \frac{az + b}{cz + d}.$$

Assume that $A \neq \pm E$, where E is the identity matrix. Then, as we see in the following, the orientation-preserving isometry of $\mathbb{H}^3$ corresponding to $A \in \mathrm{PSL}(2, \mathbb{C})$ is elliptic, parabolic, or hyperbolic according as the trace $\mathrm{tr}\, A$ (which is defined up to sign change) belongs to $(-2, 2)$, $\{\pm 2\}$, or $\mathbb{C} - [-2, 2]$.

Case 1. $\mathrm{tr}\, A \neq \pm 2$. Then A has precisely two fixed points in $\partial\mathbb{H}^3$. After conjugation in $\mathrm{PSL}(2, \mathbb{C})$, we may assume that they are 0 and ∞. Thus $A(z) = az$ for some $a = re^{\theta\sqrt{-1}} \in \mathbb{C}^* - \{1\}$. The action of the isometry A on $\mathbb{H}^3$ is given by:

$$A(z, t) = (az, |a|t) = (re^{\theta\sqrt{-1}}z, rt).$$

This is a skrew motion along the geodesic *axis* $0 \times \mathbb{R}_+$ with (signed) translation length $\log r$ and rotation angle θ. The quantity

$$\mathcal{L}_A := \log r + \theta\sqrt{-1} = \log a \in \mathbb{C}/2\pi\sqrt{-1}\mathbb{Z}$$

is called the *complex translation length* of the isometry A. If we interchange 0 and ∞ by conjugation in $\mathrm{PSL}(2, \mathbb{C})$, then the complex translation length changes into $-(\log r + \theta\sqrt{-1})$. Thus the complex translation length is defined only modulo $2\pi\sqrt{-1}\mathbb{Z}$ and up to multiplication by ± 1. In fact, a simple calculation implies $\mathcal{L}_A$ is characterized by the following identity: (Note that $\mathrm{tr}\, A$ for $A \in \mathrm{PSL}(2, \mathbb{C})$ is defined only up to sign.)

$$\pm \operatorname{tr} A = 2 \cosh \frac{\mathcal{L}_A}{2}$$

If we fix an orientation of the axis, then the complex translation length is defined as an element in $\mathbb{C}/2\pi\sqrt{-1}\mathbb{Z}$. Note that A is elliptic if and only if $r = 1$, which equivalent to the condition that $\operatorname{tr} A \in (-2, 2)$ (under the assumption that $A \neq \pm E$). Thus A is hyperbolic or elliptic according to whether $\operatorname{tr} A$ is contained in $(-2, 2)$ or $\mathbb{C} - [-2, 2]$.

Case 2. $\operatorname{tr} A = \pm 2$. Then A has a unique fixed point in $\partial \mathbb{H}^3$ and hence parabolic. After conjugation in $\mathrm{PSL}(2, \mathbb{C})$, we may assume that it is ∞. Thus $A(z) = z + \tau$ for some $\tau \in \mathbb{C}^*$. The action of the isometry A on $\mathbb{H}^3$ is given by:

$$A(z, t) = (z + \tau, t).$$

In this case, the complex translation length $\mathcal{L}_A$ is defined to be 0.

The following lemma can be easily proved.

Lemma 3.5.1 *Let A be a nontrivial element in* $\mathrm{PSL}(2, \mathbb{C})$ *which is not an elliptic element of order* 2*. Then the centralizer $C(A)$ in* $\mathrm{PSL}(2, \mathbb{C})$ *is as follows.*

(1) If A is elliptic or hyperbolic, then $C(A) - \{E\}$ consists of elliptic/hyperbolic elements which share the same axis with A. Thus $C(A)$ is isomorphic to the multiplicative group $\mathbb{C}^*$
(2) If A is parabolic, then $C(A) - \{E\}$ consists of parabolic elements which share the same parabolic fixed point with A. Thus $C(A)$ is isomorphic to the additive group $\mathbb{C}$

3.5.2 Basic Facts for Hyperbolic Manifolds

By a *hyperbolic structure* on an n-manifold M, we mean a Riemannian metric on M of constant sectional curvature -1: the curvature condition means that every point in M has a neighborhood isometric to an open set of $\mathbb{H}^n$. A hyperbolic structure on M induces a hyperbolic structure on the universal cover $\tilde{M}$ of M which is invariant by the action of the covering transformation group. Thus we obtain a local isometry $D : \tilde{M} \to \mathbb{H}^n$, called the *developing map*, and a homomorphism $\rho : \pi_1(M) \to \operatorname{Isom} \mathbb{H}^n$, called the *holonomy representation*, such that D is ρ-equivariant, i.e., $D \circ \gamma = \rho(\gamma) \circ D : \tilde{M} \to \mathbb{H}^n$.

A hyperbolic structure on M is *complete* if the induced metric on M is complete. This condition is equivalent to the condition that the induced metric on $\tilde{M}$ is complete, which in turn is equivalent to the condition that the developing map $D : \tilde{M} \to \mathbb{H}^n$ is an isometry. Then the holonomy representation $\rho : \pi_1(M) \to \operatorname{Isom} \mathbb{H}^n$ is faithful and discrete, nameley ρ gives an isomorphism from $\pi_1(M)$ to a discrete torsion-free subgroup, Γ, of $\operatorname{Isom}\mathbb{H}^n$. Thus the complete hyperbolic manifold M is identified with $\mathbb{H}^n/\Gamma$.

By a *Kleinian group* we mean a discrete subgroup of $\operatorname{Isom}^+\mathbb{H}^3$, and by a *Fuchsian group* we mean a discrete subgroup of $\operatorname{Isom}^+\mathbb{H}^2$. By Lemma 3.5.1, any commutative torsion-free Kleinian group is conjugate to one of the three groups in the following example.

Example 3.5.2 (Commutative Torsion-Free Kleinian Groups)

(1) The infinite cyclic group $J_0 = J_0(re^{\theta\sqrt{-1}})$ generated by the hyperbolic element $A(z) = re^{\theta\sqrt{-1}}z$ with $r > 1$. The hyperbolic manifold $\mathbb{H}^3/J_0$ is homeomorphic to the interior of the solid torus, and it has the unique closed geodesic with length $\Re\mathcal{L}_A = \log r$. For any $r > 0$, the closed r-neighborhood of axis A is invariant by J_0, and its quotient by J_0 is called a *tube* around the closed geodesic.
(2) The infinite cyclic group J_1 generated by the parabolic transformation $A(z) = z+1$. The hyperbolic manifold $\mathbb{H}^3/J_1$ is homeomorphic to the product $\operatorname{int} D^* \times \mathbb{R}_+$, where $D^* = D^2 - \{0\}$ is a once-punctured disk. This hyperbolic manifold does not contain a closed geodesic. For any $c > 0$, the horoball $H_{\infty,c}$ is invariant by J_1 and and its quotient by J_1 is called an *annulus cusp*.
(3) The rank 2 free abelian group $J_2 = J_2(\tau)$ generated by the two parabolic transformations $A(z) = z+1$ and $B(z) = z+\tau$ with $\tau \in \mathbb{C}-\mathbb{R}$. The hyperbolic manifold $\mathbb{H}^3/J_2$ is homeomorphic to the product $T^2 \times \mathbb{R}_+$, and it does not contain a closed geodesic. For any $c > 0$, the horoball $H_{\infty,c}$ is invariant by J_2 and its quotient by J_2 is called a *torus cusp*. The boundary torus $\partial H_{\infty,c}/J_2$ admits a Euclidean structure which is conformally equivalent to the Euclidean torus $\mathbb{C}/\langle 1, \tau\rangle$. Though the cusp neighborhood $H_{\infty,c}/J_2$ is noncompact, its volume $\operatorname{vol}(H_{\infty,c}/J_2) = \frac{1}{2}\operatorname{area}(\partial H_{\infty,c}/J_2)$ is finite. The complex number τ is called the *modulus* of the cusp torus with respect to the basis $\{A, B\}$.

For an orientable complete hyperbolic 3-manifold $M = \mathbb{H}^3/\Gamma$ and a point $x \in M$ the *injectivity radius* $r(x, M)$ of M at x is defined by

$$r(x, M) = \sup\{r > 0 \mid \text{the } r\text{-neigborhood of } x \text{ in } M \text{ is isometric to an } r\text{-ball in } \mathbb{H}^3\}.$$

For a given constant $\epsilon > 0$, we can decompose M into the ϵ-*thick part*

$$M_{\geq\epsilon} = \{x \in M \mid r(x, M) \geq \frac{1}{2}\epsilon\}$$

and its complement

$$M_{<\epsilon} = \{x \in M \mid r(x, M) < \frac{1}{2}\epsilon\}.$$

The closure of $M_{<\epsilon}$ is denoted by $M_{\leq\epsilon}$ and is called the ϵ-*thin part* of M. (This complicated definition eliminates the trouble which occurs when there is a closed geodesic of length ϵ [307, p.254].) The following is a consequence of the Margulis lemma (see [300, Theorem 5.10.1 and Corollary 5.10.2]).

Theorem 3.5.3 *There is a universal constant $\epsilon_0 > 0$, such that for any positive constant $\epsilon < \epsilon_0$ and for any orientable complete hyperbolic manifold $M = \mathbb{H}^3/\Gamma$, the ϵ-thin part $M_{\leq\epsilon}$ is a disjoint union of tubes around (short) simple closed geodesics, annulus cusps, and torus cusps.*

The following proposition can be proved by using the above theorem and the concept of convex core introduced in Sect. 3.10.1. (See [300, Proposition 5.11.1]).

Proposition 3.5.4 *If an orientable complete hyperbolic manifold $M = \mathbb{H}^3/\Gamma$ has finite volume, then M is the union of a compact submanifold (bounded by tori) and finitely many torus cusps $C_1, \ldots, C_m$ for some $m \geq 0$. In particular, M is identified with the interior of a compact 3-manifold $\bar{M}$ with (possibly empty) toral boundary.*

To end this subsection, we recall an important consequence of Thurston's hyperbolization theorem for Haken manifolds.

Definition 3.5.5 A knot or link L in S^3 is *hyperbolic* if its complement $S^3 - L \cong \operatorname{int} E(L)$ admits a complete hyperbolic structure of finite volume.

The following theorem is a special case of Theorem 3.3.5, which in turn is a special case of the geometrization theorem.

Theorem 3.5.6 *A prime knot in S^3 is hyperbolic if and only if it is neither a torus knot nor a satellite knot. More generally, an unsplittable prime link L is hyperbolic if and only if $E(L)$ is atoroidal and is not a Seifert fibered space.*

3.5.3 Rigidity Theorem for Complete Hyperbolic Manifolds of Finite Volume

For complete hyperbolic structures of finite volume of dimension ≥ 3, the following strong rigidity theorem is established by Mostow [225] and Prasad [257] (cf. [300, Theorem 5.7.2]).

Theorem 3.5.7 (The Mostow–Prasad Rigidity Theorem) *If an orientable n-manifold with $n \geq 3$ admits a complete hyperbolic structure of finite volume, then this structure is unique. To be precise, the following holds. Let Γ_i $(i = 1, 2)$ be discrete torsion free subgroups of $\operatorname{Isom}^+ \mathbb{H}^n$ with $n \geq 3$ of cofinite volume, i.e., $\operatorname{vol}(\mathbb{H}^n/\Gamma_i) < \infty$. Then any isomorphism $\phi : \Gamma_1 \to \Gamma_2$ is realized by a unique isometry $f : \mathbb{H}^n/\Gamma_1 \to \mathbb{H}^n/\Gamma_2$.*

This theorem together with Thurston's hyperbolization theorem had tremendous impact in knot theory. Because Theorem 3.5.6 says that almost all knots are hyperbolic (moreover, the Geometrization Theorem 3.3.5 reduces the study of knots to the study of hyperbolic links) and the above theorem imply that geometric invariants, such as volumes, cusp shapes, and lengths of shortest closed geodesics, of the complete hyperbolic structures on knot/link complements are topological invariants of the knots/links.

3.6 Computation of Hyperbolic Structures and Canonical Decompositions of Cusped Hyperbolic Manifolds

Epstein and Penner proved that every cusped hyperbolic manifold of finite volume admits a natural ideal polyhedral decomposition, called the canonical decomposition. This fact (together with the rigidity theorem and the Gordon–Luecke knot complement theorem) has the following striking consequence in knot theory. *The combinatorial structure of the canonical decomposition of a hyperbolic knot complement is a complete knot invariant.* Moreover the marvelous computer program *SnapPea* developed by Jeffrey Weeks enabled us to compute the canonical decompositions of knot complements. For example, SnapPea immediately tells us that the Kinoshita–Terasaka knot and the Conway's knot are different and that they admit no symmetry.

In this section, we recall the Epstein–Penner canonical decomposition and its impact on knot theory. We also recall a method for constructing hyperbolic structures by using ideal triangulation, which was first explained in Thurston's lecture notes [300, Chapter 4], and explain a method for finding the canonical decomposition. In the final subsection, we give a list of geometric invariants of hyperbolic knots, which are guaranteed to be knot invariants by the rigidity theorem, and introduce their study from the viewpoint of *effective geometrization*.

3.6.1 The Canonical Decompositions of Cusped Hyperbolic Manifolds

Let $M = \mathbb{H}^n/\Gamma$ be an orientable complete hyperbolic n-manifold of finite volume with $m \geq 1$ cusps. Pick mutually disjoint cusps $C_1, \ldots, C_m$ of M, and set $C = \cup_{i=1}^m C_i$. Then we can canonically construct a spine $\mathcal{F}$ and a canonical ideal polyhedral decomposition $\mathcal{D}$ of M as follows.

Observe that a generic point in $M - C$ has a unique shortest geodesic path to C but that there are exceptional points which have more than one shortest geodesic paths to C. Let $\mathcal{F}$ be the subset of $M - C$ consisting of these exceptional points. Namely, $\mathcal{F}$ is the *cut locus* in M with respect to the cusps $C = \cup_{i=1}^m C_i$. Then $\mathcal{F}$ is a locally finite totally geodesic cell complex in M, and there is a deformation retraction of M onto $\mathcal{F}$. We call it the *Ford complex* or *Ford spine* of M, with respect to the choice of cusps $C_1, C_2, \ldots, C_m$.

By taking the geometric dual to $\mathcal{F}$ as follows, we obtain an ideal polyhedral decomposition $\mathcal{D}$ of M. Let $\tilde{\mathcal{F}}$ and $\tilde{C}$ be the inverse images of $\mathcal{F}$ and C in the universal covering $\mathbb{H}^n$ of M. Pick a vertex x of $\tilde{\mathcal{F}}$. Then there are finitely many shortest geodesic paths from x to $\tilde{C}$. Let $\{v_i\}$ be the ideal points in $\partial\mathbb{H}^n$ forming the centers of the horoball components of $\tilde{C}$ which are joined to x by a shortest geodesic path. The convex hull of the ideal points $\{v_i\}$ forms an n-dimensional *ideal polyhedron* of $\mathbb{H}^n$, and the collection of all such ideal polyhedra, where x runs over the vertices of $\tilde{\mathcal{F}}$, determines a Γ-invariant tessellation of $\mathbb{H}^n$. The tessellation descends to an ideal polyhedral decomposition $\mathcal{D}$ of $M = \mathbb{H}^n/\Gamma$.

Epstein and Penner [82] gave a description of the decomposition $\mathcal{D}$ by using a convex hull construction in Minkowski space. Their description shows that each cell of $\mathcal{D}$ admits a natural (incomplete) Euclidean structure: so, these decompositions are called *Euclidean decompositions*. In [14], a generalization of the Epstein–Penner construction to cusped hyperbolic manifolds of infinite volume is given, and their relationship to the convex cores are discussed.

The Ford complex $\mathcal{F}$ and its geometric dual $\mathcal{D}$ depend only on the ratio of the volumes $\mathrm{vol}(C_1) : \mathrm{vol}(C_2) : \cdots : \mathrm{vol}(C_m)$. Moreover, it is proved by Akiyoshi [13] that the combinatorial structures of $\mathcal{F}$, when the ratio varies, are finite. The ideal polyhedral decomposition $\mathcal{D}$, for the case when $C_1, C_2, \ldots, C_m$ have the same volume, is uniquely determined by the hyperbolic manifold M, and is called the *canonical decomposition* of the cusped hyperbolic manifold M.

Example 3.6.1

(1) Let M be the hyperbolic thrice-punctured sphere, obtained by gluing two ideal hyperbolic triangles through identification of their boundaries via the identity map. Then this decomposition of M into the two copies of ideal triangles is the canonical decomposition of M. The corresponding Ford complex of M is a θ-shaped geodesic spine of M consisting of two vertices and three edges.
(2) As shown by Thurston [301, Chapter 4], the complete hyperbolic structure of the figure-eight knot complement M is obtained by glueing two copies of the regular ideal tetrahedron. The decomposition of M into the two copies of the regular ideal tetrahedron is the canonical decomposition of M.

Since the complete hyperbolic structure of a given knot complement is unique by the Mostow–Prasad rigidity theorem (Theorem 3.5.7), and since by the knot complement theorem (Theorem 3.2.4) a knot is determined by its complement, it follows that the combinatorial structure of the canonical decomposition of a hyperbolic knot complement is a complete topological invariant of the knot.

Theorem 3.6.2

(1) Two hyperbolic knots are equivalent, if and only if the canonical decompositions of their complements are combinatorially equivalent.
(2) Let K be a hyperbolic knot and $\mathcal{D}$ the canonical decomposition of $S^3 - K$. Then

$$\mathrm{Sym}(S^3, K) \cong \mathrm{Isom}(S^3 - K) \cong \mathrm{Aut}(\mathcal{D}).$$

In the above theorem, $\mathrm{Sym}(S^3, K) := \pi_0 \mathrm{Diff}(S^3, K)$ denotes the *symmetry group* of the knot K, and $\mathrm{Aut}(\mathcal{D})$ denotes the combinatorial automorphism group of $\mathcal{D}$.

Example 3.6.3 (1) It is a simple exercise to see that the automorphism group of the canonical decomposition $\mathcal{D}$ of the complement figure-eight knot K is isomorphic to the order 8 dihedral group D_8. Thus we have $\mathrm{Sym}(S^3, K) \cong \mathrm{Isom}(S^3 - K) \cong D_8$.

(2) The canonical decompositions of the complements of the Kinoshita–Terasaka knot and the Conway knot consist of 12 and 14 ideal tetrahedra, respectively. Hence they are inequivalent, even though they are mutants of each other and so they share the same Alexander polynomial, the Jones polynomial, the hyperbolic volume and the same double branched coverings. Moreover, the automorphism groups of both canonical decompositions are trivial. Thus the symmetry groups of these two knots are trivial. In particular, both of them are neither amphicheiral nor invertible. The noninvertibility of 8_{17} can be also proved by using the canonical decomposition of the knot complement.

As explained in the next subsection, the canonical decompositions are amenable to computer calculation, and wonderful computer programs were developed: *SnapPea* by Weeks [320], *Snap* by Coulson–Goodman–Hodgson–Neumann [68], *SnapPy* by Culler–Dunfield–Goerner [75], and a computer verified program *HIKMOT* by Hoffman–Ichihara–Kashiwagi–Masai–Oishi–Takayasu [143]. The results in Example 3.6.3(2) are, of course, obtained by any of these programs.

This enabled Hoste, Thistlethwaite and Weeks [144] to extend (and correct) Conway's enumeration of all 11 crossing knots to include all prime knots up 16 crossings. There are 1,701,936 such knots, and all except for 32 knots are hyperbolic! To be precise, Hoste and Weeks used the canonical decomposition, and Thistelethwaite used the "universal method" described at the end of Sect. 3.2.7. Thus their table is double checked, and this fact shows the strength of both methods.

This is something like a magic wand for knot theorists as long as finitely many knots of reasonable crossing numbers are concerned. However, to understand the canonical decompositions of infinite families of knots or cusped hyperbolic manifolds is not easy. For the Farey manifolds, namely punctured torus bundles and 2-bridge knot complements, the combinatorial structures of the canonical decompositions are determined by Jorgensen [152] and Guéritaud [121] (cf. [15, 120, 281]).

To end this subsection, we remark that it is still an open problem to see whether every orientable cusped hyperbolic 3-manifold of finite volume admits a *ideal triangulation*, namely an ideal polyhedral decomposition consisting of ideal tetrahedra. Here an *ideal tetrahedron* is a closed convex hull in $\mathbb{H}^3$ of 4 ideal points in $\partial\mathbb{H}^3$, called the *ideal vertices*. (Since any such manifold M admits an ideal polyhedral decomposition by [82] and since every ideal polyhedron is decomposed into ideal tetrahedra, M admits a *partially flat* ideal triangulation, namely one in which some of the tetrahedra degenerate into flat quadrilaterals with distinct vertices (see [256]). But this does not necessarily lead to a genuine ideal triangulation of M.) Wada, Yamashita and Yoshida [315] and Yoshida [326] proved the existence of such triangulations under certain combinatorial conditions on the polyhedral decomposition, and Luo, Schleimer and Tillman [188] proved that every such manifold virtually admits an ideal triangulation, namely some finite cover has an ideal triangulation. Hodgson, Rubinstein and Segerman [141] considered a relaxed version of the problem, and proved, in particular, that every hyperbolic link complement in S^3 admits a topological ideal triangulation with a "strict angled structure".

3.6.2 *Ideal Triangulations and Computations of the Hyperbolic Structures*

Let $M = \mathbb{H}^3/\Gamma$ be an orientable complete hyperbolic 3-manifold of finite volume with $m \geq 1$ cusps, and let $\rho : \pi_1(M) \to \Gamma < \mathrm{PSL}(2, \mathbb{C})$ be the holonomy representation. Then, as we have observed in the previous section, M admits an ideal polyhedral decomposition $\mathcal{D}$. We now assume that $\mathcal{D}$ is an ideal triangulation, namely $\mathcal{D}$ consists of ideal tetrahedra. Any ideal tetrahedron Δ (up to isometry) is represented by a complex number z with positive imaginary part, such that the Euclidean triangle cut out of any vertex of Δ by a horosphere is similar to the triangle in $\mathbb{C}$ with vertices 0, 1, and z. In fact, Δ is isometric to the ideal tetrahedron $\Delta(z)$ spanned by $0, 1, \infty$ and z in the upper half-space model $\mathbb{C} \times \mathbb{R}_+$ of $\mathbb{H}^3$. We call z the *shape parameter* of the ideal tetrahedron $\Delta(z)$. (If z has negative imaginary part, then $\Delta(z)$ is regarded as negatively oriented. If z is a real number different from 0 and 1, then $\Delta(z)$ is regarded as a degenerate ideal tetrahedron.) The complex numbers z, $(z-1)/z$, and $1/(1-z)$ give isometric ideal tetrahedra, and we give each edge e of $\Delta = \Delta(z)$ one of the three complex numbers by the following rule, and call it the *edge parameter* of Δ associated with e.

- Edges $[0, \infty]$ and $[1, z]$ have edge parameter z.
- Edges $[1, \infty]$ and $[0, z]$ have edge parameter $1/(1-z)$.
- Edges $[z, \infty]$ and $[0, 1]$ have edge parameter $(z-1)/z$.

Let e be an edge of the ideal triangulation $\mathcal{D}$ of M, and let $z_1, \ldots, z_k$ be the edge parameter of the edges of ideal tetrahedra glued to e. Since these ideal tetrahedra close up as one goes around e, the parameters satisfies the following equation.

$$\prod_{i=1}^{k} z_j = 1 \quad \text{and} \quad \sum_{j=1}^{k} \arg(z_j) = 2\pi$$

This condition is identical to the following equation, which is called the *gluing equation* around e.

$$\sum_{j=1}^{k} \log(z_j) = 2\pi\sqrt{-1},$$

where $\log : \mathbb{C} - \mathbb{R}_{\leq 0} \to \mathbb{C}$ is the branch of the logarithm function whose imaginary part lies in $(-\pi, \pi)$.

Let T be a torus boundary component of the compact 3-manifold $\bar{M}$ whose interior is homeomorphic to the hyperbolic manifold $M = \mathbb{H}^3/\Gamma$, and let μ be an oriented essential simple loop on T. (A simple loop on T is *essential* if it does not bound a disk in T.) By identifying T with a cusp torus, and considering the intersection with the cusp torus with the ideal triangulation $\mathcal{D}$, we obtain a triangulation of T, whose vertices correspond to the edges of $\mathcal{D}$ and whose triangles correspond to truncations of ideal tetrahedra around ideal vertices. We may assume that μ intersects the edges of the triangulation transversely and does not intersect the vertices of the triangulation. Each segment of μ in a triangle cuts off a single vertex of the triangle, and so has an associated edge parameter z_j. Define $\epsilon_j = +1$ or -1 according to whether the vertex lies to the *left* of μ or not. (Here we assume that ∞ is a parabolic fixed point of Γ, $\pi_1(T)$ is identified with the stabilizer Γ_∞ of ∞, and T is identified with the Euclidean torus $\mathbb{C}/\Gamma_\infty$ via the projection from a horosphere centered at ∞ to $\mathbb{C}$. The left/right convention is determined by the standard orientation of $\mathbb{C}$.) Then we can see that the complex translation length of the image $\rho(\mu)$ of μ by the holonomy representation ρ of the complete hyperbolic manifold M is represented by the complex number

$$\mathcal{L}_\mu := \sum_j \epsilon_j \log(z_j).$$

Since $\rho(\mu)$ is parabolic, we have $\mathcal{L}_\mu = 0$. Thus we have the following *completeness equation*

$$\sum_i \epsilon_i \log(z_i) = 0.$$

Conversely, let $\bar{M}$ be an orientable compact manifold whose boundary is non-empty and consists of tori, and let $\mathcal{D}$ be a *topological ideal triangulation* of $M = \operatorname{int} \bar{M}$. Namely $\mathcal{D}$ is a topological trianglulation (a cell decomposition whose cells are identified with simplices) of the space $\hat{M} = \bar{M}/\sim$, where $\sim$ is the equivalence relation which identifies all points of each boundary component of $\bar{M}$, such that the vertex set of $\mathcal{D}$ is equal to the finite set consisting of the image of $\partial \bar{M}$. By a simple argument using the Euler characteristic, we see that the number of edges in $\mathcal{D}$ is equal to the number, t, of tetrahedra in $\mathcal{D}$. Now let $\mathbb{H}_+ = \{z \in \mathbb{C} \mid \Im z > 0\}$ be the upper-half space of the complex plane. Pick a t-tuple of complex numbers $\boldsymbol{z} = (z_1, \ldots, z_t) \in (\mathbb{H}_+)^t \subset \mathbb{C}^t$ with positive imaginary parts, and identify the topological ideal tetrahedra $\{\Delta_1, \ldots, \Delta_t\}$ with hyperbolic ideal tetrahedra $\{\Delta(z_1), \ldots, \Delta(z_t)\}$. Since all hyperbolic ideal triangles are isometric, we can realize the topological gluing maps among the faces of the topological ideal tetrahedra by hyperbolic isometries. Thus we obtain a hyperbolic structure on the complement of the 1-skeleton of $\mathcal{D}$. We have the following theorem (see [242, 300]).

Theorem 3.6.4 *Under the above setting, the following hold for each* $z = (z_1, \ldots, z_t) \in (\mathbb{H}_+)^t \subset \mathbb{C}^t$.

(1) The hyperbolic structure on the complement of the 1*-skeleton of* $\mathcal{D}$ *extends a hyperbolic structure on the whole* M *if and only if* z *satisfies the gluing equation at every edge of* $\mathcal{D}$.

(2) When condition (1) is satisfied, the resulting hyperbolic structure on M *is complete if and only if* z *also satisfies the completeness equation at every boundary component of* $\bar{M}$ *(for a single choice of an oriented essential simple loop* μ *for each boundary component).*

Remark 3.6.5 Let $\mathcal{X}$ be the variety of $z = (z_1, \ldots, z_t) \in \mathbb{C}^t$ consisting of the solutions of the gluing equations. Then, by a combinatorial argument, we can see that $\mathcal{X}$ has dimension m over $\mathbb{C}$, where m is the number of boundary components of $\bar{M}$ ([300, Theorem 5.6], [242, Proposition 2.3]). By the rigidity theorem, there is a unique point $z^0 \in \mathcal{X} \cap (\mathbb{H}_+)^t$ which satisfies the completeness equation. It is proved by [242, Section 4] that z^0 is a smooth point of $\mathcal{X} \cap (\mathbb{H}_+)^t$, namely there is a neighborhood of z^0 in $\mathcal{X} \cap (\mathbb{H}_+)^t$ which is biholomorphically equivalent to an open set in $\mathbb{C}^m$. (Moreover, it was proved by Choi [62] that $\mathcal{X} \cap (\mathbb{H}_+)^t$ is a smooth complex manifold.) This fact plays a crucial role in a proof of the hyperbolic Dehn filling theorem (see Sect. 3.7.2).

On the other hand, there is a convenient method for obtaining topological ideal triangulations of knot/link complements from diagrams (see [206, 297, 322]). Thus we have a good chance to construct a complete hyperbolic structure on a given knot/link complement by applying Theorem 3.6.4. In fact, this works extremely well, though the proof of Thurston's uniformization theorem is very difficult.

Moreover, if a given ideal triangulation $\mathcal{D}$ of M satisfies a certain inequality at each codimension 1 face of $\mathcal{D}$, then $\mathcal{D}$ is the canonical decomposition (see [321]). If the inequality was not satisfied at some face of $\mathcal{D}$, then apply the Pachner $3-2$ move to $\mathcal{D}$ at the face, if it is geometrically realizable, and check if the conditions for the faces hold. If this does not lead to the canonical decomposition, then retriangulate $\mathcal{D}$ randomly, and repeat the above procedure. This is the way SnapPea finds the canonical decompositions. Though there is no theoretical guarantee, this method is extremely efficient (see [321, 322]). For the treatment of the case when the canonical decomposition is not an ideal triangulation, see the work of Hodgson and Weeks [140].

3.6.3 Other Geometric Invariants for Hyperbolic Knots and Effective Geometrization

In addition to the canonical decomposition, there are various important geometric invariants of hyperbolic knots and links.

- The volumes and the Chern–Simons invariants of the hyperbolic link complements.
- The volumes of the maximal cusps.
- The moduli of the Euclidean cusp tori.
- Length spectrum, i.e., the multi-set of lengths of closed geodesics, in particular the length of the shortest closed geodesic.
- Lengths of the vertical geodesic paths, joining maximal cusps to themselves.
- Euclidean length spectrum of the maximal cusp torus.

Volumes of hyperbolic manifolds are treated in Sect. 3.8.

In the recent beautiful survey [96], Futer, Kalfagianni, and Purcell discuss these invariants from the viewpoint of *effective geometrization* or *WYSIWYG topology*, where WYSIWYG stands for "what you see is what you get", which aims to determine the geometry of link complements directly from the link diagrams. A typical example in this direction is the following estimate by Lackenby [176] of the volume of alternating link complements in terms of the twist number.

Theorem 3.6.6 *Let D be a reduced alternating diagram of a hyperbolic link L in S^3, and let $t(D)$ be the twist number of the diagram D. Then*

$$\frac{1}{2}V_{\mathrm{tet}}(t(D)-2) \le \mathrm{vol}(S^3 - L) \le V_{\mathrm{tet}}(16t(D)-16)$$

where $V_{\mathrm{tet}} = 1.0149416\ldots$ is the volume of the regular ideal tetrahedron.

Here the *twist number* $t(D)$ of a link digram D is the number of twists of D, where a *twist* of D is either a connected collection of bigon regions in D arranged in a row which is maximal in the sense that it is not part of a longer row of bigons, or a single crossing adjacent to no bigon regions.

The article [96] presents a nice survey on the recent great progress towards effective geometrization, including a refinement of the above result.

3.7 Flexibility of Incomplete Hyperbolic Structures and the Hyperbolic Dehn Filling Theorem

By the Mostow–Prasad rigidity theorem, the complete hyperbolic structure on a 3-manifold M of finite volume is rigid. However, when M has a cusp, the complete hyperbolic structure admits nontrivial continuous deformations into incomplete hyperbolic structures (see Remark 3.6.5). In the generic case, the metric completion yields a pathological topological space which is not even Hausdorff. However, in certain special isolated cases, the metric completion produces a complete hyperbolic manifold. This is a rough idea of Thurston's hyperbolic Dehn filling Theorem. This theorem has stimulated keen attention of many mathematicians and enormous amount of research grew out of this result. In this section, we give an outline of a proof of this theorem and a brief survey of its influence on knot theory.

3.7.1 Hyperbolic Dehn Filling Theorem

We begin by recalling the topological operation, Dehn filling. By an *oriented slope* on a torus T, we mean the isotopy class of an oriented essential simple loop on T. Each oriented slope represents a primitive element of $H_1(T;\mathbb{Z})$, and conversely any primitive element of $H_1(T;\mathbb{Z})$ is represented by a unique oriented slope. If we fix a basis $\{\mu,\lambda\}$ of $H_1(T;\mathbb{Z})$, then a primitive element of $H_1(T;\mathbb{Z})$ is expressed as $p\mu + q\lambda$ where (p,q) is a pair of relatively prime integers. Thus we can identify the set of oriented slopes on T with the set of pairs of relatively prime integers $(p,q) \in \mathbb{Z}^2 \subset \mathbb{R}^2 \cup \{\infty\} \cong S^2$.

Let M be a connected compact orientable 3-manifold whose boundary consists of m tori $T_1,\ldots,T_m$. Pick an oriented slope ν_j on T_j for each j, and attach a solid torus $V_j = D_j^2 \times S^1$ to M along T_j, so that the meridian $\partial D_j^2 \times \{*\}$ is identified with the slope ν_j. The resulting manifold is denoted by $M(\boldsymbol{\nu}) = M(\nu_1,\ldots,\nu_m)$ and called the result of *Dehn filling* of M along the tuple $\boldsymbol{\nu} = (\nu_1,\ldots,\nu_m)$ of oriented slopes. We extend this operation to the case where some ν_j is the symbol ∞, by the rule that if $\nu_j = \infty$ then we leave the boundary T_j as it is. In particular, $M(\infty,\ldots,\infty) = M$.

The following theorem is proved by Thurston [300, Chapters 4 and Section 5.8].

Theorem 3.7.1 (Hyperbolic Dehn Filling Theorem) *Let M be a connected compact orientable 3-manifold whose boundary consists of m tori, and suppose that* int M *admits a complete hyperbolic structure of finite volume. Then, except for finitely many choices of the slopes of ν_j for each $1 \le j \le m$, the manifold $M(\nu_1,\ldots,\nu_m)$ admits a complete hyperbolic structure. To be more precise, there exists a neighborhood V of $(\infty,\ldots,\infty)$ in $(\mathbb{R}^2\cup\{\infty\})^m$ such that $M(\nu_1,\ldots,\nu_m)$ admits a complete hyperbolic structure for every slope $(\nu_1,\ldots,\nu_m)$ contained in V.*

Remark 3.7.2

(1) The operation at T_j is actually determined by the *slope* (the isotopy class of an unoriented essential simple loop on a torus) obtained from ν_j by forgetting the orientation.
(2) When M is the exterior of an m-component link $L = \cup_{j=1}^m K_j$ in S^3, we fix an orientation of each component K_j of L, and choose the meridian-longitude systems $\{\mu_j,\lambda_j\}$ as a preferred basis for $H_1(T_j;\mathbb{Z})$, and represent an oriented slope, ν_j, on T_j by a pair of relatively prime integers (p_j,q_j) with $\nu_j = p_j\mu_j + q_j\lambda_j$. The slope obtained from ν_j by forgetting the orientation is uniquely determined by the rational number $p_j/q_j \in \mathbb{Q}\cup\{1/0\}$. (It should be noted that slope $1/0$ and the symbol ∞ have different meanings.) Moreover, this does not depend on the choice of the orientation of K_j. We denote the manifold $M(\nu_1,\ldots,\nu_m)$ by $M(p_1/q_1,\ldots,p_m/q_m)$, and call it the result of *Dehn surgery* on L with slope $(p_1/q_1,\ldots,p_m/q_m)$.

In Theorem 3.7.1, a slope (or a tuple of slopes) which does not produce a hyperbolic manifold is called an *exceptional slope*.

Example 3.7.3

(1) The exceptional slopes of the figure-eight knot K are the slopes p/q with $-4 \leq p \leq 4$ and $-1 \leq q \leq 1$. Thus the set of exceptional slopes is $\{1/0, 0, \pm 1, \pm 2, \pm 3, \pm 4\}$ (see [300, Section 4.6]).
(2) Let M be the exterior of the Whitehead link $L = K_1 \cup K_2$ in S^3. Consider the Dehn filling only along $T_1 = \partial N(K_1)$. Then the exceptional slopes for this Dehn filling are those slopes contained in the parallelogram with vertices $\pm(4, -1)$ and $\pm(0, 1)$ (see [240, Section 6]).

3.7.2 Outline of a Proof and Generalized Dehn Filling Coefficients

We give an outline of the proof of Theorem 3.7.1 by Neumann–Zagier [242] (cf. [19, Section E.6]), when the hyperbolic manifold int M admits an ideal triangulation $\mathcal{D}$. (See Petronio–Porti [256] for a proof without assuming the existence of an ideal triangulation, and using a partially flat ideal triangulation of M.) Let $\Delta_1, \ldots, \Delta_t$ be the ideal tetrahedra in $\mathcal{D}$, and let $\boldsymbol{z}^0 = (z_1^0, \ldots, z_t^0)$ be their shape parameters. By the rigidity theorem and Theorem 3.6.4, $\boldsymbol{z}^0$ is the unique solution of the gluing and the completeness equations. Let $\mathcal{X}$ be the variety of $\boldsymbol{z} = (z_1, \ldots, z_t) \in \mathbb{C}^t$ consisting of the solutions of the gluing equations. For $\boldsymbol{z} \in \mathcal{X} \cap (\mathbb{H}_+)^t$, let $M_{\boldsymbol{z}}$ be the (almost certainly incomplete) hyperbolic manifold determined by the parameter $\boldsymbol{z}$, and let $\rho_{\boldsymbol{z}} : \pi_1(M) \to \mathrm{PSL}(2, \mathbb{C})$ be the holonomy representation of $M_{\boldsymbol{z}}$. For each boundary component T_j of M ($1 \leq j \leq m$), fix an oriented slope μ_j. For $\boldsymbol{z} \in \mathcal{X} \cap (\mathbb{H}_+)^t$, let $u_j(\boldsymbol{z})$ be the complex number $\mathcal{L}_{\mu_j}$, defined as in Sect. 3.6.2, which represents the complex translation length of $\rho_{\boldsymbol{z}}(\mu_j)$. (Though the complex translation length is defined only modulo $2\pi\sqrt{-1}\mathbb{Z}$ and up to multiplication by ± 1, the construction in Sect. 3.6.2 gives a well-defined continuous lift to $\mathbb{C}$.)

For each boundary component T_j, pick an oriented slope λ_j which intersects μ_j transversely in a single point and so $\{\mu_j, \lambda_j\}$ forms a generator system of $H_1(T_j; \mathbb{Z})$. Let $\boldsymbol{v} := (v_1, \ldots, v_m)$ be the map from $\mathcal{X}$ to $\mathbb{C}^m$, where $v_j(\boldsymbol{z})$ is the complex number $\mathcal{L}_{\lambda_j}$, defined as in Sect. 3.6.2, which represents the complex translation length of $\rho_{\boldsymbol{z}}(\lambda_j)$.

Recall the key Remark 3.6.5 that there is a neighborhood of $\boldsymbol{z}^0$ in $\mathcal{X} \cap (\mathbb{H}_+)^t$ which is biholomorphically equivalent to an open set in $\mathbb{C}^m$. By using this fact, we can see that $\boldsymbol{u} := (u_1, \ldots, u_m)$ maps a neighborhood of $\boldsymbol{z}^0 \in \mathcal{X} \cap (\mathbb{H}_+)^t$ biholomorphically onto a neighborhood, $\mathcal{X}_0$, of $0 \in \mathbb{C}^m$ (cf. [242, Section 4]).

We now change notation as follows. For $\boldsymbol{u} \in \mathcal{X}_0$, we denote the corresponding hyperbolic manifold and the holonomy representation by $M_{\boldsymbol{u}}$ and $\rho_{\boldsymbol{u}}$, respectively, and we regard $\boldsymbol{v}$ as a map from $\mathcal{X}_0$ to $\mathbb{C}^m$.

By replacing $\mathcal{X}_0$ with a smaller neighboorhood of 0, if necessary, we can assume that $\boldsymbol{u}$ and $\boldsymbol{v}$ are independent over $\mathbb{R}$, for all $\boldsymbol{u} \in \mathcal{X}_0 - \{0\}$. In fact, there is an analytic function $\boldsymbol{\tau} = (\tau_1, \dots, \tau_m) : \mathcal{X}_0 \to \mathbb{C}^m$, satisfying the following conditions [242, Lemma 4.1]:

(1) $v_j(\boldsymbol{u}) = \tau_j(\boldsymbol{u})u_j$ for every $\boldsymbol{u} = (u_1, \dots, u_j, \dots, u_m) \in \mathcal{X}_0$ and $j = 1, \dots, m$.
(2) $\tau_j(0, \dots, 0)$ is equal to the modulus of the cusp torus T_j of the complete hyperbolic manifold M with respect to $\{\mu_j, \lambda_j\}$.

In particular, we may assume that $\tau_j(\boldsymbol{u})$ is non-real for every $\boldsymbol{u} \in \mathcal{X}_0$, and so $\boldsymbol{u}$ and $\boldsymbol{v}$ are independent over $\mathbb{R}$ for every $\boldsymbol{u} \in \mathcal{X}_0 - \{0\}$.

Now we define the *generalized Dehn filling coefficients* of the j-th boundary torus component $\nu_j \in \mathbb{R}^2 \cup \{\infty\} \cong S^2$ by the formula:

$$\begin{cases} \nu_j = \infty & \text{if } u_j = 0 \\ \nu_j = (p_j, q_j) & \text{where } p_j u_j + q_j v_j = 2\pi\sqrt{-1} \quad \text{if } u_j \neq 0 \end{cases}$$

The hyperbolic Dehn filling Theorem 3.7.1 is a consequence of the following theorem.

Theorem 3.7.4 *Under the above setting, the "generalized Dehn filling coefficients map"* $\boldsymbol{u} \mapsto \boldsymbol{\nu} = (\nu_1, \dots, \nu_m)$ *gives a homeomorphism from a neighborhood* $U \subset \mathcal{X}_0$ *of* 0 *in* $\mathbb{C}^m$ *onto a neigborhood* V *of* $(\infty, \dots, \infty)$ *in* $(\mathbb{R}^2 \cup \{\infty\})^m$*. Moreover, the following hold.*

- *If* $\nu_j = \infty$*, the hyperbolic structure at the* j*-th end is complete.*
- *If* $\nu_j = (p_j, q_j)$ *where* $p_j, q_j \in \mathbb{Z}$ *are coprime, then the completion of the* j*-th end is a hyperbolic* 3*-manifold, which is topologically the Dehn filling such that the simple loop* $p_j\mu_j + q_j\lambda_j$ *on* T_j *bounds a disk.*
- *When* $p_j/q_j \in \mathbb{Q} \cup \{\infty\}$*, let* $m_j, n_j \in \mathbb{Z}$ *be coprime integers such that* $(p_j, q_j) = d(m_j, n_j)$ *for some* $d > 0$*. The completion is a hyperbolic cone* 3*-manifold obtained by gluing a solid torus with singular core, such that the simple loop* $m_j\mu_j + n_j\lambda_j$ *on* T_j *bounds a disk which has a singularity at the center, and that the cone angle of the singular locus is* $2\pi/d$*.*
- *If* $p_j/q_j \in \mathbb{R} - \mathbb{Q}$*, then the metric completion of the* j*-th end is not even topologically a manifold.*

In the above, a *hyperbolic cone* 3*-manifold* is a smooth 3-manifold C equipped with a complete metric (distance function) which is locally isometric to $\mathbb{H}^3$ or to the space $\mathbb{H}^3(\alpha)$ obtained from a geodesic cheese-cake-shaped polyhedron of angle $\alpha > 0$ by identifying two sides. The singular locus $\Sigma \subset C$ is the set of points modeled on the singular line of some $\mathbb{H}^3(\alpha)$, and α is called the *cone angle* at a singular point modeled on this singular line (for precise definition, see [135, Section 1], [65, Chapter 3], [32, Chapter 1], [36, Section 3]). Hyperbolic 3-cone manifolds play a key role in the proof of the orbifold theorem (Theorem 3.4.4).

Remark 3.7.5

(1) Assume that a tuple of oriented slopes $\boldsymbol{\nu} = (\nu_1, \dots, \nu_m)$ is the image of a paremeter $\boldsymbol{u} = (u_1, \cdots, u_m) \in U$ in Theorem 3.7.4, namely the metric

completion of the hyperbolic manifold $M_{\boldsymbol{u}}$ is homeomorphic to the manifold $M(\boldsymbol{\nu})$ obtained from M by Dehn filling along $\boldsymbol{\nu}$. Let $\boldsymbol{\nu}'$ be the tuple of oriented slopes obtained from $\boldsymbol{\nu}$ by replacing some component $\nu_j = (p_j, q_j)$ with $-\nu_j = (-p_j, -q_j)$. Then $\boldsymbol{\nu}'$ is the image of the parameter $\boldsymbol{u}'$ obtained from $\boldsymbol{u}$ by replacing the component u_j with $-u_j$. Since $M(\boldsymbol{\nu}')$ is homeomorphic to $M(\boldsymbol{\nu})$ by a homeomorphism preserving the subspace M, the rigidity theorem implies that $M_{\boldsymbol{u}'}$ is isometric to $M_{\boldsymbol{u}}$. In fact, such an isometry exists whenever two parameters $\boldsymbol{u}$ and $\boldsymbol{u}'$ are related by the involution $(u_1, \dots, u_j, \dots, u_m) \mapsto (u_1, \dots, -u_j, \dots, u_m)$. Thus deformations of the complete hyperbolic manifold M are parametrized by the quotient of U by the $(\mathbb{Z}/2\mathbb{Z})^m$-action, generated by the above involutions with $j = 1, \dots, m$. In other words, the space U is identified with a $(\mathbb{Z}/2\mathbb{Z})^m$-branched covering of a deformation space of M. The space U actually parametrizes the incomplete hyperbolic manifolds $M_{\boldsymbol{u}}$ endowed with an ideal triangulation (see [242, p.323]).

(2) In Theorem 3.7.4, the complete hyperbolic manifolds $\{M(\boldsymbol{\nu})\}$ are regarded as discrete points in the deformation space $U/(\mathbb{Z}/2\mathbb{Z})^m \cong V/(\mathbb{Z}/2\mathbb{Z})^m$. Thus the discrete set of complete hyperbolic manifolds $\{M(\boldsymbol{\nu})\}$ are linked together in the connected space $V/(\mathbb{Z}/2\mathbb{Z})^m$.

3.7.3 Geometry of the Hyperbolic Manifolds Obtained by Dehn Fillings

In the hyperbolic Dehn filling Theorem 3.7.1, the complete hyperbolic manifolds $M(\boldsymbol{\nu}) = M(\nu_1, \dots, \nu_m)$ *geometrically converge* to the original complete hyperbolic manifold $\operatorname{int} M$ as $\boldsymbol{\nu} = (\nu_1, \dots, \nu_m) \to \infty = (\infty, \dots, \infty)$ [300, Section 5.11]. Namely, there are positive numbers $\epsilon(\boldsymbol{\nu})$ converging to 0 as $\boldsymbol{\nu} \to \infty$, and numbers $k(\boldsymbol{\nu}) > 1$ converging to 1 as $\boldsymbol{\nu} \to \infty$, such that there is a $k(\boldsymbol{\nu})$-bi-Lipschitz diffeomorphism

$$\phi_{\boldsymbol{\nu}} : M(\boldsymbol{\nu})_{\geq \epsilon(\boldsymbol{\nu})} \to (\operatorname{int} M)_{\geq \epsilon(\boldsymbol{\nu})}$$

between that $\epsilon(\boldsymbol{\nu})$-thick parts. This in particular implies that the lengths of core loops of the attached solid tori in $M(\boldsymbol{\nu})$ converge to 0 as $\boldsymbol{\nu} \to \infty$. This fact plays an essential role in various researches, including [27, 140, 172, 177, 269].

This also implies that the volumes $\operatorname{vol}(M(\boldsymbol{\nu}))$ of the Dehn filled manifolds converge to the volume $\operatorname{vol}(\operatorname{int} M)$ of the original hyperbolic manifold as $\boldsymbol{\nu} \to \infty$. Moreover, Thurston [300] proved, by using the Gromov norm (cf. Sect. 3.8.4), that $\operatorname{vol}(M(\boldsymbol{\nu}))$ is strictly smaller than $\operatorname{vol}(M)$ if $\boldsymbol{\nu} \neq \infty$. This is refined to quantitative estimates of $\operatorname{vol}(M(\boldsymbol{\nu}))$ by Neumann–Zagier [242], Hodgson–Kerckhoff [136] and Futer–Kalfagianni–Purcell [95].

Gromov and Thurston obtained the following result, by constructing a Riemannian metric of negative curvature on $M(\boldsymbol{\nu})$, when each surgery curve is "sufficiently long", by modifying the complete hyperbolic metric of int M (see [26] for a detailed proof).

Theorem 3.7.6 (The 2π-Theorem) *Let M be an orientable complete hyperbolic 3-manifold of finite volume, and let $C_1, \dots, C_m$ be disjoint torus cusps of M. Suppose ν_i is a slope on ∂C_i represented by a geodesic with length $> 2\pi$ with respect to the Euclidean metric. Then $M(\nu_1, \dots, \nu_m)$ has a Riemannian metric of negative curvature.*

The metric on $M(\nu_1, \dots, \nu_m)$ outside the filling solid tori is identical to the hyperbolic metric on $M - \cup_{j=1}^{m} C_j$. The geometrization theorem (Theorem 3.3.3) established by Perelman guarantees that the resulting manifold $M(\nu_1, \dots, \nu_m)$ is actually hyperbolic.

The 2π-theorem was refined to the 6-theorem by Agol [5] and Lackenby [176], and it plays a key role in the study of exceptional surgeries (see the next subsection).

3.7.4 Exceptional Surgeries

For a given hyperbolic knot K in S^3, or more generally an orientable complete hyperbolic manifold with one cusp, there are only finitely many exceptional slopes ν which produce non-hyperbolic manifolds. For example, the figure-eight knot has 10 exceptional slopes (Example 3.7.3(1)). In the survey [112], Gordon proposed various interesting conjectures, including one which says that 10 is the largest possible number of exceptional slopes of a hyperbolic knot complement.

The natural and important problem of determining exceptional surgery slopes has attracted attention of many mathematicians, and an enormous amount of research grew out of this problem, including:

- the 2π-theorem of Gromov–Thurston [118] and its improvement to the 6-theorem by Agol [5] and Lackenby [176],
- the cyclic surgery theorem by Culler–Gordon–Luecke–Shalen [76], obtained by combining two different kinds of arguments, namely (i) arguments using the $\mathrm{SL}(2, \mathbb{C})$-character varieties (cf. Sect. 3.11.3) and (ii) combinatorial, graph-theoretic analysis of the intersection of two incompressible, planar surfaces in knot exteriors,
- study of finite surgery by Boyer–Zhang [46, 47] and Ni–Zhang [243], by mainly using the $\mathrm{SL}(2, \mathbb{C})$-character varieties (Heegaard Floer homology and the Casson–Walker invariant are also used in [243]),
- the proof of the Property R conjecture by Gabai [102], by using taut foliations,
- a universal upper bound of the number of exceptional slopes by Hodgson–Kerckhoff [136, 137], by developing deformation theory of hyperbolic structures (cf. [135]).

- the optimal universal upper bound, 10, on the number of exceptional slopes of a one-cusped hyperbolic manifold by Lackenby–Meyerhoff [178] (see Agol [6] for related work),
- the optimal universal upper bound, 8, on the geometric intersection numbers of pairs of exceptional slopes of one-cusped hyperbolic manifolds by Lackenby–Meyerhoff [178],
- the complete classification of exceptional surgeries on hyperbolic alternating knots by Ichihara–Masai [145], building on a result of [175] and through computer-aided verified computation [143] using a super-computer.

The last three results give affirmative answers to some conjectures in [112]. See the survey articles [45, 112, 113] for background and further information.

As for *Seifert surgeries* of knots, namely surgeries which produce Seifert fibered spaces, Deruelle, Miyazaki and Motegi [77] embarked on the project to understand the whole shape of relationships among all such surgeries, and various interesting results are obtained in this direction.

Among Seifert surgeries, *lens space surgeries* are particularly interesting. Berge [21] presented a conjecturally complete list of lens space surgery on knots in S^3. Based on Berge's conjecture, Goda and Teragaito [108] conjectured that if a p-surgery on a hyperbolic knot K produces a lens space then K is a fibered knot and its genus g satisfies the inequality $2g + 8 \leq |p| \leq 4g - 1$. (Note that by the cyclic surgery theorem p is an integer.) Rasmussen [259] attacked this problem by using the Heegaard Floer homology, and obtained the estimate $|p| \leq 4g + 3$. This in fact relies on the fact that lens spaces belong to larger class of spaces, known as *L-spaces*, which are rational homology 3-spheres with the "simplest Heegaard–Floer homology" (see Ozsváth–Szabó [248]). See Greene [116] and references therein for further information on L-space surgery, and see the reviews [153, 249] for the background.

A nice overall survey (in Japanese) on surgery was recently written by Motegi [226], and its English translation will appear soon. This survey is strongly recommended.

3.8 Volumes of Hyperbolic 3-Manifolds

The volume is the most basic invariant of hyperbolic manifolds. After quickly recalling a method for calculating hyperbolic volumes, we explain (i) the Jorgensen-Thurston theory concerning the volume spectrum of hyperbolic 3-manifolds, (ii) results concerning small volume hyperbolic manifolds, (iii) relation to the Gromov norm, and finally (iv) the volume conjecture, which lies in the two innovations, hyperbolic geometry and quantum topology, in knot theory.

3.8.1 Calculation of Hyperbolic Volumes

We explain a method for calculating the volumes of hyperbolic 3-manifolds, which is implemented in SnapPea. The method depends on the fact that every hyperbolic 3-manifold M is obtained by hyperbolic Dehn filling on a cusped hyperbolic manifold, say M_0. This follows from the facts that the complement of a simple closed geodesic is a cusped hyperbolic manifold (see [277]) and that the shortest closed geodesic in M is simple. SnapPea usually succeeds in finding an ideal triangulation of the complete hyperbolic manifold M_0, which can be deformed into an ideal triangulation of the incomplete hyperbolic structure on M_0 whose completion yields the complete hyperbolic structure of M (cf. Sect. 3.7.2). Thus the calculation of $\mathrm{vol}(M)$ is reduced to that of the volumes of ideal tetrahedra.

Recall that the isometry type of an ideal tetrahedron is determined by its shape parameter $z \in \mathbb{H} \subset \mathbb{C}$, which in turn represent the similarity class of the Euclidean triangle with vertex set $\{0, 1, z\}$. Let α, β, γ be the inner angles of this triangle. Then the volume of the ideal tetrahedron $\Delta(z)$ of shape parameter z is given by the following formula:

$$\mathrm{vol}(\Delta(z)) = \Lambda(\alpha) + \Lambda(\beta) + \Lambda(\gamma),$$

where $\Lambda(\theta)$ is the *Lobachevsky function* defined by

$$\Lambda(\theta) = -\int_0^\theta \log|2\sin t| dt = \frac{1}{2}\sum_{n=1}^{\infty} \frac{\sin(2n\theta)}{n^2}.$$

The volume function $\mathrm{vol}(\Delta(z))$ takes the maximal value $V_{\mathrm{tet}} = 3\Lambda(\pi/3) = 1.0149416\ldots$ precisely at $z = \exp(\pi\sqrt{-1}/3)$, i.e., exactly when $\Delta(z)$ is a regular ideal tetrahedron. See [300, Chapters 6 and 7] for details.

3.8.2 Jørgensen–Thurston Theory for the Volumes of Hyperbolic 3-Manifolds

Let $V_n \subset \mathbb{R}_+$ be the ordered set consisting of the volumes of complete hyperbolic n-manifolds. If $n \neq 3$ then V_n is isomorphic to $\mathbb{N}$, by Gauss–Bonnet theorem for $n = 2$ and by Wang's theorem [319] for $n \geq 4$. For dimension $n = 3$, we have the following surprising theorem due to Jørgensen and Thurston (see [300]), which forms a sharp contrast to Wang's theorem.

Theorem 3.8.1 (Jørgensen–Thurston Theorem) *V_3 is a well-orderd closed set which is isomorphic to ω^ω. Moreover, the map*

$$\mathrm{vol} : \{\textit{complete hyperbolic } 3\textit{-manifolds of finite volume}\}/(\mathrm{isometry}) \to V_3$$

is finite to one.

This means that there is a smallest volume v_1, a next smallest volume v_2, and so on, and these are the volumes of closed hyperbolic 3-manifolds. The increasing sequence $v_1 < v_2 < \cdots < v_k < \cdots$ has a limit v_ω, and this is the volume of a complete hyperbolic 3-manifold with one cusp (cf. Sect. 3.7.3). There is a smallest volume $v_{\omega+1}$ bigger than v_ω, a second smallest volume $v_{\omega+2}$ bigger than $v_{\omega+1}$, and so on, and these are the volumes of closed hyperbolic 3-manifolds, and their limit $v_{2\omega}$ is the second smallest volume of a complete hyperbolic 3-manifold with one cusp. The increasing sequence $v_\omega < v_{2\omega} < \cdots < v_{k\omega} < \cdots$ has a limit v_{ω^2}, and this is the volume of a complete hyperbolic 3-manifold with two cusps, and so on.

The second statement of Theorem 3.8.1 says that the volume is "almost" a complete invariant of complete hyperbolic manifolds.

Of course, the volume is not a complete invariant. For example, the complements of the Kinoshita–Terasaka knot and the Conway knot have the same volume 11.21911773.... In fact, Ruberman [276] proved that the hyperbolic volume, more generally the Gromov invariant (cf. Sect. 3.8.4 below), is unchanged by mutation. Hodgson and Masai [138] studied the number $N(v)$ of orientable hyperbolic 3-manifolds with given volume $v \in V_3$: they constructed infinitely many $v \in V_3$ for which $N(v) = 1$, and proved the exponential growth of $N(v)$ by showing $N(4nV_{\mathrm{oct}}) \geq 2^n/(2n)$. See Chesebro–DeBlois [61] and Millichap [209] for related results.

3.8.3 Small Volume Hyperbolic Manifolds

It is a natural and important problem to determine the small volumes, such as v_1, v_ω, v_{ω^2}, etc. For the minimal volume v_{ω^n} of orientable complete hyperbolic 3-manifolds with n-cusps, the following results are established.

- Gabai–Meyerhoff–Milley (2009) [103]: The *Fomenko–Matveev–Weeks manifold*, which is obtained by (5, 2) and (5, 1) Dehn surgery on the Whitehaed link, has the smallest volume $v_1 = 0.94270736\ldots$.
- Cao–Meyerhoff (2001) [58]: The figure-eight knot complement and its sister, namely (5, 1)-filling on one component of the Whitehead link complement, have the volume $v_\omega = 2V_{\mathrm{tet}} = 2.02988\ldots$, where $V_{\mathrm{tet}} = 1.0149416\ldots$ is the volume of the regular ideal tetrahedron. The figure-eight knot is the orientation double cover of the Gieseking manifold, the non-orientable hyperbolic 3-manifold, which has the smallest volume among the all (orientable or not) complete non-compact hyperbolic 3-manifolds (see Adams [1]).
- Agol (2010) [7]: The Whitehaed link complement and the complement of the pretzel link $P(-2, 3, 8)$ have the volume $v_{\omega^2} = V_{\mathrm{oct}} = 3.66386\ldots$, where V_{oct} is the volume of regular ideal octahedron.
- Yoshida (2013) [327]: The complement of the minimally twisted hyperbolic 4-chain link has the volume $v_{\omega^4} = 2V_{\mathrm{oct}} = 7.32772\ldots$.

See the review [105], for further information. It should be noted that all of the above small volume hyperbolic manifolds are arithmetic (cf. [124], [240, Theorem 5.1] and Subection 3.9.3).

As is noted in [104, Introduction], Thurston had long promoted the idea that volume is a good measure of the complexity of a hyperbolic 3-manifold. In fact, in [300, the end of Chapter 6], he writes the following: *One gets a feeling that volume is a very good measure of the complexity of a link complement, and that the ordinal structure is really inherent in three-manifolds.* The following conjecture, due to Thurston, Weeks, Matveev–Fomenko and Mednykh–Vesnin, states the idea more rigorously, and the results presented above can be regarded as partial answers to this conjecture.

Conjecture 3.8.2 The complete low-volume hyperbolic 3-manifolds can be obtained by filling cusped hyperbolic 3-manifolds of small topological complexity.

To end this subsection, we explain another approach to Thurston's idea above, by using the notions of *shadows* of 3 and 4-manifolds introduced by Turaev [310, 311]. Costantino and Thurston [67] introduced the *shadow complexity* $\mathrm{sc}(M)$ of a compact orientable 3-manifold M with (possibly empty) toral boundary, and proved the following estimate of the Gromov norm $||M||$ (cf. Sect. 3.8.4, below):

$$\frac{V_{\mathrm{tet}}}{2V_{\mathrm{oct}}}||M|| \le \mathrm{sc}(M) \le C||M||^2 \quad \text{for some universal constant } C.$$

In the same paper, they implicitly introduced the notion of *stable map complexity* and studied its relation between (branched) shadow complexity as well. Ishikawa and Koda [147] showed the two complexities are actually equal, and moreover, using the result of [95], they gave an elaborate refinement of the above (left) inequality when M is hyperbolic. They also defined the *branched shadow complexity* $\mathrm{bsc}(M, L)$ for a link L in a compact orientable 3-manifold M with (possibly empty) toral boundary, and gave a complete characterization of hyperbolic links L in S^3 with $\mathrm{bsc}(S^3, L) = 1$.

3.8.4 Gromov Norm

In [117], Gromov introduced the notion of *simplicial volume* $||M||$ of a closed manifold M as follows, using real singular homology:

$$||M|| := \inf\{||z|| \mid z \text{ is a singular cycle representing the fundamental class } [M]\}$$

Here, for a (real) singular chain $z = \sum_j a_j \sigma_j$, its norm $||z||$ is defined as the sum $\sum_j |a_j|$ of the absolute values of its coefficients. He used it to estimate the "*minimal volume*" of closed smooth manifold (see [117]). Building on this work, Thurston [300, Chapter 6] defined the *Gromov norm* $||M||$ of a compact orientable 3-manifold

M with (possibly empty) toral boundary as follows:

$$||M|| := \lim_{\epsilon \to 0} \inf\{||z|| \mid z \text{ is a singular chain representing } [M, \partial M] \text{ and } ||\partial z|| \leq \epsilon\}$$

He then proved the following.

(1) If M is hyperbolic (and hence $\operatorname{int} M$ admits a complete hyperbolic structure of finite volume), then

$$||M|| = \frac{1}{V_{\text{tet}}} \operatorname{vol}(\operatorname{int} M).$$

(2) If M is a Seifert fibered space, then $||M|| = 0$.
(3) Let T be a torus embedded in $\operatorname{int} M$ and let M_T be the manifold obtained by cutting M along T. Then $||M|| \leq ||M_T||$.

Soma [294] proved that when T is incompressible, equality holds in (3) and that similarly equality holds for an incompressible annulus properly embedded in M. He then defined, for a link L in S^3, the *Gromov invariant* $||L||$ of L by $||L|| = ||E(L)||$, and obtained the following theorem.

Theorem 3.8.2 (Soma) *For a link L in S^3, the following hold.*

(1) If L is a split sum of two links L_1 and L_2, then $||L|| = ||L_1|| + ||L_2||$.
(2) If L is a connected sum of two links L_1 and L_2, then $||L|| = ||L_1|| + ||L_2||$.
(3) Suppose L is a non-splittable link, and let $\{M_j\}$ be the hyperbolic pieces of the JSJ decomposition of $E(L)$. Then

$$||L|| = \sum_j ||M_j|| = \frac{1}{V_{\text{tet}}} \sum_j \operatorname{vol}(\operatorname{int} M_j).$$

3.8.5 The Volume Conjecture

In addition to the revolution caused by William Thurston, knot theory has experienced yet another revolution through the discovery of the Jones polynomial by Vaughan Jones [151]. The Volume Conjecture, first stated by Rinat Kashaev [158] and then reformulated and expanded by Hitoshi Murakami and Jun Murakami [230], provoked deep interaction between the two innovations, hyperbolic geometry and quantum topology.

The conjecture says that the hyperbolic volume of a hyperbolic knot in S^3 (more generally, the Gromov norm of a knot in S^3) is determined by the asymptotic behavior of *Kashaev's invariant* $\langle K \rangle_N$, which is shown by [230] to coincide with the evaluation, $J_N(K)$, of the *N-colored Jones polynomial* (with a certain normalization) at the primitive N-th root of unity $\exp(2\pi i/N)$.

Conjecture 3.8.4 (Volume Conjecture) For any knot K in S^3, the following holds:

$$||K|| = \frac{2\pi}{V_{\text{tet}}} \lim_{N\to\infty} \frac{\log|J_N(K)|}{N}.$$

In particular, if K is a hyperbolic knot, the following holds:

$$\text{vol}(S^3 - K) = 2\pi \lim_{N\to\infty} \frac{\log|J_N(K)|}{N}.$$

Moreover, H. Murakami and J. Murakami proved that Kashaev's invariant also coincides with an evaluation of the *generalized Alexander polynomial* defined by Y. Akutsu, T. Deguchi and T. Ohtsuki [16]. They say in [230, page 86] that *the set of colored Jones polynomials and the set of generalized Alexander polynomials of Akutsu–Deguchi–Ohtsuki intersect at Kashaev's invariants.*

Furthermore, H. Murakami, J. Murakami, M. Okamoto, T. Takata and Y. Yokota [232] proposed the following complexification of Kashaev's conjecture:

Conjecture 3.8.5 (Complexification of the Volume Conjecture) For any hyperbolic knot K in S^3, the following holds:

$$\text{vol}(S^3 - K) + \sqrt{-1}\text{CS}(S^3 - K) = 2\pi \lim_{N\to\infty} \frac{\log J_N(K)}{N}.$$

In the above conjecture $\text{CS}(S^3 - K)$ denotes the *Chern–Simons invariant* of $S^3 - K$ (see [60, 208, 325]). For further information, see the surveys [228, 229] and the recently published book [231].

3.9 Commensurability and Arithmetic Invariants of Hyperbolic Manifolds

In [300, Sections 6.7 and 6.8], Thurston studied the commensurability relation among hyperbolic knot/link complements, and gave various commensurable and incommesurable examples. This work has promoted intimate interaction between knot theory and number theory. In this section, we recall basic arithmetic invariants of commensurability classes of Kleinian groups, and describe application to knot theory. We also describe the dichotomy between arithmetic groups and non-arithmetic groups found by Margulis and Borel. In the final subsection, we recall the solution due to Gehring, Marshal and Martin of the 3-dimensional Siegel problem to determine the minimal volume of hyperbolic orbifolds, lying emphasis on the role of arithmetic groups. For further information on the topic of this section, see the textbook Maclachlan–Reid [189].

3.9.1 Commensurability Classes and Invariant Trace Fields

Two Kleinian groups Γ_1 and Γ_2 are said to be *commensurable* if there is a conjugate, $\Gamma_2^g := g^{-1}\Gamma_2 g$ ($g \in \mathrm{PSL}(2,\mathbb{C})$) such that $\Gamma_1 \cap \Gamma_2^g$ has finite index both in Γ_1 and Γ_2^g. This is equivalent to the condition that the two hyperbolic manifolds $M_1 = \mathbb{H}^3/\Gamma_1$ and $M_2 = \mathbb{H}^3/\Gamma_2$ are *commensurable*, i.e., there is a hyperbolic manifold which is a finite covering of both M_1 and M_2. As is explained in Sect. 3.6.1, the canonical decomposition provides us an efficient method for checking if two (cusped) hyperbolic manifolds are isometric. But, the method is not directly applicable for checking commensurability, though there is a nice application of the canonical decomposition for the commensurability problem (see [110] and Sect. 3.9.3).

Number theory enables us to define a very useful invariant of the commensurability classes of Kleinian groups of cofinite volume. Let $M = \mathbb{H}^3/\Gamma$ be an orientable complete hyperbolic manifold of finite volume. Consider the set $\mathrm{tr}\,\Gamma = \{\pm\,\mathrm{tr}(\gamma) \mid \gamma \in \Gamma\} \subset \mathbb{C}$ and the field $\mathbb{Q}(\mathrm{tr}\,\Gamma)$ generated by the set. (Note that the trace $\mathrm{tr}\,\gamma$ for $\gamma \in \mathrm{PSL}(2,\mathbb{C}) \cong \mathrm{Isom}^+(\mathbb{H}^3)$ is well-defined up to sign.) This is called the *trace field* of the Kleinian group Γ. It follows from the rigidity theorem that the trace field $\mathbb{Q}(\mathrm{tr}\,\Gamma)$ has finite degree over $\mathbb{Q}$, i.e., it is a *number field*. By the rigidity theorem again, this is an invariant of the topological space M.

Though the trace field $\mathbb{Q}(\mathrm{tr}\,\Gamma)$ itself is not, in general, an invariant of the commensurability class, it provides us with a very useful commensurability invariant as follows. Let $\Gamma^{(2)}$ be the subgroup of Γ generated by $\{\gamma^2 \mid \gamma \in \Gamma\}$. Then $\Gamma^{(2)}$ is normal in Γ and $\Gamma/\Gamma^{(2)}$ is a finite abelian group which is a direct sum of order 2 cyclic groups. The following theorem was proved by Reid [261].

Theorem 3.9.1 *Let Γ be a Kleinian group of finite covolume. Then $\mathbb{Q}(\mathrm{tr}\,\Gamma^{(2)})$ is an invariant of the commensurability class of Γ. Moreover*

$$\mathbb{Q}(\mathrm{tr}\,\Gamma^{(2)}) = \mathbb{Q}(\{(\mathrm{tr}\,\gamma)^2 \mid \gamma \in \Gamma\}).$$

The field $\mathbb{Q}(\mathrm{tr}\,\Gamma^{(2)})$ is denoted by $k(\Gamma)$ and is called the *invariant trace field* of Γ. By [240, Corollary 2.3], if $M = \mathbb{H}^3/\Gamma$ is a knot complement (or more generally, the complement of a link in a $\mathbb{Z}/2\mathbb{Z}$-homology sphere) then $k(\Gamma) = \mathbb{Q}(\mathrm{tr}\,\Gamma)$: thus in this case the trace field itself is an invariant of the commensurability class.

If M is a cusped hyperbolic manifold which admits an ideal triangulation into the hyperbolic ideal tetrahedra $\{\Delta(z_1), \dots, \Delta(z_t)\}$, then the following holds [240, Theorem 2.4]:

$$k(\Gamma) = \mathbb{Q}(z_1, \dots, z_t).$$

The *invariant quaternion algebra* of Γ is the $k(\Gamma)$-algebra of the 2×2 matrix algebra $M_2(\mathbb{C})$ generated over $k(\Gamma)$ by the elements of $\Gamma^{(2)}$. It is denoted by $A(\Gamma)$. This algebra is also an invariant of the commensurability class of Γ. Both $k(\Gamma)$ and $A(\Gamma)$ are preserved by mutation (see [239]).

The computer program "Snap" calculates various arithmetic invariants including the invariant trace field and the invariant quaternion algebra (see [68]).

3.9.2 Commensurators and Hidden Symmetries

For a Kleinian group Γ of cofinite volume, the *commensurator* of Γ is defined by

$$\mathrm{Comm}(\Gamma) = \{g \in \mathrm{Isom}\,\mathbb{H}^3 \mid [\Gamma; \Gamma \cap \Gamma^g] < \infty\},$$

and its orientation-preserving subgroup is denoted by $\mathrm{Comm}^+(\Gamma)$. The commensurator $\mathrm{Comm}(\Gamma)$ is identified with the group of equivalence classes of virtual automorphisms of Γ. A *virtual automorphism* of Γ is an isomorphism $\phi : \Gamma_1 \to \Gamma_2$ between subgroups of finite index in Γ, and two virtual automorphisms are defined to be equivalent if they agree on some subgroup of Γ of finite index. A virtual automorphism represents an isometry between two finite coverings $\mathbb{H}^3/\Gamma_1$ and $\mathbb{H}^3/\Gamma_2$ of the hyperbolic manifold $M = \mathbb{H}^3/\Gamma$. It is called a *hidden symmetry* of M if it is not a lift of an isometry of M. By a *hidden symmetry* of a hyperbolic knot in S^3, we mean a hidden symmetry of the knot complement. We can see as follows that the figure-eight knot K has a hidden symmetry. Recall that $S^3 - K = \mathbb{H}^3/\Gamma$ has an ideal triangulation consisting of two copies of the regular ideal tetrahedron $\Delta(\omega)$ with $\omega = \exp(\frac{\pi\sqrt{-1}}{3})$. This implies that the invariant trace field $k(\Gamma)$ is equal to $\mathbb{Q}(\omega) = \mathbb{Q}(\sqrt{-3})$. Moreover, we see that Γ is a subgroup $\mathrm{PSL}(2, \mathcal{O}_3)$ of finite index (actually equal to 24), where $\mathcal{O}_3$ is the ring of integers of the number field of $\mathbb{Q}(\sqrt{-3})$. This implies that $\mathrm{PGL}(2, \mathbb{Q}(\sqrt{-3}))$ belongs to the commensurator subgroup of Γ. In fact, we have $\mathrm{Comm}^+(\Gamma) = \mathrm{PGL}(2, \mathbb{Q}(\sqrt{-3}))$. Since $\mathrm{PGL}(2, \mathbb{Q}(\sqrt{-3}))$ is dense in $\mathrm{PSL}(2, \mathbb{C})$, the normalizer of Γ must be a proper subgroup of $\mathrm{Comm}^+(\Gamma) = \mathrm{PGL}(2, \mathbb{Q}(\sqrt{-3}))$. Hence Γ (and so the figure-eight knot) has a hidden symmetry.

In addition to the figure-eight knot, the two dodecahedral knots of Aitchison and Rubinstein [12] admit hidden symmetries, and these three are the only known such knots. Neumann and Reid [240, Question 1] conjecture that they are all. For results related to the conjecture, see [37, 38, 210, 264] and references therein.

3.9.3 Arithmetic Versus Non-arithmetic

The above explanation for the existence of hidden symmetries of the figure-eight knot is based on the fact that the figure-eight knot group belongs to the particularly nice family of Kleinian groups, called *arithmetic groups*. For the definition of arithmetic groups, see the textbook [263] or the course notes [238, Chapter 3, Section 3]. If we restrict our attention to a cofinite volume Kleinian group Γ such that $M = \mathbb{H}^3/\Gamma$ has a cusp, then Γ is arithmetic if and only if Γ is conjugate to a subgroup of $\mathrm{PGL}(2, \mathcal{O}_d)$ for some positive integer d. Here $\mathcal{O}_d$ is the ring of integers of the number field of $\mathbb{Q}(\sqrt{-d})$. In this case, we have $k(\Gamma) = \mathbb{Q}(\sqrt{-d})$ and $A(\Gamma) = M_2(\mathbb{Q}(\sqrt{-d}))$, and the invariant trace field $k(\Gamma)$ is the complete commensurability invariant of the arithmetic group Γ. However, most cusped hyperbolic manifolds are non-arithmetic; in particular, the figure-eight knot is the unique hyperbolic knot in S^3 whose complement is arithmetic (see Reid [262]).

Margulis [192] (see also Borel [43]) establised the following striking dichotomy between the arithemtic Kleinian groups and non-arithmetic Kleinian groups.

Theorem 3.9.2 *Let Γ be a cofinite volume Kleinian group. Then the following hold.*

(1) Γ *is non-arithmetic if and only if* Γ *has finite index in* $\mathrm{Comm}^+(\Gamma)$. *In this case,* $\mathrm{Comm}^+(\Gamma)$ *is the unique maximal element in the commensurability class of* Γ.
(2) Γ *is arithmetic if and only if* $\mathrm{Comm}^+(\Gamma)$ *is dense in* $\mathrm{PSL}(2, \mathbb{C})$. *In this case, there are infinitely many maximal elements in the commensurability class of* Γ.

The first assertion of Theorem 3.9.2 shows that the commensurability class of a non-arithmetic cofinite volume Kleinian group Γ is particularly simple, namely it consists only of conjugates of finite index subgroups of the Kleinian group $\mathrm{Comm}^+(\Gamma)$. In terms of orbifolds, this means that two non-arithmetic orientable hyperbolic 3-manifolds M_1 and M_2 are commensurable if and only if they cover a common orbifold. Based on this fact and by using the Epstein–Penner decomposition [82] and finiteness of Epstein–Penner decompositions of a given cusped hyperbolic manifolds (see Akiyoshi [13]), Goodman–Heard–Hodgson [110] gave a practical algorithm for determining when two cusped hyperbolic non-arithmetic 3-manifolds are commensurable. Their algorithm is based on the fact that two cusped hyperbolic n-manifolds M and M' cover a common orbifold if and only if they admit Epstein–Penner decompositions lifting to isometric tilings of $\mathbb{H}^n$ (see [110, Theorem 2.4]). Their algorithm is implemented in a computer program, which enabled them to determine the commensurability classes of the complements of all hyperbolic knots and links up to 12 crossings. In particular, they have shown that the complements of the Kinoshita–Terasaka knot and the Conway knot belong to different commensurability classes, even though they share the same invariant trace fields and invariant quaternion algebras. See Chesebro–DeBlois [61] and Millichap–Worden [210] for related works.

The second assertion of Theorem 3.9.2 shows that the commensurability class of arithmetic Kleinian groups is very complicated. Walter Neumann describes a geometric way of thinking of this situation as follows, in his course notes [238, Chapter 3, Section 6].

> A Kleinian (or Fuchsian) group is the symmetry group of some "pattern" in $\mathbb{H}^3$ (respectively $\mathbb{H}^2$). This pattern might just be a tessellation — for instance, a tesselation by fundamental domains, or it might be an Escher-style drawing. If one superposes two copies of this pattern, displaced with respect to each other, one will usually get a pattern which no longer has a Kleinian (or Fuchsian) symmetry group in our sense — the symmetry group has

> become too small to have finite volume quotient. But in the arithmetic case — and only in this arithmetic case — one can always change the displacement very slightly to make the superposed pattern have a symmetry group that is of finite index in the original group.

In the course notes [238], we can also find a beautiful introduction to the idea of Scissor congruence, with a historical background which goes back to Euclid, Dehn and Hilbert. For more details of this important topic, see [237].

3.9.4 Siegel's Problem and Arithmetic Manifolds

In Sect. 3.8.3, we surveyed various important results concerning small volume hyperbolic 3-manifolds. It is equally natural and important to study small volume hyperbolic orbifolds. In 1943, Siegel [291, 292] posed the problem of identifing the infimum

$$\mu(n) = \inf_{\Gamma} \mathrm{vol}(\mathbb{H}^n/\Gamma)$$

where the infimum is taken over the *lattices* $\Gamma < \mathrm{Isom}^+ \mathbb{H}^n$, i.e., discrete subgroups of cofinite volume. Siegel solved the problem in dimension 2, by showing that the (2, 3, 7)-triangle group is the unique Fuchsian group of minimal coarea

$$\mu(2) = 2\pi \left| \frac{1}{2} + \frac{1}{3} + \frac{1}{7} - 1 \right| = \frac{\pi}{21}.$$

In 1986, Kazhdan and Margulis [163] made an important contribution to the Siegel problem, by proving that $\mu(n)$ is positive and attained for each n.

Arithmetic groups play a crucial role in the study of the Siegel problem. One big reason is that, due to formulas of Borel [43], various explicit calculations can be made for arithmetic Kleinian groups. According to Gaven Martin [195], another reason is that *it turns out that nearly all the extremal problems one might formulate are realised by arithmetic groups, perhaps the number theory forcing additional symmetries in a group and therefore making it "smaller" or "tighter"*.

After a long term collaboration, Gehring, Marshal and Martin [106, 193] finally solved the 3-dimensional *Siegel problem*.

Theorem 3.9.3 *The minimum $\mu(3)$ of the volumes of hyperbolic 3-orbifolds is*

$$\mu(3) = \mathrm{vol}(\mathbb{H}^3/\Gamma_0) = 275^{3/2} 2^{-7} \pi^{-6} \zeta_k(2) \sim 0.03905\ldots,$$

where ζ_k is the Dedekind zeta function of the underlying number field $\mathbb{Q}(\gamma_0)$, with γ_0 a complex root of $\gamma^4 + 6\gamma^3 + 12\gamma^2 + 9\gamma + 1$, of discriminant -275. Here Γ_0 is an arithmetic Kleinian group obtained as a $\mathbb{Z}/2\mathbb{Z}$-extension of the index 2 orientation-preserving subgroup of the group generated by reflection in the faces of *the* 3-5-3-*hyperbolic Coxeter tetrahedron. The group* Γ_0 *is generated by two elliptic elements, one of order* 2 *and the other of order* 5.

Remark 3.9.4 The quotient orbifold $\mathcal{O}_0 = \mathbb{H}^3/\Gamma_0$ is as illustrated in Fig. 3.9, where the blue eyeglasses represent the generating pair. This orbifold is obtained from the "Heckoid orbifold $H(1/4; 5/2)$" in Fig. 3.10 by an orbifold surgery. Here a Heckoid orbifold is a hyperbolic 3-orbifold whose orbifold fundamental group is a *Heckoid group*, which is a Kleinian group generated by two parabolic transformations introduced by Riley [274] as an analogy of Hecke groups and formulated by [181]. Heckoid orbifolds are also intimately related to 2-bridge links. As noted by Martin [195, 196], most of small volume 3-orbifolds arise from 2-bridge links.

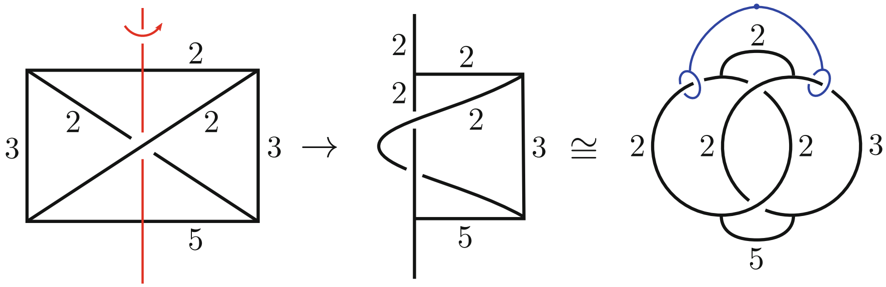


Fig. 3.9 The minimal volume 3-orbifold $\mathbb{H}^3/\Gamma_0$. The blue eyeglass frame represents the generating pair of Γ_0 consisting of elliptic elements

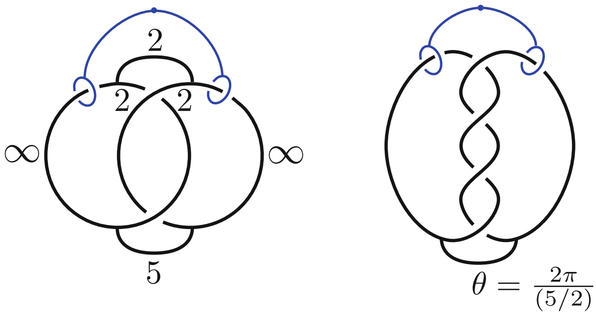


Fig. 3.10 The left picture illustrates the Heckoid orbifold $H(1/4; 5/2)$ and the parabolic generating pair of the Heckoid group $\pi_1^{\mathrm{orb}}(H(1/4; 5/2))$. The Heckoid group is identified with the image of the holonomy representation of the hyperbolic cone manifold $C\left(1/4; \frac{2\pi}{(5/2)}\right)$ depicted by the right picture

See the survey by Martin [195] for backgrounds and details concerning the 3-dimensional Siegel problem, and the surveys by Belolipetsky [20] and Kellerhals [164] for the higher dimensional Siegel problem.

The above theorem has the following application to finite group actions on hyperbolic 3-manifolds.

Corollary 3.9.5 *Let M be an orientable complete hyperbolic 3-manifold of finite volume, and let G be a finite group acting on M effectively and orientation-preservingly. Then*

$$|G| \leq \frac{\operatorname{vol} M}{\mu(3)}.$$

A refinement of this corollary for hyperbolic knot complements can be found in [96, Theorem 4.14].

3.10 Flexibility of Complete Hyperbolic Manifolds: Deformation Theory of Hyperbolic Structures

Let M be a complete hyperbolic manifold homeomorphic to the interior of a compact orientable 3-manifold $\bar{M}$. If $\partial\bar{M}$ is a (possibly empty) union of tori, then $\mathrm{vol}(M) < \infty$ and so by the Mostow–Prasad rigidity theorem the complete hyperbolic structure on M is unique. However, when $\partial\bar{M}$ contains a component different from a torus, the complete hyperbolic structure of M admits a nontrivial deformation, and there is a rich and deep deformation theory. This deformation theory is one of the central themes in Kleinian group theory and it plays a crucial role in the proof of the geometrization theorem of Haken manifolds. In particular, the existence of complete hyperbolic structures on surface bundles over the circle, e.g. the complements of hyperbolic fibered knots, was established as a consequence the double limit theorem [308, Theorem 4.1] concerning the deformation space of hyperbolic structures on $\Sigma \times \mathbb{R}$ where Σ is a (fiber) surface. The idea of a Cannon–Thurston map, a $\pi_1(\Sigma)$-equivariant sphere filling curve, grew out of this construction.

On the other hand, Agol [8] proved that a hyperbolic punctured surface bundle over the circle admits a very special topological ideal triangulation, called a *veering triangulation*, which is canonical in the sense that it is determined by the fiber structure. It was revealed by Guéritaud [122] that the veering triangulation is intimately related to the Cannon–Thurston map.

The purpose of this section is (i) to give an introduction to the deformation theory of Kleinian groups and its relation to the hyperbolic structures of surface bundles over the circle, and (ii) to explain Cannon–Thurston maps and veering triangulations. For further information on deformation theory, see Otal [246, 247], Matsuzaki–Taniguchi [197], Kapovich [157], Ohshika [245] and Marden [190, 191].

3.10.1 Convex Cores and Conformal Boundaries of Hyperbolic Manifolds

In this subsection, we recall the basic concepts of convex cores and conformal boundaries of hyperbolic manifolds.

Though the action of a Kleinian group Γ on $\mathbb{H}^3$ is properly discontinuous, the action of Γ on $\partial\mathbb{H}^3$ does not have this property. To see this, pick a point $x \in \mathbb{H}^3$ and consider its orbit Γx. Of course the orbit is discrete in $\mathbb{H}^3$. But, it has nonempty accumulation points in the 3-ball $\mathbb{H}^3 \cup \partial\mathbb{H}^3$ (provided that Γ is not a finite group). The set of all accumulation points is independent of the choice of x and forms a Γ-invariant closed set in $\partial\mathbb{H}^3$. This set is denoted by $\Lambda(\Gamma)$ and is called the *limit set* of Γ. The action of Γ on $\Lambda(\Gamma)$ is not properly discontinuous and is chaotic. The complement $\Omega(\Gamma) := \partial\mathbb{H}^3 - \Lambda(\Gamma)$ is called the *domain of discontinuity* of Γ, and it is a (possibly empty) maximal open domain in $\partial\mathbb{H}^3$ on which Γ acts properly discontinuously.

The *convex core* C_M of a complete hyperbolic manifold $M = \mathbb{H}^3/\Gamma$ is defined as the quotient $C(\Lambda(\Gamma))/\Gamma$, where $C(\Lambda(\Gamma))$ is the convex hull in $\mathbb{H}^3$ of the limit set $\Lambda(\Gamma)$. Note that any closed geodesic in M corresponds to a conjugacy class of a hyperbolic element of Γ and that the endpoints of its axis are contained in $\Lambda(\Gamma)$: this implies that the axis is contained in $C(\Lambda(\Gamma))$ and so the closed geodesic is contained in C_M. In fact, C_M is the smallest locally convex closed subset of M which contains all closed geodesics of M. The convex core C_M is also characterized as the smallest locally convex submanifold of M whose inclusion is a homotopy equivalence.

On the other hand, since the action of Γ on $\partial\mathbb{H}^3$ (and hence on $\Omega(\Gamma)$) is conformal, the quotient space $\partial_\infty M := \Omega(\Gamma)/\Gamma$ has a natural conformal structure and forms the boundary of the *Klein manifold* $(\mathbb{H}^3 \cup \Omega(\Gamma)/\Gamma)$. The Riemann surface $\partial_\infty M = \Omega(\Gamma)/\Gamma$ is called the *conformal boundary* of M.

Example 3.10.1 (Infinite Cyclic Kleinian Group) For the infinite cyclic Kleinian group Γ generated by a hyperbolic transformation $A(z) = az$ with $|a| \neq 1$ in Example 3.5.2(1), the convex core of the quotient hyperbolic manifold $\mathbb{H}^3/\Gamma \cong \mathrm{int}(D^2 \times S^1)$ is equal to the core closed geodesic $(0 \times \mathbb{R}_+)/\Gamma$, and the conformal boundary is the torus $(\mathbb{C} - \{0\})/(z \sim az)$.

In the remainder of this section $\Sigma \cong \mathrm{int}\,\Sigma_{g,b}$ denotes the closed orientable surface of genus g with b punctures, and with negative Euler characteristic.

Definition 3.10.2 (Type-Preserving Representation) A representation $\rho : \pi_1(\Sigma) \cong \pi_1(\Sigma_{g,b}) \to \mathrm{Isom}^+\,\mathbb{H}^3$ is *type-preserving* if it satisfies the following conditions.

(1) ρ maps peripheral elements (elements represented by boundary loops of $\Sigma_{g,b}$) to parabolic elements.
(2) ρ is *irreducible*, i.e., $\rho(\pi_1(\Sigma))$ does not have a common fixed point on $\partial\mathbb{H}^3$.

Example 3.10.3 (Fuchsian Group) The surface Σ admits a complete hyperbolic structure of finite area $\pi|\chi(\Sigma)|$. Pick a complete hyperbolic metric on Σ and let $\rho_0 : \pi_1(\Sigma) \to \mathrm{Isom}^+ \mathbb{H}^2$ be the holonomy representation. Then it is discrete, faithful and type-preserving, and its image $\Gamma_0 = \rho_0(\pi_1(\Sigma))$ is a Fuchsian group. The limit set of the Fuchsian group Γ_0 is equal to $\partial\mathbb{H}^2$. Regard Γ_0 as a Kleinian group, i.e., a discrete subgroup of $\mathrm{Isom}^+ \mathbb{H}^3$. Then the limit set $\Lambda(\Gamma_0)$ is the round circle $\partial\mathbb{H}^2$ in $\partial\mathbb{H}^3$, where $\mathbb{H}^2 (= \mathbb{R} \times \mathbb{R}_+ \subset \mathbb{C} \times \mathbb{R}_+ = \mathbb{H}^3)$ is the hyperplane of $\mathbb{H}^3$ invariant by Γ. The Kleinian manifold $(\mathbb{H}^3 \cup \Omega(\Gamma_0))/\Gamma_0$ is homeomorphic to the product of Σ and the closed interval $[-\infty, \infty]$, and the convex core is identified with $\Sigma \times 0$.

Example 3.10.4 (Quasifuchsian Group) The Fuchsian representation $\rho_0 : \pi_1(\Sigma) \to \mathrm{PSL}(2, \mathbb{R})$ in the previous example admits a nontrivial deformation into a faithful discrete type-preserving $\mathrm{PSL}(2, \mathbb{C})$-representation ρ, such that $(\mathbb{H}^3 \cup \Omega(\Gamma))/\Gamma \cong \Sigma \times [-\infty, \infty]$ where $\Gamma = \rho(\pi_1(\Sigma))$. This condition is equivalent to the condition that the limit set $\Lambda(\Gamma)$ is a topological circle. A Kleinian group isomorphic to $\pi_1(\Sigma)$ satisfying this condition is called a *quasifuchsian group* and the holonomy representation is called a *quasifuchsian representation*. Generically, a quasifuchsian group is not conjugate to a Fuchsian group in $\mathrm{PSL}(2, \mathbb{C})$, and in this case, the circle $\Lambda(\Gamma)$ in $\partial\mathbb{H}^3$ is very complicated; in particular its Housdorff dimension is strictly bigger than 1. The convex core C_M of the hyperbolic manifold $M = \mathbb{H}^3/\Gamma$ is identified with $\Sigma \times [-1, 1]$ in $\Sigma \times (-\infty, \infty) \cong M$. Each boundary component $\Sigma \times \{\pm 1\}$ of the convex core has the structure of "hyperbolic surface bent along a geodesic lamination". (see [300, Section 8.5], [55]). The domain of discontinuity $\Omega(\Gamma)$ consists of two components $\Omega_+(\Gamma)$ and $\Omega_-(\Gamma)$, and the Riemann surfaces $S_\pm = \Omega_\pm(\Gamma)$ correspond to the boundary components $\Sigma \times \{\pm\}$ of $\Sigma \times [\infty, +\infty]$.

Example 3.10.5 (Fiber Group) Let $\hat{M} = \mathbb{H}^3/\hat{\Gamma}$ be a complete hyperbolic manifold of finite volume, and assume that $\hat{M}$ has the structure of a Σ-bundle over S^1. Then the *fiber group*, Γ, the subgroup of $\hat{\Gamma}$ obtained as the image of the fundamental group of a fiber surface Σ, is an infinite normal subgroup. This implies that $\Lambda(\Gamma) = \Lambda(\hat{\Gamma}) = \partial\mathbb{H}^3$ (see [300, Corollary 8.1.3]). Thus the inverse image of a fiber Σ in the universal cover $\mathbb{H}^3$ of $\hat{M}$ is a topological plane whose closure contains the whole ideal boundary $\partial\mathbb{H}^3$. It is very difficult to imagine such a plane, and in this sense, the fiber group Γ is quite different from a quasifucshian group, though they are all isomorphic to $\pi_1(\Sigma)$.

3.10.2 Deformation Space

We continue to denote by Σ a closed orientable surface of genus g with b punctures, which has a negative Euler characteristic. By a *marked hyperbolic structure* on Σ, we mean a pair (S, f) of a finite area complete hyperbolic surface $S = \mathbb{H}^2/\Gamma$ and an orientation-preserving homeomorphism $f : \Sigma \to S$. Note that the composition of $f_* : \pi_1(\Sigma) \to \pi_1(S)$ and the holonomy representation $\pi_1(S) \to \Gamma < \mathrm{Isom}^+ \mathbb{H}^2$ determine a type-preserving discrete faithful representation $\rho : \pi_1(\Sigma) \to \mathrm{Isom}^+ \mathbb{H}^2$. Two marked hyperbolic structure (S_1, f_1) and (S_2, f_2) on Σ are *equivalent* if there is an orientation-preserving isometry $h : S_1 \to S_2$ such that $h \circ f_1$ is homotopic to f_2. This is equivalent to the condition that the corresponding representations ρ_1 and ρ_2 are equal up to conjugation by an element of $\mathrm{Isom}^+ \mathbb{H}^2$. Let $H(\Sigma)$ be the set of all marked hyperbolic structure on Σ up to equivalence.

In order to introduce a natural topology on $H(\Sigma)$, consider the spaces

$$\mathrm{Hom}_{\mathrm{tp}}(\pi_1(\Sigma), \mathrm{Isom}^+ \mathbb{H}^2) := \{\rho : \pi_1(\Sigma) \to \mathrm{Isom}^+ \mathbb{H}^2 \mid \rho \text{ is type-preserving}\};$$
$$\mathcal{R}_{\mathrm{tp}}(\Sigma) := \mathrm{Hom}_{\mathrm{tp}}(\pi_1(\Sigma), \mathrm{Isom}^+ \mathbb{H}^2)/\,\mathrm{Isom}^+ \mathbb{H}^2.$$

By choosing a finite generating set of $\pi_1(\Sigma)$ of cardinality k, $\mathrm{Hom}_{\mathrm{tp}}(\pi_1(\Sigma), \mathrm{Isom}^+ \mathbb{H}^2)$ is identified with a subset of the product (topological) space $(\mathrm{Isom}^+ \mathbb{H}^2)^k$, and the subspace topology it inherits is independent of the choice of a finite generating set. The group $\mathrm{Isom}^+ \mathbb{H}^2$ acts by conjugation on the space $\mathrm{Hom}_{\mathrm{tp}}(\pi_1(\Sigma), \mathrm{Isom}^+ \mathbb{H}^2)$, and $\mathcal{R}_{\mathrm{tp}}(\Sigma)$ is defined to be the quotient space. The set $H(\Sigma)$ is identified with a subset of $\mathcal{R}_{\mathrm{tp}}(\Sigma)$, and we denote the resulting topological space by $AH(\Sigma)$.

The space $AH(\Sigma)$ is nothing other than the *Teichmüller space* $\mathrm{Teich}(\Sigma)$ of Σ. The *Fenchel–Nielsen coordinate* gives a homeomorphism from $AH(\Sigma) = \mathrm{Teich}(\Sigma)$ onto the Euclidean space $\mathbb{R}^{6g-6+3b}$ (see [146], [300, Theroem 5.3,5]). It should be noted that $\mathrm{Teich}(\Sigma)$ can be also identified with the space of *marked Riemann surface structures on* Σ.

Now we consider hyperbolic structures on the oriented 3-manifold $\Sigma \times \mathbb{R}$. By a *marked hyperbolic structure* on $\Sigma \times \mathbb{R}$, we mean a pair (N, f) where $N = \mathbb{H}^3/\Gamma$ is an oriented complete hyperbolic 3-manifold and $f : \Sigma \times \mathbb{R} \to N$ an orientation-preserving homeomorphism which satisfies the following conditions.

- Let $\rho : \pi_1(\Sigma) \to \Gamma < \mathrm{Isom}^+ \mathbb{H}^3$ be the homomorphism obtained as the composition of the homomorphism $(f \circ j)_* : \pi_1(\Sigma) \to \pi_1(N)$, where $j : \Sigma \to \Sigma \times 0 \to \Sigma \times \mathbb{R}$ is the inclusion map, and the holonomy representation $\pi_1(N) \to \mathrm{Isom}^+(\mathbb{H}^3)$ of the hyperbolic manifold N. Then we require that ρ is type-preserving. (In other words, we require that the homeomorphism f maps (ends of Σ) $\times\, \mathbb{R}$ into the *main* cusp of N carrying the parabolic elements ρ(peripheral elements).)

Thus we restrict our attention to the hyperbolic structures on the *pared manifold* $(\Sigma_{g,b} \times I, \partial\Sigma_{g,b} \times I)$ with $I = [-\infty, \infty]$ (see [305, Section 7]) for the teminology).

Two marked hyperbolic structures (N_1, f_1) and (N_2, f_2) on $\Sigma \times \mathbb{R}$ are regarded as *equivalent* if there is an orientation-preserving isometry $h : N_1 \to N_2$ such that $h \circ f_1$ is homotopic to f_2. This condition is equivalent to the condition that the corresponding representations ρ_1 and ρ_2 are equal up to conjugation by an element of $\mathrm{Isom}^+ \mathbb{H}^3$. Thus the set $H(\Sigma \times \mathbb{R})$ of all marked hyperbolic structures on $\Sigma \times \mathbb{R}$

up to equivalence is identified with the subset of the space

$$\mathcal{R}_{\mathrm{tp}}(\Sigma \times \mathbb{R}) := \{\rho : \pi_1(\Sigma) \to \mathrm{Isom}^+ \mathbb{H}^3 \ \rho \text{ is type-preserving}\}/\,\mathrm{Isom}^+ \mathbb{H}^3$$

consisting of (the images of) discrete faithful representations. The set $H(\Sigma \times \mathbb{R})$ with the subspace topology is denoted by $AH(\Sigma \times \mathbb{R})$. This topology is called the *algebraic topology* of $H(\Sigma \times \mathbb{R})$. It is well-known that $\mathcal{R}_{\mathrm{tp}}(\Sigma \times \mathbb{R})$ is Hausdorff, and $AH(\Sigma \times \mathbb{R})$ is a closed subset of $\mathcal{R}_{\mathrm{tp}}(\Sigma \times \mathbb{R})$ (cf. [190, Section 4]).

Let $\mathcal{QF}(\Sigma \times \mathbb{R})$ be the subspace of $AH(\Sigma \times \mathbb{R})$ consisting of the quasifuchsian representations. For each quasifuchsian representation $\rho : \pi_1(\Sigma) \to \mathrm{PSL}(2, \mathbb{C})$ with $\Gamma = \rho(\pi_1(\Sigma))$, the Kleinian manifold $(\mathbb{H}^3 \cup \Omega(\Gamma))/\Gamma \cong \Sigma \times [-\infty, \infty]$ is bounded by two marked Riemann surfaces $S_\pm = \Omega_\pm(\Gamma)/\Gamma$, where $S_\pm$ correspond to $\Sigma \times \{\pm\infty\} \subset \Sigma \times [-\infty, \infty]$. The pair (S_-, S_+) is regarded as a point in the product $\mathrm{Teich}(\bar{\Sigma}) \times \mathrm{Teich}(\Sigma)$, where $\bar{\Sigma}$ is the surface Σ with the reverse orientation. This determines a map

$$\nu : \mathcal{QF}(\Sigma \times \mathbb{R}) \to \mathrm{Teich}(\bar{\Sigma}) \times \mathrm{Teich}(\Sigma).$$

Bers' simultaneous uniformization theorem says that ν is a homeomorphism (see [146]).

The positive solution to Thurston's Density Conjecture by Brock, Canary and Minsky [50], obtained as a consequence of deep results by a number of researchers in the deformation theory of Kleinian groups, says that $AH(\Sigma \times \mathbb{R})$ is equal to the closure of its open subset $\mathcal{QF}(\Sigma \times \mathbb{R})$:

$$AH(\Sigma \times \mathbb{R}) = \overline{\mathcal{QF}(\Sigma \times \mathbb{R})}$$

Thus any discrete faithful type-preserving $\mathrm{PSL}(2, \mathbb{C})$-representation of $\pi_1(\Sigma)$ is a limit of quasifuchsian representations. In particular, a fiber Kleinian group of a hyperbolic surface bundle over S^1 is obtained as the limit of quasifuchsian groups. Historically, the existence of the fiber Kleinian group (and so the existence of a complete hyperbolic structure on surface bundles) was first proved in the case where Σ is a once-punctured torus by Jørgensen [152]: the simplest case of the figure-eight knot complement was also proved by Riley [271]. Thurston was impressed by these works. He proved the hyperbolization theorem for surface bundles in [308] (cf. Otal [247]) via his double limit theorem [308, Theorem 4.1].

Cannon and Thurston [57] found the following surprising fact. Let $\rho_0 : \pi_1(\Sigma) \to \mathrm{PSL}(2, \mathbb{C})$ be a Fuchsian representation, and let $\rho : \pi_1(\Sigma) \to \mathrm{PSL}(2, \mathbb{C})$ be the type-preserving discrete faithful representation whose image Γ gives the fiber group of a hyperbolic Σ-bundle over S^1. Recall that $\Lambda(\Gamma_0) = \partial\mathbb{H}^2$ and $\Lambda(\Gamma) = \partial\mathbb{H}^3$ (see Example 3.10.5), and $\pi_1(\Sigma)$ acts on these sets via ρ_0 and ρ, respectively.

Theorem 3.10.6 (Cannon–Thurston Map) *There is a (ρ_0, ρ)-equivariant surjective continuous map*

$$\kappa : \partial\mathbb{H}^2 = \Lambda(\Gamma_0) \to \Lambda(\Gamma) = \partial\mathbb{H}^3.$$

The map κ is called the *Cannon–Thurston map*. This theorem was first proved by Cannon and Thurston [57] for the closed surface case, and then proved by Bowditch [44] for the general case. Work of many authors has extended the results in various ways (see the review [213]). For the simplest case where Σ is the once-punctured torus, the computer program *OPTi* developed by Wada [314] visualizes deformations of the limit sets of quasifuchsian punctured torus groups (see [15] for background). We can also see a lot of breathtaking pictures related to the Cannon–Thurston maps (mainly for the once-punctured torus) in the book Indra's Pearls [227].

3.10.3 Nielsen–Thurston Classification of Surface Homeomorphisms and Geometrization of Surface Bundles

We quickly recall the Nielsen–Thurston classification of surface homeomorphisms (see [84, 85, 306]). Let $\mathrm{MCG}(\Sigma)$ be the *mapping class group* of Σ (the closed orientable surface of genus g with b punctures such that $\chi(\Sigma) < 0$), the group of the orientation-preserving homeomorphisms of Σ modulo isotopy. We do not distinguish between a homeomorphism of Σ and the element (mapping class) of $\mathrm{MCG}(\Sigma)$ represented by it, as long as there is no fear of confusion. Then Nielsen–Thurston theory says that for any $\varphi \in \mathrm{MCG}(\Sigma)$, one of the following holds.

(1) φ is *periodic*, namely φ has finite order in $\mathrm{MCG}(\Sigma)$. In this case, φ is represented by a (periodic) isometry with respect to some finite-volume complete hyperbolic structure on Σ.
(2) φ is *reducible*, i.e., there is a nonempty family of mutually disjoint essential simple loops whose union is preserved by (a representative of) φ.
(3) φ is *pseudo-Anosov*. This means that Σ has a "half-translation structure" such that the homeomorphism φ is "realized by" a diagonal matrix $\begin{pmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{pmatrix}$ with $\alpha > 1$.

The precise meaning of the last condition is as follows. A *half-translation structure* on Σ is a singular Euclidean metric on Σ, with a finite number of conical singularities of cone angle $k\pi$ ($k \geq 3$), and total cone angle $k'\pi$ ($k' \geq 1$) around each puncture. The surface Σ with cone points removed admits an isometric atlas over $\mathbb{R}^2$ whose transition maps are of the form $(x, y) \mapsto \pm(x, y) + (a, b)$ for some $(a, b) \in \mathbb{R}^2$. Then φ is pseudo-Anosov if there is a half-translation structure on Σ, such that the homeomorphism φ has a local expression $(x, y) \mapsto (\alpha x, \alpha^{-1} y)$ with respect to isometric atlas of the half-translation structure. The constant α is called the *expansion factor* of the map φ.

This condition is described as follows in Thurston's original paper [306, Theorem 4]: there is a real number $\alpha > 1$ and a pair of transverse measured foliations $\mathcal{F}^s$

and $\mathcal{F}^u$ such that $\varphi(\mathcal{F}^s) = \alpha^{-1}\mathcal{F}^s$ and $\varphi(\mathcal{F}^u) = \alpha\mathcal{F}^u$. Here a *measured foliation* on Σ is a singular foliation endowed with a measure in the transverse direction, where only finitely many singularities of "k-pronged saddle" ($k = 1$ or $k \geq 3$) are allowed. The notation $\mathcal{F}_1 = \alpha\mathcal{F}_2$ means that $\mathcal{F}_1$ and $\mathcal{F}_2$ agree as foliations, but the transverse measure of $\mathcal{F}_1$ is α times that of $\mathcal{F}_2$. With respect to the half-translation structure of Σ discussed in the above, the measured foliations $\mathcal{F}^s$ and $\mathcal{F}^u$ are the vertical and holizontal foliations, λ^+ and λ^-, equipped with the transverse measures $|dx|$ and $|dy|$ respectively. (Note that every straight line segment in Σ belongs to a unique (singular) foliation by parallel straight lines, and so the vertical and horizontal foliations make sense.) Since φ is locally expressed by $(x, y) \mapsto (\alpha x, \alpha^{-1} y)$, it preserves the vertical and horizontal measured foliations up to the factors α^{-1} and α, respectively.

By considering the "projective classes" of measured foliations, Thurston constructed the *projective measured foliation space* $\mathrm{PMF}(\Sigma)$ and proved that it forms the boundary of a natural compactification of the Teichmüller space $\mathrm{Teich}(\Sigma)$.

$$\overline{\mathrm{Teich}}(\Sigma) = \mathrm{Teich}(\Sigma) \sqcup \mathrm{PMF}(\Sigma) \cong \mathrm{int}\, B^{6g-6+2b} \sqcup \partial B^{6g-6+2b} \cong B^{6g-6+2b}$$

The compactification is natural in the following sense. The action of $\mathrm{MCG}(\Sigma)$ on $\mathrm{Teich}(\Sigma)$ defined by the rule

$$\varphi(S, f) := (S, f \circ \varphi^{-1}) \quad \text{for } (S, f) \in \mathrm{Teich}(\Sigma)$$

extends to the action on the compactification, so that its restriction to the boundary $\mathrm{PML}(\Sigma)$ is the natural action given by

$$\int_\gamma \varphi_*(\mathcal{F}) = \int_{\varphi^{-1}(\gamma)} \mathcal{F}.$$

Here γ is an arc transverse to the foliation $\varphi(\mathcal{F})$, and $\int_\gamma \varphi_*(\mathcal{F})$ is the measure of γ with respect to the measured foliation $\varphi_*(\mathcal{F})$. It should be noted that the set of all essential simple loops in Σ up to isotopy is identified with a dense subset of $\mathrm{PMF}(\Sigma)$ and that the above action is an extension of the natural action of $\mathrm{MCG}(\Sigma)$ on $\mathcal{S}$.

By using this natural compactification of Teichmüller space, Thurston established the classification of surface homeomorphisms, as follows. For a given $\varphi \in \mathrm{MCG}(\Sigma)$, its action on $\overline{\mathrm{Teich}}(\Sigma) \cong B^{6g-6+2b}$ has a fixed point, by Brower's fixed point theorem. If there is a fixed point in $\mathrm{Teich}(\Sigma)$, then φ is periodic. Suppose there is no fixed points in $\mathrm{Teich}(\Sigma)$ and so all fixed points lie in $\mathrm{PMF}(\Sigma)$. If the underlying foliation of some fixed point contains a closed leaf, then φ is reducible. Thurston managed to prove that φ is pseudo-Anosov in the remaining case.

Now, let $M_\varphi := \Sigma \times \mathbb{R}/(x, t) \sim (\varphi(x), t + 1)$ be the Σ-bundle over S^1 with monodromy φ. Then it is easy to observe that if φ is periodic then M_φ is a Seifert fibered space, and that if φ is reducible then M_φ admits a nontrivial torus decomposition. For the remaining case when φ is pseudo-Anosov, the following theorem was proved by Thurston, as a special case of the geometrization Theorem 3.3.4.

Theorem 3.10.7 *The surface bundle M_φ is hyperbolic if and only if φ is pseudo-Anosov.*

As noted in Sect. 3.10.2, the corresponding fiber group $\rho \in AH(\Sigma \times I)$ is a limit of quasi-fuchsian groups. Actually, for any $(S_-, S_+) \in \mathrm{Teich}(\bar{\Sigma}) \times \mathrm{Teich}(\Sigma)$, ρ is obtained as a limit of a subsequence of the sequence of quasifuchsian groups $\{\nu^{-1}(\varphi^{-k}(S_-), \varphi^k(S_+))\}_{k\geq 0}$ (see McMullen [199, Theorem 3.8]).

3.10.4 Cannon–Thurston Maps and Veering Triangulations

We now describe the combinatorial structure of the Cannon–Thurston map associated with the Σ-bundle M_φ with pseudo-Anosov monodromy φ. Let $\rho_0 : \pi_1(\Sigma) \to \mathrm{PSL}(2,\mathbb{C})$ be a Fuchsian representation with image Γ_0, and let $\rho : \pi_1(\Sigma) \to \mathrm{PSL}(2,\mathbb{C})$ be the discrete faithful representation whose image Γ gives the fiber group of M_φ.

Let j be the inclusion map from $\Sigma = \mathbb{H}^2/\Gamma_0$ to the infinite cyclic cover $\tilde{M}_\varphi = \mathbb{H}^3/\Gamma$ of M_φ, and consider its lift $\tilde{j} : \mathbb{H}^2 \to \mathbb{H}^3$ to the universal cover. Then the Cannon Thurston map $\kappa : \partial\mathbb{H}^2 \to \partial\mathbb{H}^3$ is the boundary map of the extension of $\tilde{j}$ to a map from $\mathbb{H}^2 \cup \partial\mathbb{H}^2$ to $\mathbb{H}^3 \cup \partial\mathbb{H}^3$.

In order to describe the combinatorics of the Cannon–Thurston map κ, let $\tilde{\lambda}^\pm$ be the singular foliations of $\mathbb{H}^2$ obtained as the lifts of the vertical/horizontal foliations $\lambda^\pm$, invariant by φ. Then the endpoints of each leaf of $\tilde{\lambda}^\pm$ are mapped by κ into the same point, and this turns out to generate the combinatorics of κ. To be precise, let $\sim^\pm$ be the equivalence relation on $\partial\mathbb{H}^2$ which identifies the endpoints of each leaf of $\tilde{\lambda}^\pm$ by allowing for leaves that pass through singularities. Let $\sim$ be the equivalence relation on $\partial\mathbb{H}^2$ generated by $\sim^+$ and $\sim^-$. Here the relations $\sim^+$ and $\sim^-$ are "almost independent" in the sense that if $x \sim y$ then $x \sim^+ y$ or $x \sim^- y$ or else there is a parabolic fixed point p of Γ_0 such that either $(x \sim^+ p \sim^- y)$ or $(x \sim^- p \sim^+ y)$. Moreover distinct parabolic fixed points of Γ_0 cannot be equivalent under $\sim$. It was proved by Bowditch [44, Section 9] (cf. [57, Section 5]) that

$$\kappa(x) = \kappa(y) \text{ if and only if } x \sim y.$$

In the remainder of this subsection, we assume that the singularities of the invariant foliations $\lambda^\pm$ occur only at punctures of the fiber. (This condition is satisfied if Σ is a once-punctured torus.) Then it follows that for a point $q \in \partial\mathbb{H}^3$, the inverse image $\kappa^{-1}(q)$ consists of 1, 2 or countably infinitely many points. The last case happens if and only if q is a parabolic fixed point of Γ, and in this case $\partial\mathbb{H}^2 - \kappa^{-1}(q)$ consists of countably infinitely many open intervals. Cannon and Dicks [56] studied the way these intervals are mapped onto the complex plane

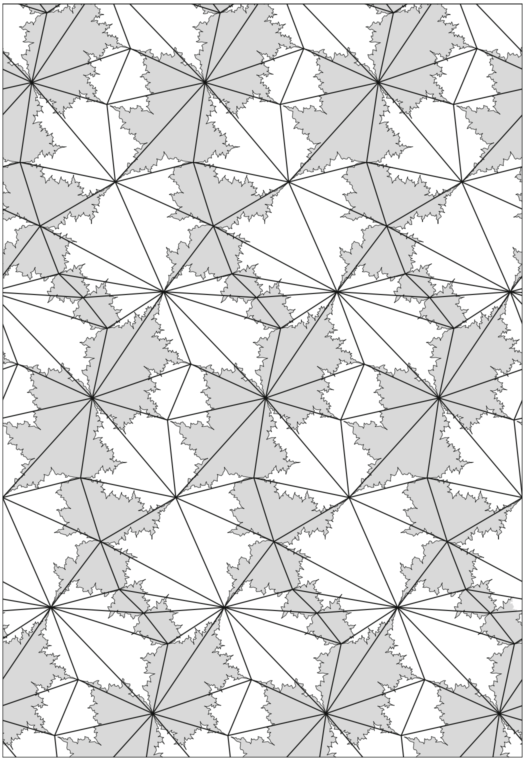


Fig. 3.11 Projected -horosphere triangulation induced by the canonical decomposition and fractal tessellation for a once -punctured torus bundle. Straight line segments etch the projected-horosphere triangulation while fractal arcs etch the fractal tessellation. This picture is taken from [78, Fig. 1]

$\mathbb{C} \cong \partial\mathbb{H}^3 - \{q\}$, and constructed a certain fractal tessellation of $\mathbb{C}$ in the case where Σ is a once-punctured torus. Dicks and Sakuma [78] then observed that there is an intimate relation between the fractal tessellation and the cusp triangulation (lifted to the universal cover $\mathbb{C}$) induced by the canonical triangulation of the hyperbolic once-punctured torus bundles (see Fig. 3.11).

On the other hand, Agol [8] introduced *veering triangulations*, which are (topological) ideal triangulations of cusped hyperbolic 3-manifolds with a very special combinatorial structure. He proved that every hyperbolic surface bundle, for which the singularities of the invariant foliations $\lambda^{\pm}$ occur only at punctures of the fiber, admits a veering triangulation, which is canonical in the sense that it is uniquely determined by the fiber structure. (More strongly, it is determined by Thurston's fiber face to which the fibration belongs [212].)

In the beautiful paper [122], Guéritaud revealed an intimate relation between the veering triangulation and the fractal tessellation arising from the Cannon–Thurston map for every such hyperbolic surface bundle M_φ. To this end, he gave a natural construction of the veering triangulation in terms of the invariant foliations. The construction works in the universal cover $\tilde{\Sigma}$, endowed with the half-translation structure associated with the pseudo-Anosov monodromy. He considered maximal rectangles in $\tilde{\Sigma}$ whose sides are vertical and horizontal in $\tilde{\Sigma}$ and whose interiors are disjoint from the singularities. Such maximal rectangles have one singularities on each side; connecting these 4 singularities produces the ideal tetrahedra of the veering triangulation. This construction enabled Guéritaud to describe the relation between the veering triangulation and the fractal tessellation associated with the Cannon–Thurston map.

Roughly speaking, Guéritaud's construction of the veering triangulation is an analogue of the Delaunay triangulation relative to the singular set, with respect to the ℓ^{∞}-metric arising from the half-translation structure. On the other hand, the canonical decomposition of a cusped hyperbolic manifold is an analogue of the Delaunay triangulation relative to cusps, with respect to the hyperbolic metric. For hyperbolic once-punctured torus bundles, these two decompositions are equal. However, these two decompositions are quite different in general. In fact, it was shown by Hodgson, Issa, Ahmad and Segerman [142] that there exist veering triangulations which are not geometric, in the sense that they are not isotopic to hyperbolic ideal triangulations. Moreover, it was recently proved by Futer, Taylor and Worden [97] that generically veering triangulations are not geometric. In spite of this defect from the viewpoint of hyperbolic geometry, nice applications of veering triangulations to the study of curve complexes were given by Minsky and Taylor [212].

3.11 Representations of 3-Manifold Groups

In Sects. 3.7 and 3.10, we treated deformations of hyperbolic structures. In Sect. 3.7, we considered complete hyperbolic manifolds of finite volume and studied deformations into incomplete hyperbolic structures, whereas in Sect. 3.10, we considered complete hyperbolic manifolds of infinite volume and studied deformations keeping the completeness. In both sections, deformations are described in terms of deformations of holonomy representations.

One purpose of this section is to present the definition of $\mathrm{SL}(2,\mathbb{C})$*character varieties*, which forms a common base ground for both treatments in Sects. 3.7 and 3.10, and then to give a description of the hyperbolic Dehn filling theorem independent of ideal triangulations, following Boileau–Heusener–Porti [32, Appendix B]. For another treatment, see Hodgson–Kerckhoff [135, p.49, Remark].

Another purpose of this section is to describe applications of the character varieties to knot theory and 3-manifold theory. We have already observed in Sect. 3.2.7 that study of representations of knot groups to finite groups gives us a powerful tool in knot theory. The character variety, which is essentially the space of representations of a knot group or a 3-manifold group into the Lie group $\mathrm{SL}(2,\mathbb{C})$ up to conjugation by elements of $\mathrm{SL}(2,\mathbb{C})$, leads to new versatile tools in knot theory and 3-manifold theory. We give a quick review to the Culler–Shalen theory [73, 74, 76] and the A-polynomials due to Cooper, Culler, Gillet, Long and Shalen [64]. For further information, see the survey Shalen [290].

3.11.1 Character Variety

Let M be a compact connected manifold, and let $R(M) = \mathrm{Hom}(\pi_1(M), \mathrm{SL}(2, \mathbb{C}))$ be the space of all representations of $\pi_1(M)$ into $\mathrm{SL}(2, \mathbb{C})$. This set has the structure of a complex affine algebraic set, because it is identified with a subspace of $(\mathrm{SL}(2, \mathbb{C}))^k \subset \mathbb{C}^{4k}$, where k is the cardinality of a generating set of $\pi_1(M)$, defined by a system of polynomial equations. For a representation $\rho \in R(M)$, the function $\chi_\rho : \pi_1(M) \to \mathbb{C}$ defined by $\chi_\rho(\gamma) = \mathrm{tr}(\rho(\gamma))$ is called the *character* of ρ. (We don't distinguish between a representation and the element of $R(M)$.) The set $X(M)$ of all characters also has the structure of an affine algebraic set, and it is called the *character variety* of M. This can be seen as follows. For each $\gamma \in \pi_1(M)$, consider the function $I_\gamma : X(M) \to \mathbb{C}$, defined by $I_\gamma(\rho) = \chi(\gamma)$. Then there are finitely many elements $\gamma_1, \dots, \gamma_d$ for which $I_{\gamma_1} \times \cdots \times I_{\gamma_d} : X(M) \to \mathbb{C}^d$ is an embedding, and its image forms an affine algebraic set [73, Corollary 1.4.5].

The natural projection from the space $R(M)/\mathrm{SL}(2, \mathbb{C})$ of all conjugacy classes of representations onto $X(M)$ fails to be injective only at the conjugacy classes of reducible representations (see [290, Proposition 1.1.1]). In this sense, $X(M)$ is regarded as the quotient $R(M)//\mathrm{SL}(2, \mathbb{C})$ in the category of affine algebraic sets.

If M is a hyperbolic 3-manifold, i.e., if $\mathrm{int}\, M$ admits a complete hyperbolic structure, then the holonomy representation $\rho : \pi_1(M) \to \mathrm{PSL}(2, \mathbb{C})$ lifts to an $\mathrm{SL}(2, \mathbb{C})$-representation (see [72]). In particular, the space $\mathcal{R}_{\mathrm{tp}}(\Sigma \times \mathbb{R})$ of conjugacy classes of type-preserving $\mathrm{PSL}(2, \mathbb{C})$-representations of $\pi_1(\Sigma \times \mathbb{R})$ (see Sect. 3.10.2) is covered by a subspace of $X(\Sigma \times \mathbb{R})$.

3.11.2 Hyperbolic Dehn Filling Theorem and Character Variety

Consider the setting in Sect. 3.7.1, namely M is a connected compact orientable 3-manifold with $\partial M = \sqcup_{j=1}^m T_j$ a non-empty union of tori, such that $\mathrm{int}\, M$ admits a complete hyperbolic structure. Let $\{\mu_j, \lambda_j\}$ is a pair of oriented slopes in the boundary torus T_j, which forms a generator system of $H_1(T_j; \mathbb{Z})$. Let ρ_0 be a lift of the holonomy representation of the complete hyperbolic structure of $\mathrm{int}\, M$, and let χ_0 be its character. Consider the map $I_\mu = (I_{\mu_1}, \dots, I_{\mu_m}) : X(M) \to \mathbb{C}^m$. Then the following theorem holds (see [32, Theorem B.1.2]).

Theorem 3.11.1 *The map $I_\mu : X(M) \to \mathbb{C}^m$ is locally bianalytic at χ_0.*

Using this theorem, we can associate each character in some neighborhood of χ_0 with generalized Dehn filling coefficients. To describe this, recall that the complex translation length $\mathcal{L}_A$ of an element $A \in \mathrm{SL}(2, \mathbb{C})$ is defined as an element of $\mathbb{C}/2\pi\sqrt{-1}\mathbb{Z}$ up to multiplication by ± 1, by the following formula (see Sect. 3.5.1).

$$\mathrm{tr}\, A = 2 \cosh \frac{\mathcal{L}_A}{2}$$

In order to have a well-defined complex translation length as an element in $\mathbb{C}$, we consider the $(\mathbb{Z}/2\mathbb{Z})^m$-branched covering map $\psi : \tilde{U} \to W$ from a neighborhood $\tilde{U} \subset \mathbb{C}^m$ of the origin onto a neighborhood $W \subset X(M)$ of χ_0 such that

$$I_{\mu_j}\psi(\boldsymbol{u}) = \epsilon_j \cosh \frac{u_j}{2} \quad \text{for every } \boldsymbol{u} = (u_1, \dots, u_m) \in \tilde{U}$$

where $\epsilon_j \in \{\pm\}$ is chosen so that $I_{\mu_j}(\psi(\mathbf{0})) = \chi_0(\mu_j) = \mathrm{tr}(\rho_0(\mu_j)) = \epsilon_j 2$. Note that Theorem 3.11.1 guarantees the existence of this covering.

One can define a generalized Dehn surgery coefficients map as in Sect. 3.7.1, as Theorem 3.7.4 holds in this setting, where a certain open neighborhood, $U \subset \tilde{U}$, of the origin plays the role of the open neighborhood U in the theorem (see [32, Proposition B.1.9]). (In fact, by Remark 3.7.5, the space U is bi-holomorphic to the space U in Theorem 3.7.4, when M admits an ideal triangulation.)

In the setting of Sect. 3.7.2, each parameter $\boldsymbol{u} \in U$ corresponds to a parameter z representing the shapes of ideal tetrahedra, and so it determines an (incomplete) hyperbolic structure on $\mathrm{int}\, M$. In the current setting, we appeal to the fact that a small deformation of a hyperbolic structure is parametrized by deformation of the holonomy representation (see [301, Proposition 5.1] and [32, Proposition B.1.10]). This is an outline of the proof the hyperbolic Dehn filling Theorem 3.7.1 without using an ideal triangulation of M, given by [32, Appendix B].

Note that the above proof is not effective in the sense that it gives no information about the size or shape of hyperbolic Dehn surgery space $V \subset (\mathbb{R}^2 \cup \{\infty\})^m$. In [137] (cf. [135, 136]), Hodgson and Kerckhoff developed a new theory of infinitesimal harmonic deformations for compact hyperbolic 3-manifolds with "tubular neighborhood", and gave an effective proof of the hyperbolic Dehn filling theorem; they proved that all generalized Dehn filling coefficients outside a disc of "uniform" size yield hyperbolic structures.

3.11.3 The Culler–Shalen Theory and the Cyclic Surgery Theorem

We give a quick survey of the Culler–Shalen theory. See [290] for a detailed self-contained review. Let M be a compact, connected, orientable, irreducible 3-manifold. Suppose M contains an essential surface F. Then by considering the inverse image $\tilde{F}$ of F in the universal covering $\tilde{M}$, we can construct a tree T, such that the vertices correspond to the components of $\tilde{M} - \tilde{F}$ and the edges correspond to the connected components of $\tilde{F}$, where the edge corresponding to a component of $\tilde{F}$ joins the two vertices corresponding to the components of $\tilde{M} - \tilde{F}$ abutting the component of $\tilde{F}$. The covering transformation group $\pi_1(M)$ acts on the tree T simplicially, and this action is *nontrivial* (i.e., no vertex is stabilized by the whole group $\pi_1(M)$) and *without inversion* (i.e., if an element $\gamma \in \pi_1(M)$ leaves an edge invariant, then γ fixes the edge pointwise). Conversely, it is known that if $\pi_1(M)$

acts simplicially on a tree nontrivially and without inversion, then M contains an essential surface.

In [73], Culler and Shalen established a method for constructing such actions of $\pi_1(M)$ on trees, by using the character variety $X(M)$. The theory says that if $X(M)$ contains an algebraic curve $\mathcal{C}$, then each *ideal point* of the curve gives rise to such an action of $\pi_1(M)$ and hence an essential surface in M. In this theory, Tits–Bass–Serre theory [289] on the structure of subgroups of $\mathrm{SL}(2, F)$, where F is a field with a discrete valuation, plays a key role. Various applications of this theory are given, including (a) a simpler proof and generalization of the Smith conjecture [73] and (b) a proof of the Neuwirth conjecture which says that every nontrivial knot group is a free product of two proper subgroups amalgamated along a free product [74].

In [76], Culler, Gordon, Luecke and Shalen introduced a norm $||\cdot|| : H_1(\partial M; \mathbb{R}) \to \mathbb{R}$ for a compact orientable hyperbolic 3-manifold M with a single torus boundary. A key fact behind this definition is the following: Let X_0 be the irreducible component of the character variety $X(M)$ containing the character χ_0 of the (lifted) holonomy representation of the complete hyperbolic structure on $\mathrm{int}\, M$. Then X_0 has complex dimension 1 (see Theorem 3.11.1). Let $\hat{X}_0$ be the projective completion of the affine algebraic curve X_0 in which the ideal points are smooth. Then for each $\gamma \in H_1(\partial M) = \pi_1(\partial M)$, the restriction to X_0 of the function $I_\gamma : X(M) \to \mathbb{C}$, defined by $I_\gamma(\chi) = \chi(\gamma)$, extends to a rational function, $\hat{I}_\gamma : \hat{X}_0 \to \mathbb{C} \cup \{\infty\}$, where the ideal points of $\hat{X}_0$ (i.e., the points in $\hat{X}_0 - X_0$) are the poles of this rational function. The norm $||\cdot|| : H_1(\partial M; \mathbb{R}) \to \mathbb{R}$ is defined to be the norm obtained as the continuous extension of the function $H_1(\partial M; \mathbb{Z}) \to \mathbb{Z}$ which associates γ with the degree of $\hat{I}_\gamma$. The norm plays a crucial role in the proof of the cyclic surgery theorem below, established by Culler, Gordon, Luecke, and Shalen [76]. The theorem was proved by combining (i) arguments using the norm and (ii) graph-theoretic analysis of the intersection of two incompressible, planar surfaces in knot exteriors.

Theorem 3.11.2 (Cyclic Surgery Theorem) *Let M be a compact, connected, orientable, irreducible* 3*-manifold such that ∂M is a single torus, and suppose that M is not a Seifert fibered space. Let α and β be two non-isotopic essential simple loops on ∂M, such that $\pi_1(M(\alpha))$ and $\pi_1(M(\beta))$ are cyclic. Then the geometric intersection number of α and β is equal to* 1.

In [46, 47], Boyer and Zhang generalized the above idea and proved an analogue of the above theorem for finite surgeries. See [45], for further information.

The Culler–Shalen theory was extended by Morgan and Shalen [220–222] to the theory of $\mathbb{R}$-trees. Here an $\mathbb{R}$-tree is a metric space in which any two points are joined by a unique topological arc. The theory plays a key role in Otal's proof [247] of the double limit theorem. See the reviews [24, 217] for further information.

3.11.4 A-Polynomials

We give a short review of the A-polynomial of a knot K, which is introduced by Cooper, Culler, Gillet, Long, and Shalen [64] by using the character variety $X(M)$ of the knot exterior M of K. The idea is to consider the restriction map $r : X(M) \to X(\partial M)$ induced by the inclusion of $\pi_1(\partial M)$ into $\pi_1(M)$. Then even though $X(M)$ is complicated, its image $r(X(M))$ can be very simple. Note that $\pi_1(\partial M)$ is the free abelian group freely generated by the longitude λ and the meridian μ. Thus, for any irreducible 1-dimensional component $\mathcal{C}$ in the image $r(X(M)) \subset X(\partial M)$, there is a holomorphic map $f : \mathcal{C} \to \mathbb{C} \times \mathbb{C}$ which assigns the pair of the "eigen values" of the images of λ and μ by the corresponding representations. Then the closure of the image $f(\mathcal{C})$ becomes an algebraic curve in $\mathbb{C}^2$. Such a curve is equal to the zero set of a single defining polynomial, $F_{\mathcal{C}}(x, y)$. Now consider the product $\prod_{\mathcal{C}} F_{\mathcal{C}}(x, y)$ of the defining polynomials $F_{\mathcal{C}}(x, y)$ where $\mathcal{C}$ runs over the 1-dimensional irreducible components of $r(X(M))$. Then the *A-polynomial* of K is defined as

$$A_K(x, y) = \frac{1}{x-1} \prod_{\mathcal{C}} F_{\mathcal{C}}(x, y)$$

The reason of dividing out by the factor $x - 1$ is that $H_1(M)$ is the free abelian group generated by μ and so we always have a component corresponding to abelian representations, which gives rise to the factor $x - 1$. By normalizing $A_K(x, y)$ so that it is an integral polynomial, it is defined up to multiplication by $\pm x^a y^b$.

It is obvious that $A_O(x, y) = 1$ for the trivial knot O, and it is proved that the converse also holds (see Boyer–Zhang [48] and Dunfield–Garoufalidis [80]). The most important properties of the A-polynomials come from the fact that they encode information about the boundary slopes of the knot, via the Newton polygon of $A_K(x, y)$. Recall that a *boundary slope* of a knot K is a slope (isotopy class of an essential simple loop) in the boundary torus of the knot exterior M, such that there is an essential surface in M whose boundary consists of loops representing the slope. The Newton polygon of the polynomial $A_K(x, y)$ is the convex hull of the finite set:

$$\{(i, j) \in \mathbb{Z}^2 \mid \text{the coefficient of } x^i y^j \text{ in } A_K(x, y) \text{ is non-zero}\}.$$

The following striking theorem is proved by [64, Theorem 3.4].

Theorem 3.11.3 *Slopes of the edges of the Newton polygon of $A_K(x, y)$ are boundary slopes of the knot K.*

3.12 Knot Genus and Thurston Norm

By generalizing the genus of a knot, Thurston [304] defined a (semi-)norm on $H^1(M;\mathbb{R}) \cong H_2(M,\partial M;\mathbb{R})$ for a compact orientable 3-manifold M. It is called the *Thurston (semi-)norm* of M. By the work of Gabai [98], the Thurston norm is identical to the Gromov norm on $H_2(M,\partial M;\mathbb{Z})$. The Thurston norm can be used to study the set of fiberings of M over the circle, and the work of Fried and McMullen enabled a unified treatment of the fiberings of M. After recalling these works, we explain two *Thurstonian connections* between the topology and geometry of 3-manifolds, related to Thurston norms. Namely, we survey (i) the relation of the Thurston norm with the *hyperbolic torsion polynomial* due to Dunfield–Friedl–Jackson [81] and Agol–Dunfield [10], and (ii) that with the *harmonic L^2-norm* with respect to the hyperbolic metric due to Brock–Dunfield [49].

3.12.1 Thurston Norm

Let M be a compact oriented 3-manifolds with ∂M a possibly empty union of tori. For a compact possibly disconnected surface Σ, let Σ_0 be the surface consisting of the components of Σ which are neither homeomorphic to D^2 nor S^2, and define its *complexity* by $\chi_-(\Sigma) := |\chi(\Sigma_0)|$. For an integral homology class $\alpha \in H_2(M,\partial M;\mathbb{Z})$, define its *Thurston norm* $||\alpha||_{\mathrm{Th}}$ by

$$||\alpha||_{\mathrm{Th}} = \min\{\chi_-(\Sigma) \mid [\Sigma] = \alpha\}$$

Theorem 3.12.1

(1) $||\cdot||_{\mathrm{Th}}$ *extends to a continuous map* $||\cdot||_{\mathrm{Th}} : H^1(M;\mathbb{R}) \cong H_2(M,\partial M;\mathbb{R}) \to \mathbb{R}_{\geq 0}$, *and this gives a semi-norm on* $H^1(M;\mathbb{R})$. *Moreover, if any compact orientable surface properly embedded in* M, *representing a nontrivial homology class, has a negative Euler characteristic, then* $||\cdot||_{\mathrm{Th}}$ *is a norm.*

(2) *Suppose* $||\cdot||_{\mathrm{Th}}$ *is a norm, then the unit ball*

$$B_M = \{\alpha \in H^1(M;\mathbb{R}) \mid ||\alpha||_{\mathrm{Th}} \leq 1\}$$

is a finite-sided polyhedron whose vertices are rational points.

(3) *Suppose* $||\cdot||_{\mathrm{Th}}$ *is a norm. Then there are codimension one faces* $F_1, \ldots, F_k$, *of* B_M *satisfying the following conditions.*

   (i) Any integral cohomology class in the interior of the cone $\mathbb{R}_+ \cdot F_i$ *is a fiber class.*

   (ii) Conversely, any fiber class is contained in the interior of some cone $\mathbb{R}_+ \cdot F_i$.

Here a class $\phi \in H^1(M;\mathbb{Z})$ is called a *fibered class* if it is an integral multiple of the cohomology class represented by a bundle projection $p : M \to S^1$. In the above theorem, each F_i is called a *fibered face*.

The fiber structures contained in the interior of the cone on a fibered face can be given a unified treatment, and various interesting results can be obtained. In particular, building on the results of Fried, McMullen [200] proved that each fibered face F determines a 2-dimensional "lamination" $\mathcal{L}$ of M transverse to every fiber surface Σ with (Poincaré dual of $[\Sigma]$) $\in \mathbb{R}_+ \cdot F$, where $\Sigma \cap \mathcal{L}$ is the stable lamination of the monodromy of the fibration. By using this result, he defined the *Teichmüller polynomial* $\theta_F \in \mathbb{Z}[H_1(M;\mathbb{Z})/\operatorname{Tor} H_1(M;\mathbb{Z})]$ and proved the following results [200].

- The Teichmüller polynomial is symmetric, i.e., if $\theta_F = \sum_g a_g g$ then $\theta_F = \sum_g a_g g^{-1}$ up to a unit in $\mathbb{Z}[H_1(M)/\operatorname{Tor} H_1(M)]$.
- For any integral cohomology class $\phi \in \mathbb{R}_+ \cdot F$, the expansion factor $k(\varphi)$ of the corresponding monodromy φ is equal to the largest root of the one-variable polynomial obtained by evaluating θ_F by ϕ.
- The function $\phi \mapsto 1/\log k(\varphi)$ extends to a real-analytic function on $\mathbb{R}_+ \cdot F$ which is strictly concave.
- The cone $\mathbb{R}_+ \cdot F$ is dual to a vertex of the Newton polygon $\subset H_1(M;\mathbb{R})$ of θ_F.
- If the lamination $\mathcal{L}$ is transversely orientable, then the (multivariable) Alexander polynomial of M divides the Teichmüller polynomial θ_F

To end this subsection, we recall an important result of Gabai [98], obtained as a corollary of his construction of codimension 1 transversely oriented foliations without Reeb components which contain a given Thurston norm minimizing surface as a closed leaf. To explain this, we consider another (semi-)norm $||\cdot||^{\mathrm{s}}_{\mathrm{Th}}$ on $H^1(M;\mathbb{R}) \cong H_2(M,\partial M;\mathbb{R})$ for a compact irreducible orientable 3-manifold M, defined by using immersed surfaces instead of embedded surfaces. Namely, for an integral homology class $\alpha \in H_2(M,\partial M;\mathbb{Z})$, define $||\alpha||^{\mathrm{s}}_{\mathrm{Th}}$ to be the minimum of $\chi_-(\Sigma)$ of a compact oriented surface Σ for which there is a proper immersion $f : (\Sigma,\partial\Sigma) \to (M,\partial M)$ such that $f_*([\Sigma]) = \alpha$, namely,

$$||\alpha||^{\mathrm{s}}_{\mathrm{Th}} = \min\{\chi_-(\Sigma) \mid \exists f : (\Sigma,\partial\Sigma) \looparrowright (M,\partial M) \text{ such that } f_*([\Sigma]) = \alpha\}.$$

The new norm $||\cdot||^{\mathrm{s}}_{\mathrm{Th}}$ is defined as a continuous extension of the above norm on the integral homology.

In addition to this, as in Sect. 3.8.4, the *Gromov norm* $||\cdot||_{\mathrm{Gr}}$ is defined by

$$||\alpha||_{\mathrm{Gr}} := \inf\{||z|| \mid z \text{ is a singular cycle representing the homology class } \alpha\},$$

where, for a (real) singular chain $z = \sum_j a_j\sigma_j$, its norm $||z||$ is defined as the sum $\sum_j |a_j|$ of the absolute values of its coefficients. The following theorem was proved by Gabai [98].

Theorem 3.12.2 *Let M be a connected compact irreducible orientable 3-manifold with possibly empty toral boundary. Then the three norms on $H^1(M;\mathbb{R}) \cong H_2(M,\partial M;\mathbb{R})$ coincide, namely,*

$$||\cdot||_{\mathrm{Th}} = ||\cdot||^{\mathrm{s}}_{\mathrm{Th}} = ||\cdot||_{\mathrm{Gr}}.$$

In particular, for a knot K in S^3, its genus $g(K)$ is equal to the *immersed genus* of K, which is defined as the minimum of the genus $g(\Sigma)$ of a compact connected oriented surface Σ such that there is an immersion $f : \Sigma \to S^3$, with $f^{-1}(K) = \partial\Sigma$, whose singular set is disjoint from K. This is a generalization of Dehn's lemma for higher genus, and gives a partial affirmative answer to a question raised by Papakyriakopoulos [252], who established Dehn's lemma.

3.12.2 Evaluation of Thurston Norms in Terms of Twisted Alexander Polynomials

The *twisted Alexander polynomials*, defined by Lin [183] for classical knots and by Wada [313] in the general setting, give a powerful tool for studying the Thurston norm. Such a "polynomial" $\Delta(M,\phi,\rho)$ depends on a class $\phi \in H^1(M;\mathbb{Z})$ and a linear representation $\rho : \pi_1(M) \to \mathrm{GL}(V)$, where V is a finite-dimensional vector space over a field F. Then $\Delta(M,\phi,\rho)$ is defined as an element of the quotient field $F(t^{\pm 1})$ of the group ring $F[t^{\pm 1}]$, and analogies of Theorem 3.2.7 on the classical Alexander polynomial are obtained by several authors (see the surveys [89, 160]). Friedl and Viddusi [90, 91] proved the surprising results that the twisted Alexander polynomials can detect fiber classes and the Thurston norms.

When K is a hyperbolic knot in S^3, it is natural to consider the twisted Alexander polynomial for the representation $\rho : G(K) \to \mathrm{SL}(2,\mathbb{C})$ which projects to the holonomy representation of the complete hyperbolic structure of $S^3 - K$. Though there are precisely two such representations up to conjugacy, there is unique one for which $\mathrm{tr}\,\rho(\mu) = +2$, where μ a meridian of K. (For the other lift ρ', we have $\mathrm{tr}\,\rho'(\mu) = -2$.) Thus we can consider the twisted Alexander ploynomial $\Delta(E(K),\phi,\rho)$, where $\phi \in H^1(E(K);\mathbb{Z}) \cong \mathbb{Z}$ is the generator. The invariant is called the *hyperbolic torsion polynomial* of K and is denoted by $\mathcal{T}_K(t)$ (see [81]). The artificial choice of the lift ρ is irrelevant, because if ρ is replaced with the other lift ρ', then the corresponding polynomial $\mathcal{T}'_K(t)$ is equal to $\mathcal{T}_K(-t)$. As a special case of the general results on the twisted Alexander polynomial, the following hold for every hyperbolic knot K in S^3.

(1) $4g(K) - 2 \geq \deg \mathcal{T}_K(t)$.
(2) If K is fibered, then $\mathcal{T}_K(t)$ is monic.

These may be regarded as analogies of Theorem 3.2.7(2) and (3) on the classical Alexander polynomial. Dunfield, Friedl and Jackson [81] made extensive computer experiments, and confirmed that for all hyperbolic knots with at most 15 crossings, the estimate (1) is sharp and that (2) detect all non-fibered knots. In particular, the hyperbolic torsion polynomial detects that the genera of the Kinoshita–Terasaka knot and the Conway knot are 3 and 5, respectively. (The genera of arborescent links, including these two knots, had been determined by Gabai [101] through the topological study of complementary sutured manifolds.) Thus the hyperbolic torsion polynomials can distinguish knots which are mutants of each other. In [10], Agol and Dunfield studied the conjecture posed by [81], that the estimate (1) is sharp for every hyperbolic knot, and they verified the conjecture for *libroid* hyperbolic knots in S^3. The libroid knots form a broad class of knots, which is closed under Murasugi sum, and in particular all arborescent are libroid knots.

3.12.3 Harmonic Norm and Thurston Norm

Let M be a closed orientable hyperbolic 3-manifold. Then in addition to the topologically defined Thurston norm $||\cdot||_{\mathrm{Th}}$, there is yet another canonically defined geometric norm on $H^1(M;\mathbb{R})$. By the rigidity theorem, M admits a unique hyperbolic metric, and by applying Hodge theory to this Riemannian metric, we can identify $H^1(M;\mathbb{R})$ with the space of harmonic 1-forms. Thus the *harmonic norm* $||\cdot||_{L^2}$ determines another norm on $H^1(M;\mathbb{R}) \cong H_2(M;\mathbb{R})$. Here the harmonic norm is the one associated with the usual inner product for 1-forms:

$$\langle \alpha, \beta \rangle = \int_M \alpha \wedge *\beta,$$

where $*$ is the Hodge $*$-operator. Since it comes from a positive-definite inner product, the unit ball of $||\cdot||_{L^2}$ is a smooth ellipsoid. Brock and Dunfield [49] proved the following relation between the topological norm and the geometric norm.

Theorem 3.12.3 *For all closed orientable hyperbolic* 3*-manifold M one has*

$$\frac{\pi}{\sqrt{\mathrm{vol}(M)}}||\cdot||_{\mathrm{Th}} \le ||\cdot||_{L^2} \le \frac{10\pi}{\sqrt{\mathrm{inj}(M)}}||\cdot||_{\mathrm{Th}}.$$

In the above theorem, $\mathrm{inj}(M)$ denotes the *injectivity radius* of M, i.e., half of the length of the shortest closed geodesic. Moreover, they also showed that the above estimates are in some sense sharp, by giving families of examples.

These results were obtained as refinements of a result of Bergeron, Sengün and Venkatesh [22], which in turn is preceded by the work by Kronheimer and Mrowka [174] that characterize the Thurston norm as the infimum (over all possible Rimennian metrics) of certain scaled harmonic metrics.

3.13 Finite-Index Subgroups of Knot Groups and 3-Manifold Groups

As explained in Sect. 3.2.7, finite branched/unbranched coverings of knots are a powerful tool for distinguishing knots. This fact reflects the richness of finite-index subgroups of knot groups. In this section, we survey the following topics which illustrate this richness: (i) universal groups, which produce all closed orientable 3-manifolds, (ii) positive solution of the virtual fibering conjecture, (iii) Grothendieck rigidity of 3-manifold groups, and (iv) mysterious relation between the Gromov norm and the homology growth of finite coverings.

3.13.1 Universal Knots/Links and Universal Groups

In an unpublished preprint [303], W. Thurston presented a very complicated six component link in S^3, and proved the surprising fact that every closed orientable 3-manifold can be expressed as a branched cover of the 3-sphere branched over this link. He called links with this property *universal links*. He asked if a universal knot exists, and if even the figure-eight knot was universal. This question was answered affirmatively by Hilden, Lozano and Montesinos in [130], where they proved that every hyperbolic 2-bridge knot and link is universal.

Moreover, it was later proved by Hilden, Lozano, Montesinos and Whitten in [131] that every closed orientable 3-manifold M is a covering of S^3 branched over the Borromean rings and having branching indices 1, 2 and 4 (cf. Fig. 3.12). This implies that the hyperbolic orbifold $\mathcal{U} = \mathbb{H}^3/\mathrm{U}$ with underlying space S^3 and with singular set the Borromean ring where all components have cone angle $\pi/2$, is a *universal orbifold* in the following sense: for any closed orientable 3-manifold M, there is a finite orbifold covering $\mathcal{O} \to \mathcal{U}$ with underlying space $|\mathcal{O}|$ homeomorphic to M. In other words, the orbifold fundamental group $\mathrm{U} = \pi_1^{\mathrm{orb}}(\mathcal{U})$ is a *universal group*, i.e., for any closed orientable 3-manifold M, there is a finite index subgroup Γ of U such that $|\mathbb{H}^3/\Gamma| \cong M$. It is surprising that all closed orientable 3-manifolds are constructed from a single group U and its finite index subgroups. Moreover, universal groups seem to be ubiquitous (cf. [132]).

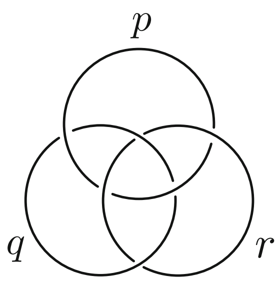


Fig. 3.12 The *Borromean orbifold* $\boldsymbol{B}(p, q, r)$ is a universal orbifold if $p \geq 3$ and both q and r are even integers ≥ 4 by [132]

3.13.2 Virtual Fibering Conjecture

The positive solution of Thurston's virtual fibering conjecture by Agol [9] and the geometric solution of Waldhausen's conjecture for hyperbolic manifolds due to Kahn–Markovic [154], which play a key role in the proof of the virtual fibering conjecture, also reflect the richness of subgroups of Kleinian groups. Please see Bestvina [25] for a survey of this important topic.

Here, I only recall Walsh's simple construction [318] of a nontrivial example of virtual fibering using knot theory. Let K be a spherical Montesinos knot/link which is not fibered, e.g., the 5_2 knot, the 2-bridge knot of slope $2/7$. Then the double branched covering $M_2(K)$ of S^3 branched over K is a spherical manifold and so its universal covering $\tilde{M}_2(K)$ is the 3-sphere. The inverse image, $\tilde{K}$, of K in the universal cover is a great circle link in S^3, because it is the singular set of the isometric group action of the π-orbifold group of K (cf. Sect. 3.4.3). Pick a component O of $\tilde{K}$, and observe that the remaining components form a closed braid around O, because $\tilde{K}$ consists of great circles. This shows that the covering $\tilde{M}_2(K) - \tilde{K}$ of $S^3 - K$ is a punctured disk bundle over the circle, though $S^3 - K$ itself does not admit a fiber structure over the circle.

3.13.3 Profinite Completions of Knot Groups and 3-Manifold Groups

As explained in Sect. 3.2.7, representations of knot groups onto finite groups serve a powerful tool for distinguishing knots. Thus it is natural to ask the following question (cf. [30]).

Question 3.13.1 To what degree does the set of finite quotients of knot groups distinguish knots? More generally, what properties of 3-manifolds are determined by the set of finite quotients of their fundamental groups?

The geometrization theorem and Hempel's argument [128] show that every 3-manifold group is *residually finite*, namely, for any nontrivial element $g \in \pi_1(M)$, where M is a compact connected orientable 3-manifold, there is a finite quotient of $\pi_1(M)$ in which g remain nontrivial. This implies that the above question can be formulated in terms of the profinite completion of the fundamental group.

Recall that the *profinite completion* of a group Γ, is the inverse limit of the inverse system of finite quotients of Γ: we denote it by $\hat{\Gamma}$. (The profinite completion is actually defined to be a topological group endowed with the *profinite topology*. By Nikolov–Segal [244], the topology of any "finitely generated profinite group" is determined by the algebraic structure. So we do not care about the topological structure in this subsection.) The natural map $\Gamma \to \hat{\Gamma}$ is injective if and only if Γ is residually finite. Let $\mathcal{C}(\Gamma)$ denote the family of finite quotients of Γ. Then the following holds (see [268, p.88-89], [185, Theorem 2.2]).

Theorem 3.13.2 *For two finitely generated residually finite groups* Γ_1 *and* Γ_2*, the equality* $\mathcal{C}(\Gamma_1) = \mathcal{C}(\Gamma_2)$ *holds if and only if* $\hat{\Gamma}_1 \cong \hat{\Gamma}_2$*, i.e., the profinite completions are isomorphic.*

Thus Question 3.13.1 is reformulated by using the profinite completion; in particular, the following question arises as a special case.

Question 3.13.3 Let M_1 and M_2 be connected compact orientable 3-manifolds, for which the profinite completions $\widehat{\pi_1(M_1)}$ and $\widehat{\pi_1(M_2)}$ are isomorphic. Are $\pi_1(M_1)$ and $\pi_1(M_2)$ isomorphic?

The answer to the above question is no. In fact, Funar [94] and Hempel [129] showed that the profinite completion of the fundamental group cannot always distinguish certain pairs of torus bundles nor certain pairs of Seifert fibered spaces. It is still an open question though whether the profinite completion can distinguish any two hyperbolic 3-manifolds. Boileau and Friedl [30] considered a more relaxed Question 3.13.1 and obtained various interesting results concerning fiberedness and the Thurston norm, and have shown that the figure-eight knot and torus knots are distinguished from other knots by the profinite completions of their knot groups.

On the other hand, the following problem had been posed by Grothendieck [119].

Problem 3.13.4 (Grothendieck) Let $\varphi : \Gamma_1 \to \Gamma_2$ be a homomorphism of finitely presented residually finite groups for which the extension $\hat{\varphi} : \hat{\Gamma}_1 \to \hat{\Gamma}_2$ is an isomorphism. Is φ an isomorphism?

If $\hat{\varphi} : \hat{\Gamma}_1 \to \hat{\Gamma}_2$ is an isomorphism, then the composition $\Gamma_1 \to \hat{\Gamma}_1 \to \hat{\Gamma}_2$ is an injection and so $\varphi : \Gamma_1 \to \Gamma_2$ must be an injection. Therefore Grothendieck's problem reduces to the case where Γ_1 is a subgroup of Γ_2. Long and Reid [185] introduced the following terminology. For a group G and its subgroup $H < G$, the pair (G, H) is a *Grothendieck pair* if the inclusion $j : H \to G$ provides a negative answer to Grothendieck's problem. If for all finitely generated subgroups $H < G$, (G, H) is never a Grothendieck pair then G is *Grothendieck rigid*.

The following theorem was proved by Long and Reid [185] for the case where M is closed and by Boileau and Friedl [31] for general case.

Theorem 3.13.5 *Let* M *be a connected, orientable, irreducible, compact* 3-*manifold. Then* $\pi_1(M)$ *is Grothendieck rigid.*

In the examples of Funar [94] and Hempel [129], the isomorphisms between the profinite completions are not induced by a homomorphism between the 3-manifold groups.

3.13.4 Homology Growth

Investigation of the first homology groups of finite (branched or unbranched) coverings has a long history (cf. Sect. 3.2.7). For the homology of finite abelian coverings of links, it was proved that they are essentially determined by the Alexander invariants of links (see [86, 198, 280]). In [111], Gordon studied the asymptotic behavior the homology of finite cyclic branched coverings of a knot, and gave a necessary and sufficient condition for $H_1(M_n(K); \mathbb{Z})$ to be periodic with respect to n, in terms of the Alexander invariants. This in particular implies that if the Alexander polynomial $\Delta_K(t)$ has a root which is not a primitive root of 1 then $H_1(M_n(K); \mathbb{Z})$ cannot be periodic. In fact, he showed that the order $|H_1(M_n(K); \mathbb{Z})|$ is unbounded under the same assumption, and then asked if the order $|H_1(M_n(K); \mathbb{Z})|$ tends to ∞. Riley [273] and González-Acuña and Short [109], independently, proved that $|H_1(M_n(K); \mathbb{Z})|$ grows exponentially. To be precise, the following was proved:

$$\lim_{n_j \to \infty} \frac{1}{n_j} \log |H_1(M_{n_j}(K); \mathbb{Z})| = \log \mathbb{M}(\Delta_K),$$

where $\{n_j\}$ runs over the natural numbers such that $|H_1(M_{n_j}(K); \mathbb{Z})|$ is finite, and $\mathbb{M}(\Delta_K)$ is the *Mahler measure* of the Alexander polynomial $\Delta_K(t)$. The Mahler measure of a polynomial $f(t)$ is defined by

$$\begin{aligned} \mathbb{M}(f) &= \exp\left(\int_{S^1} \log |f(s)| ds\right) \\ &= \exp\left(\int_0^1 \log |f(e^{2\pi\sqrt{-1}t})| dt\right) \\ &= |c| \prod_{f(\omega)=0} \max(|\omega|, 1) \quad (c \text{ is the leading coefficient of } f(t)). \end{aligned}$$

Here, the last equality is a consequence of Jensen's formula [11, p. 205].

This result was extended by Silver and Williams [293] to links in S^3, by using the result of Schmidt [282] on the entropy of a certain dynamical system. Let L be an m-component oriented link in S^3, with the complement $X = S^3 - L$. For a subgroup $\Lambda \subset H_1(X; \mathbb{Z}) \cong \mathbb{Z}^m$ of rank m, let X_Λ^{br} be the corresponding branched covering of Z. Set

$$\langle \Lambda \rangle = \min\{|x| \mid x \in \Lambda - \{0\}\},$$

where $|x| = \sqrt{\sum_i |x_i|^2}$ for $x = (x_1, \ldots, x_m) \in \mathbb{Z}^m$. Let $\Delta_L \in \mathbb{Z}[t_1, \ldots, t_m]$ be the (0-th) Alexander polynomial of L.

Theorem 3.13.6 (Silver–Williams) *Under the above setting, suppose that* $\Delta_L \neq 0$. *Then the following holds:*

$$\limsup_{\langle \Lambda \rangle \to \infty} \frac{\ln |\operatorname{Tor}_{\mathbb{Z}} H_1(X_\Lambda^{\mathrm{br}}; \mathbb{Z})|}{|\mathbb{Z}^m / \Lambda|} = \log \mathbb{M}(\Delta_L).$$

Here $\mathbb{M}(\Delta_L)$ is the Mahler measure of Δ_L, defined by

$$\mathbb{M}(\Delta_L) = \exp\left(\int_{T^m} \log|\Delta_L(s)|ds\right),$$

where $T^m := (S^1)^m \subset \mathbb{C}^m$ is the multiplicative subgroup in $\mathbb{C}^m$, and ds indicates integration with respect to normalized Haar measure on T^m.

In [179], Thang Le solved a conjecture of Schmidt [282], and by using the solution, he extended the above result to links in oriented integral homology 3-spheres, and to include the case where the 0-th Alexander polynomial vanishes, by replacing the 0-th Alexander polynomial with the first non-vanishing Alexander polynomial. He also proved that the same formula holds for unbranched abelian coverings.

In [180], Le also studied the asymptotic behavior of the homology of non-abelian coverings, by using the result on L^2-torsion by Lück [186, Theorems 4.3 and 4.9]. Let X be an irreducible compact orientable 3-manifold with infinite fundamental group with (possibly empty) toral boundary. For a subgroup Γ of $\pi_1(X)$ of finite index, let X_Γ be the corresponding finite covering of X. A sequence $\{\Gamma_k\}$ of subgroups of $\pi_1(X)$ of finite index is said to be *nested*, if $\Gamma_{k+1} < \Gamma_k$. It is said to be *exhaustive* if $\cap_k \Gamma_k = \{1\}$.

Theorem 3.13.7 (Le) *Under the above setting, the following holds for any nested exhaustive sequence* $\{\Gamma_k\}$ *of normal subgroups of* $\pi_1(X)$ *of finite index:*

$$\limsup_{k\to\infty} \frac{\ln|\operatorname{Tor} H_1(X_{\Gamma_k};\mathbb{Z})|}{[\pi_1(X):\Gamma_k]} \le \frac{V_{\mathrm{tet}}||X||}{6\pi},$$

where $||X||$ *is the Gromov norm of* X.

For a knot K in S^3 with exterior X and a finite index subroup $\Gamma < G(K)$, let X_Γ^{br} be the corresponding branched covering of S^3 branched over K.

Theorem 3.13.8 (Le) *Under the above setting, the following holds for any nested exhaustive sequence* $\{\Gamma_k\}$ *of normal subgroups of* $G(K) = \pi_1(X)$ *of finite index.*

$$\limsup_{k\to\infty} \frac{\ln|\operatorname{Tor} H_1(X_{\Gamma_k}^{\mathrm{br}};\mathbb{Z})|}{[\pi_1(X):\Gamma_k]} \le \frac{V_{\mathrm{tet}}||X||}{6\pi},$$

where $||X||$ *is the Gromov norm of* X.

For the sake of simplicity, we stated Le's theorem only for regular coverings. However, the actual statement of his theorem is much more general and it does not restrict to regular coverings. For a precise statement, see the original paper [180]. Moreover, he conjectures that the identity holds in both theorems.

The homology of finite (branched/unbranched) coverings is a common and well-known invariant in knot theory. It is impressive that the asymptotic behavior of this familiar invariant reflects the deep geometric structure of the knot.

Acknowledgments The author would like to thank Ken'ichi Ohshika for giving him the challenging opportunity to survey the tremendous impact of Thurston's work on knot theory. He is also grateful to François Guéritaud, Luisa Paoluzzi, and Han Yoshida for correcting errors and providing valuable comments for Sects. 3.10, 3.4 and 3.9, respectively. The author would also like to thank Yuya Koda, Gaven Martin, and Hitoshi Murakami for reading through an early version and for sending him a large number of valuable suggestions and corrections. The author's thanks also go to Hirotaka Akiyoshi, Warren Dicks, Hiroshi Goda, Kazuhiro Ichihara, Yuichi Kabaya, Takuya Katayama, Akio Kawauchi, Eiko Kin, Thang Le, Hidetoshi Masai, José María Montesinos, Kimihiko Motegi, Kunio Murasugi, Shunsuke Sakai, Masakazu Teragaito, Ken'ichi Yoshida, and Bruno Zimmermann for their valuable information and suggestions on early versions of this survey. The author is grateful to the referee for his/her very careful reading and valuable suggestions, including Remark 3.7.5(1). Finally, the author would like to thank Athanase Papadopoulos for his extremely careful check of the final draft.

The author was supported by JSPS Grants-in-Aid 15H03620, 20K03614, and by Osaka City University Advanced Mathematical Institute (MEXT Joint Usage/Research Center on Mathematics and Theoretical Physics JPMXP0619217849).

References

1. C. Adams, The noncompact hyperbolic 3-manifold of minimum volume. Proc. Am. Math. Soc. **100**, 601–606 (1987)
2. C. Adams, *The Knot Book* (W.H. Freeman & Co./American Mathematical Society, San Francisco/Providence, 1994)
3. C. Adams, Knot theory and its applications: expository articles on current research. Chaos Solitons Fractals **9**(4–5), 531–824 (1998)
4. C. Adams, Hyperbolic knots, in *Handbook of Knot Theory* (Elsevier B.V., Amsterdam, 2005), pp. 1–18
5. I. Agol, Bounds on exceptional Dehn filling. Geom. Topol. **4**, 431–449 (2000)
6. I. Agol, Bounds on exceptional Dehn filling II. Geom. Topol. **14**, 1921–1940 (2010)
7. I. Agol, The minimal volume orientable hyperbolic 2-cusped 3-manifolds. Proc. Am. Math. Soc. **138**, 3723–3732 (2010)
8. I. Agol, Ideal triangulations of pseudo-Anosov mapping tori, in *Topology and Geometry in Dimension Three*. Contemporary Mathematics, vol. 560 (American Mathematical Society, Providence, 2011), pp. 1–17
9. I. Agol, The virtual Haken conjecture. With an appendix by Agol, Daniel Groves, and Jason Manning. Doc. Math. **18**, 1045–1087 (2013)
10. I. Agol, N. Dunfield, Certifying the Thurston norm via $SL(2, \mathbb{C})$-twisted homology, in *The Thurston Memorial Conference Proceedings* (Princeton University Press, Princeton, 2015), 23 p. arXiv:1501.02136
11. L.V. Ahlfors, *Complex Analysis*, 2nd edn. (McGraw-Hill, New York, 1966)
12. I.R. Aitchison, J.H. Rubinstein, Combinatorial cubings, cusps, and the Dodecahedral knots, in *Topology '90*, Columbus, OH, 1990, ed. by B. Apanasov, W.D. Neumann, A.W. Reid, L. Siebenmann. Ohio State University Mathematical Research Institute Publications, vol. 1 (de Gruyter, Berlin, 1992), pp. 17–26
13. H. Akiyoshi, Finiteness of polyhedral decompositions of cusped hyperbolic manifolds obtained by the Epstein–Penner's method. Proc. Am. Math. Soc. **129**, 2431–2439 (2001)

14. H. Akiyoahi, M. Sakuma, Comparing two convex Hull constructions for cusped hyperbolic manifolds, in *Kleinian Groups and Hyperbolic 3-Manifolds*, Warwick, 2001. London Mathematical Society Lecture Note Series, vol. 299 (Cambridge University Press, Cambridge, 2003), pp. 209–246
15. H. Akiyoahi, M. Sakuma, M. Wada, Y. Yamashita, *Punctured Torus Groups and 2-Bridge Knot Groups (I)*. Lecture Notes in Mathematics, vol. 1909 (Springer, Berlin, 2007)
16. Y. Akutsu, T. Deguchi, T. Ohtsuki, Invariants of colored links. J. Knot Theory Ramifications **1**, 161–184 (1992)
17. J.W. Alexander, Note on Riemann spaces. Bull. Am. Math. Soc. **26**, 370–372 (1920)
18. J. Anderson, *Hyperbolic Geometry*. Springer Undergraduate Mathematics Series (Springer-Verlag London, Ltd., London, 1999), x+230 pp.
19. R. Benedetti, C. Petronio, *Lectures on Hyperbolic Geometry*. Universitext (Springer, Berlin, 1992), xiv+330 pp.
20. M. Belolipetsky, Hyperbolic orbifolds of small volume, in *Proceedings of the International Congress of Mathematicians*, Seoul, 2014, vol. II (Kyung Moon Sa, Seoul, 2014), pp. 837–851
21. J. Berge, Some knots with surgeries yielding lens spaces. Unpublished manuscript. arXiv:1802.09722 [math.GT]
22. N. Bergeron, M.H. Sengün, A. Venkatesh, Torsion homology growth and cycle complexity of arithmetic manifolds. Duke Math. J. **165**, 1629–1693 (2016)
23. L. Bessiéres, G. Besson, M. Boileau, S, Maillot, J. Porti, *Geometrisation of 3-Manifolds*. EMS Tracts in Mathematics, vol. 13 (European Mathematical Society, Zürich, 2010), x+237 pp.
24. M. Bestvina, $\mathbb{R}$-trees in topology, geometry, and group theory, in *Handbook of Geometric Topology* (North-Holland, Amsterdam, 2002), pp. 55–91
25. M. Bestvina, Geometric group theory and 3-manifolds hand in hand: the fulfillment of Thurston's vision. Bull. Am. Math. Soc. **51**, 53–70 (2014)
26. S. Bleiler, C. Hodgson, Spherical space forms and Dehn filling. Topology **35**, 809–833 (1996)
27. S. Bleiler, C. Hodgson, J.R. Weeks, Cosmetic surgery on knots, in *Proceedings of the Kirbyfest*, Berkeley, CA, 1998. Geometry & Topology Monographs, vol. 2 (Geometry & Topology Publications, Coventry, 1999), pp. 23–34
28. M. Boileau, E. Flapan, Uniqueness of free actions on S^3 respecting a knot. Can. J. Math. **39**, 969–982 (1987)
29. M. Boileau, E. Flapan, On π-hyperbolic knots which are determined by their 2-fold and 4-fold cyclic branched coverings. Topol. Appl. **61**, 229–240 (1995)
30. M. Boileau, S. Friedl, The profinite completion of 3-manifold groups, fiberedness and the Thurston norm. arXiv:1505.07799 [math.GT]
31. M. Boileau, S. Friedl, Grothendieck rigidity of 3-manifold groups. Groups Geom. Dyn. **13**, 1133–1150 (2019)
32. M. Boileau, J. Porti, Geometrization of 3-orbifolds of cyclic type. Astérisque **272**, 208 (2001). Appendix A by M. Heusener, J. Porti
33. M. Boileau, B. Zimmermann, Symmetries of nonelliptic Montesinos links. Math. Ann. **277**, 563–584 (1987)
34. M. Boileau, B. Zimmermann, The π-orbifold group of a link. Math. Z. **200**, 187–208 (1989)
35. M. Boileau, S, Maillot, J. Porti, Three-dimensional orbifolds and their geometric structures, Panoramas et Synthèses, vol. 15 (Société Mathématique de France, Paris, 2003), viii+167 pp.
36. M. Boileau, B. Leeb, J. Porti, Geometrization of 3-dimensional orbifolds. Ann. Math. **162**, 195–290 (2005)
37. M. Boileau, S. Boyer, R. Cebanu, G.S. Walsh, Knot commensurability and the Berge conjecture. Geom. Topol. **16**(2), 625–664 (2012)
38. M. Boileau, S. Boyer, R. Cebanu, G.S. Walsh, Knot complements, hidden symmetries and reflection orbifolds, Ann. Fac. Sci. Toulouse Math. (6) **24**, 1179–1201 (2015)
39. M. Boileau, C. Franchi, M. Mecchia, L. Paoluzzi, B. Zimmermann, Finite group actions on 3-manifolds and cyclic branched coverings of knots. J. Topol. **11**, 283–308 (2018)

40. F. Bonahon, Geometric structures on 3-manifolds, in *Handbook of Geometric Topology* (North-Holland, Amsterdam, 2002), pp. 93–164
41. F. Bonahon, L. Siebenmann, The characteristic toric splitting of irreducible compact 3-orbifolds. Math. Ann. **278**, 441–479 (1987)
42. F. Bonahon, L. Siebenmann, New geometric splittings of classical knots, and the classification and symmetries of arborescent knots. http://www-bcf.usc.edu/~fbonahon/Research/Publications.html
43. A. Borel, Commensurability classes and volumes of hyperbolic 3-manifolds. Ann. Scuola Norm. Sup. Pisa **8**, 1–33 (1981)
44. B.H. Bowditch, The Cannon–Thurston map for punctured-surface groups. Math. Z. **255**, 35–76 (2007)
45. S. Boyer, Dehn surgery on knots, in *Handbook of Geometric Topology* (North-Holland, Amsterdam, 2002), pp. 165–218
46. S. Boyer, X. Zhang, On Culler–Shalen seminorms and Dehn filling. Ann. Math. **148**, 737–801 (1998)
47. S. Boyer, X. Zhang, A proof of the finite filling conjecture. J. Differ. Geom. **59**, 87–176 (2001)
48. S. Boyer, X. Zhang, Every nontrivial knot in S^3 has nontrivial A-polynomial. Proc. Am. Math. Soc. **133**, 2813–2815 (2005)
49. J. Brock, N. Dunfield, Norms on the cohomology of hyperbolic 3-manifolds. Invent. Math. **210**, 531–558 (2017)
50. J. Brock, R.D. Canary, Y.N. Minsky, The classification of Kleinian surface groups, II: The ending lamination conjecture. Ann. Math. **176**, 1–149 (2012)
51. R. Budney, JSJ-decompositions of knot and link complements in S^3. Enseign. Math. **52**, 319–359 (2006)
52. G. Burde, H. Zieschang, *Knots*. De Gruyter Studies in Mathematics, vol. 5 (Walter de Gruyter & Co., Berlin, 1985), xii+399 pp.
53. G. Burde, H. Zieschang, M. Heusener, *Knots. Third, Fully Revised and Extended Edition*. De Gruyter Studies in Mathematics, vol. 5 (De Gruyter, Berlin, 2014)
54. P.J. Callahan, A.W. Reid, Hyperbolic structures on knot complements. Chaos Solitons Fractals **9**, 705–738 (1998)
55. R.D. Canary, D.B.A. Epstein, P. Green, Notes on Notes of Thurston, in *Analytical and Geometric Aspects of Hyperbolic Space* (Coventry/Durham, 1984), pp. 3–92. London Mathematical Society Lecture Note Series, vol. 111 (Cambridge University Press, Cambridge, 1987)
56. J.W. Cannon, W. Dicks, On hyperbolic once-punctured-torus bundles II. Geom. Dedicata **126**, 11–63 (2006). Errata and addenda: http://mat.uab.es/dicks/Cannon.html
57. J.W. Cannon, W.P. Thurston, Group invariant Peano curves. Geom. Topol. **11**, 1315–1355 (2007)
58. C. Cao, G.R. Meyerhoff, The orientable cusped hyperbolic 3-manifolds of minimum volume. Invent. Math. **146**, 451–478 (2001)
59. H.-D. Cao, X.-P. Zhu, A complete proof of the Poincaré and geometrization conjectures-application of the Hamilton–Perelman theory of the Ricci flow. Asian J. Math. **10**, 165–492 (2006). Erratum, ibid, 663
60. S.-S. Chern, J. Simons, Characteristic forms and geometric invariants. Ann. Math. (2) **99**, 48–69 (1974)
61. E. Chesebro, J. DeBlois, Algebraic invariants, mutation, and commensurability of link complements. Pac. J. Math. **267**, 341–398 (2014)
62. Y.E. Choi, Positively oriented ideal triangulations on hyperbolic three-manifolds. Topology **43**, 1345–1371 (2004)
63. J.H. Conway, An enumeration of knots and links, and some of their algebraic properties, in *1970 Computational Problems in Abstract Algebra* (Proceedings Conference, Oxford, 1967) (Pergamon Press, 1970), pp. 329–358
64. D. Cooper, M. Culler, H. Gillet, D.D. Long, P.B. Shalen, Plane curves associated to character varieties of 3-manifolds. Invent. Math. **118**, 47–84 (1994)

65. D. Cooper, C. Hodgson, S. Kerckhoff, *Three-Dimensional Orbifolds and Cone-Manifolds*. MSJ Memoirs, vol. 5, (Mathematical Society of Japan, Tokyo, 2000), x+170 pp.
66. F. Costantino, On a proof of the JSJ theorem. Rend. Sem. Mat. Univ. Politec. Torino **60**, 129–146 (2002/2003)
67. F. Costantino, D. Thurston, 3-manifolds efficiently bound 4-manifolds. J. Topol. **1**, 703–745 (2008)
68. D. Coulson, O. Goodman, C. Hodgson, W. Neumann, Computing arithmetic invariants of 3-manifolds. Exp. Math. **9**, 127–152 (2000)
69. P.R. Cromwell, *Knots and Links* (Cambridge University Press, Cambridge, 2004), xviii+328 pp.
70. R. Crowell, Genus of alternating link types. Ann. Math. **69**, 258–275 (1959)
71. R. Crowell, R. Fox, *Introduction to Knot Theory*. Graduate Texts in Mathematics, vol. 57 (Springer, New York/Heidelberg, 1977), x+182 pp. Reprint of the 1963 original
72. M. Culler, Lifting representations to covering groups. Adv. Math. **59**, 64–70 (1986)
73. M. Culler, P.B. Shalen, Varieties of group representations and splittings of 3-manifolds. Ann. Math. **117**, 109–146 (1983)
74. M. Culler, P.B. Shalen, Bounded, separating, incompressible surfaces in knot manifolds. Invent. Math. **75**, 537–545 (1984)
75. M. Culler, N. Dunfield, M. Goerner, *SnapPy*. Computer Software available at https://www.math.uic.edu/t3m/SnapPy/
76. M. Culler, C.McA. Gordon, J. Luecke, P.B. Shalen, Dehn surgery on knots. Ann. Math. **125**, 237–300 (1987)
77. A. Deruelle, K. Miyazaki, K. Motegi, Networking Seifert surgeries on knots. Mem. Am. Math. Soc. **217**(1021), viii+130 pp. (2012)
78. W. Dicks, M. Sakuma, On hyperbolic once-punctured-torus bundles III: comparing two tessellations of the complex plane. Topol. Appl. **157**, 1873–1899 (2010)
79. J. Dinkelbach, B. Leeb, Equivariant Ricci flow with surgery and applications to finite group actions on geometric 3-manifolds. Geom. Topol. **13**, 1129–1173 (2009)
80. N. Dunfield, S. Garoufalidis, Non-triviality of the A-polynomial for knots in S^3. Algebr. Geom. Topol. **4**, 1145–1153 (2004)
81. N. Dunfield, S. Friedl, N. Jackson, Twisted Alexander polynomials of hyperbolic knots. Exp. Math. **21**, 329–352 (2012)
82. D. Epstein, R. Penner, Euclidean decompositions of noncompact hyperbolic manifolds. J. Differ. Geom. **27**, 67–80 (1988)
83. J.B. Etnyre, Lectures on open book decompositions and contact structures, in *Floer Homology, Gauge Theory, and Low Dimensional Topology*. Clay Mathematics Proceedings, vol. 5 (American Mathematical Society, Providence, RI, 2006), pp. 103–141
84. B. Farb, D. Margalit, *A Primer on Mapping Class Groups*. Princeton Mathematical Series, vol. 49 (Princeton University Press, Princeton, 2012), xiv+472 pp.
85. A. Fathi, F. Laudenbach, V. Poénaru, *Travaux de Thurston sur les Surfaces*, 2nd edn. Astérisque (Société Mathématique de France, Paris, 1991), pp. 66–67
86. R.H. Fox, Free differential calculus III. Ann. Math. **59**, 195–210 (1954)
87. R.H. Fox, A quick trip through knot theory, in *Topology of 3-Manifolds and Related Topics*, ed. by M.K. Fort (Prentice-Hall, Englewood Cliff, 1961), pp. 120–167
88. R.H. Fox, Some problems in knot theory, in *Topology of 3-Manifolds and Related Topics* (Proceedings of The University of Georgia Institute, 1961), pp. 168–176 (Prentice-Hall, Englewood Cliffs, 1962)
89. R.H. Fox, N. Smythe, An ideal class invariant of knots. Proc. Am. Math. Soc. **15**, 707–709 (1964)
90. F. Frankl, L. Pontrjagin, Ein Knotensatz mit Anwendung auf die Dimensionstheorie. Math. Ann. **102**, 785–789 (1930)
91. S. Friedl, S. Viddusi, A survey of twisted Alexander polynomials, in *The Mathematics of Knots*. Contributions in Mathematical and Computational Sciences, vol. 1 (Springer, Heidelberg, 2011), pp. 45–94

92. S. Friedl, S. Viddusi, Twisted Alexander polynomials detect fibered 3-manifolds. Ann. Math. **173**, 1587–1643 (2011)
93. S. Friedl, S. Viddusi, The Thurston norm and twisted Alexander polynomials. J. Reine Angew. Math. **707**, 87–102 (2015)
94. L. Funar, Torus bundles not distinguished by TQFT invariants. Geom. Topol. **17**, 2289–2344 (2013)
95. D. Futer, E. Kalfagianni, J.S. Purcell, Dehn filling, volume, and the Jones polynomial. J. Differ. Geom. **78**, 429–464 (2008)
96. D. Futer, E. Kalfagianni, J.S. Purcell, A survey of hyperbolic knot theory, in *Knots, Low-Dimensional Topology and Applications*. Springer Proceedings in Mathematics & Statistics, vol. 284 (Springer, Cham, 2019), pp. 1–30
97. D. Futer, S.J. Taylor, W. Worden, Random veering triangulations are not geometric. arXiv:1808.05586 [math.GT]
98. D. Gabai, Foliations and the topology of 3-manifolds. J. Differ. Geom. **18**, 445–503 (1983)
99. D. Gabai, The Murasugi sum is a natural geometric operation, in *Low-Dimensional Topology*, San Francisco, CA, 1981. Contemporary Mathematics, vol. 20 (American Mathematical Society, Providence, 1983), pp. 131–143
100. D. Gabai, Foliations and genera of links. Topology **23**, 381–394 (1984)
101. D. Gabai, Genera of the arborescent links. Mem. Am. Math. Soc. **59**(339), i–viii and 1–98 (1986)
102. D. Gabai, Foliations and the topology of 3-manifolds. III. J. Differ. Geom. **26**, 479–536 (1987)
103. D. Gabai, R. Meyerhoff, P. Milley, Minimum volume cusped hyperbolic three-manifolds. J. Am. Math. Soc. **22**, 1157–1215 (2009)
104. D. Gabai, R. Meyerhoff, P. Milley, Mom technology and volumes of hyperbolic 3-manifolds. Comment. Math. Helv. **86**, 145–188 (2011)
105. D. Gabai, R. Meyerhoff, P. Milley, Volumes of hyperbolic 3-manifolds, in *The Poincaré Conjecture*. Clay Mathematics Proceedings, vol. 19 (American Mathematical Society, Providence, RI, 2014), pp. 65–79
106. F.W. Gehring, G.J. Martin, Minimal co-volume hyperbolic lattices. I. The spherical points of a Kleinian group. Ann. Math. **170**, 123–161 (2009)
107. E. Giroux, Géométrie de contact: de la dimension trois vers les dimensions supérieures, in *Proceedings of the International Congress of Mathematicians* (Beijing, 2002), vol. II (Higher Education Press, Beijing, 2002), pp. 405–414
108. H. Goda, M. Teragaito, Dehn surgeries on knots which yield lens spaces and genera of knots. Math. Proc. Camb. Philos. Soc. **129**, 501–515 (2000)
109. F. González-Acuña, H. Short, Cyclic branched coverings of knots and homology spheres. Revista Math. **4**, 97–120 (1991)
110. O. Goodman, D. Heard, C. Hodgson, Commensurators of cusped hyperbolic manifolds. Exp. Math. **17**, 283–306 (2008)
111. C.McA. Gordon, Knots whose branched coverings have periodic homology. Trans. Am. Math. Soc. **168**, 357–370 (1972)
112. C.McA. Gordon, Dehn filling: a survey, in *Knot Theory* (Warsaw, 1995). Banach Center Publications, vol. 42 (Polish Academy of Sciences Institute of Mathematics, Warsaw, 1998), pp. 129–144
113. C.McA. Gordon, Exceptional Dehn filling, in *Introductory Lectures on Knot Theory*. Series on Knots Everything, vol. 46 (World Scientific Publishing, Hackensack, 2012), pp. 124–134
114. C.McA. Gordon, J. Luecke, Knots are determined by their complements. J. Am. Math. Soc. **2**, 371–415 (1989)
115. J.E. Greene, Lattices, graphs, and Conway mutation. Invent. Math. **192**, 717–750 (2013)
116. J.E. Greene, L-space surgeries, genus bounds, and the cabling conjecture. J. Differ. Geom. **100**, 491–506 (2015)
117. M. Gromov, Volume and bounded cohomology. Inst. Hautes Études Sci. Publ. Math. **56**, 5–99 (1982/1983)

118. M. Gromov, W. Thurston, Pinching constants for hyperbolic manifolds. Invent. Math. **89**, 1–12 (1987)
119. A. Grothendieck, Représentations linéaires et compactification profinie des groupes discrets. Manuscripta Math. **2**, 375–396 (1970)
120. F. Guéritaud, On canonical triangulations of once-punctured torus bundles and two-bridge link complements. With an appendix by David Futer. Geom. Topol. **10**, 1239–1284 (2006)
121. F. Guéritaud, Géometrie hyperbolique effective et triangulations idéales canoniques en dimension 3. Ph.D. Thesis (2006)
122. F. Guéritaud, Veering triangulations and Cannon–Thurston maps. J. Topol. **9**, 957–983 (2016)
123. R. Hartley, Identifying noninvertible knots. Topology **22**, 137–145 (1983)
124. A. Hatcher, Hyperbolic structures of arithmetic type on some link complements. J. Lond. Math. Soc. **27**, 345–355 (1983)
125. A. Hatcher, Notes on Basic 3-Manifold Topology. http://www.math.cornell.edu/hatcher/3M/3Mdownloads.html
126. A. Hatcher, W. Thurston, Incompressible surfaces in 2-bridge knot complements. Invent. Math. **79**, 225–246 (1985)
127. J. Hempel, *3-Manifolds*. Annals of Mathematics Studies, vol. 86 (Princeton University Press/University of Tokyo Press, Princeton/Tokyo, 1976), xii+195 pp.
128. J. Hempel, Residual finiteness for 3-manifolds, in *Combinatorial Group Theory and Topology*, Alta, Utah, 1984. Annals of Mathematics Studies, vol. 111 (Princeton University Press, Princeton, 1987), pp. 379–396
129. J. Hempel, Some 3-manifold groups with the same finite quotients. arXiv:1409.3509 [math.GT]
130. H.M. Hilden, M.T. Lozano, J.M. Montesinos-Amilibia, On knots that are universal. Topology **24**, 499–504 (1985)
131. H.M. Hilden, M.T. Lozano, J.M. Montesinos-Amilibia, W.C. Whitten, On universal groups and three-manifolds. Invent. Math. **87**, 441–456 (1987)
132. H.M. Hilden, M.T. Lozano, J.M. Montesinos-Amilibia, On universal hyperbolic orbifold structures in S^3 with the Borromean rings as singularity. Hiroshima Math. J. **40**, 357–370 (2010)
133. J.A. Hillman, *Alexander Ideals of Links*. Lecture Notes in Mathematics, vol. 895 (Springer, Berlin/New York, 1981), v+178 pp.
134. E. Hironaka, Author Interview: Dale Rolfsen. Blog: Book Ends - Conversation about math books. https://blogs.ams.org/bookends/
135. C.D. Hodgson, S.P. Kerckhoff, Rigidity of hyperbolic cone-manifolds and hyperbolic Dehn surgery. J. Differ. Geom. **48**, 1–59 (1998)
136. C.D. Hodgson, S.P. Kerckhoff, Universal bounds for hyperbolic Dehn surgery. Ann. Math. **162**, 367–421 (2005)
137. C.D. Hodgson, S.P. Kerckhoff, The shape of hyperbolic Dehn surgery space. Geom. Topol. **12**, 1033–1090 (2008)
138. C. Hodgson, H. Masai, On the number of hyperbolic 3-manifolds of a given volume, in *Geometry and Topology Down Under*. Contemporary Mathematics, vol. 597 (American Mathematical Society, Providence, 2013), pp. 295–320
139. C.D. Hodgson, H.J. Rubinstein, Involutions and isotopies of lens spaces, in *Knot Theory and Manifolds* (Vancouver, BC, 1983). Lecture Notes in Mathematics, vol. 1144 (Springer, Berlin, 1985), pp. 60–96
140. C.D. Hodgson, J.R. Weeks, Symmetries, isometries and length spectra of closed hyperbolic three-manifolds. Exp. Math. **3**, 261–274 (1994)
141. C.D. Hodgson, H.J. Rubinstein, H. Segerman, Triangulations of hyperbolic 3-manifolds admitting strict angle structures. J. Topol. **5**, 887–908 (2012)
142. C. Hodgson, A. Issa, H. Segerman, Non-geometric veering triangulations. Exp. Math. **25**, 17–45 (2016)
143. N. Hoffman, K. Ichihara, M. Kashiwagi, H. Masai, S. Oishi, A. Takayasu, Verified computations for hyperbolic 3-manifolds. Exp. Math. **25**, 66–78 (2016)

144. J. Hoste, M. Thistlethwaite, J. Weeks, The first 1, 701, 936 knots. Math. Intell. **20**, 33–48 (1998)
145. K. Ichihara, H. Masai, Exceptional surgeries on alternating knots. Comm. Anal. Geom. **24**, 337–377 (2016)
146. Y. Imayoshi, M. Taniguchi, *An Introduction to Teichmüller Spaces*. Translated and Revised from the Japanese by the Authors (Springer, Tokyo, 1992), xiv+279 pp.
147. M. Ishikawa, Y. Koda, Stable maps and branched shadows of 3-manifolds. Math. Ann. **367**, 1819–1863 (2017)
148. W. Jaco, *Lectures on Three-Manifold Topology*. CBMS Regional Conference Series in Mathematics, vol. 43 (American Mathematical Society, Providence, 1980), xii+251 pp.
149. W.H. Jaco, P.B. Shalen, Seifert fibered spaces in 3-manifolds. Mem. Am. Math. Soc. **21**(220), viii+192 pp. (1979)
150. K. Johannson, *Homotopy Equivalences of 3-Manifolds with Boundaries*. Lecture Notes in Mathematics, vol. 761 (Springer, Berlin, 1979), ii+303 pp.
151. V.F.R. Jones, A polynomial invariant for knots via von Neumann algebras. Bull. Am. Math. Soc. (N.S.) **12**, 103–111 (1985)
152. T. Jørgensen, On pairs of punctured tori. Unfinished manuscript. Available in *Proceedings of the Workshop "Kleinian Groups and Hyperbolic 3-Manifolds"*, ed. by Y. Komori, V. Markovic, C. Series. London Mathematical Society Lecture Note Series, vol. 299 (2003), pp. 183–207
153. A. Juhász, A survey of Heegaard Floer homology, in *New Ideas in Low Dimensional Topology*. Series on Knots Everything, vol. 56 (World Scientific Publishing, Hackensack, 2015), pp. 237–296
154. J. Kahn, V. Markovic, Immersing almost geodesic surfaces in a closed hyperbolic three manifold. Ann. Math. **175**, 1127–1190 (2012)
155. T. Kanenobu, The augmentation subgroup of a pretzel link. Math. Sem. Notes Kobe Univ. **7**, 363–384 (1979)
156. T. Kanenobu, Unions of knots as cross sections of 2-knots. Kobe J. Math. **4**, 147–162 (1988)
157. M. Kapovich, *Hyperbolic Manifolds and Discrete Groups: Notes on Thurston's Hyperbolization*. Progress in Mathematics, vol. 183 (Birkhäuser, Basel, 2000)
158. R.M. Kashaev, A link invariant from quantum dilogarithm. Modern Phys. Lett. A **10**, 1409–1418 (1995)
159. L.H. Kauffman, *On Knots*. Annals of Mathematics Studies, vol. 115 (Princeton University Press, Princeton, 1987), pp. xvi+481
160. A. Kawauchi, The invertibility problem on amphicheiral excellent knots. Proc. Jpn. Acad. **55**, 399–402 (1979)
161. A. Kawauchi, *A Survey of Knot Theory*. Translated and revised from the 1990 Japanese original by the author (Birkhäuser Verlag, Basel, 1996)
162. A. Kawauchi, Topological imitations and Reni–Mecchia–Zimmermann's conjecture. Kyungpook Math. J. **46**, 1–9 (2006)
163. D.A. Kazhdan, G.A. Margulis, A proof of Selberg's conjecture. Math. USSR-Sbornik **4**, 147–152 (1968)
164. R. Kellerhals, Hyperbolic orbifolds of minimal volume. Comput. Methods Funct. Theory **14**, 465–481 (2014)
165. S. Kinoshita, On knots and periodic transformations. Osaka Math. J. **10**, 43–52 (1958)
166. S. Kinoshita, H. Terasaka, On unions of knots. Osaka Math. J. **9**, 131–153 (1957)
167. R. Kirby, Problems in low dimensional manifold theory, in *Algebraic and Geometric Topology* (Proceedings of the Symposium in Pure Mathematics, Stanford University, Stanford, CA, 1976), Part 2. Proceedings of the Symposium in Pure Mathematics, vol. XXXII (American Mathematical Society, Providence, 1978), pp. 273–312
168. T. Kitano, Twenty years of twisted Alexander polynomials: refinement of the Alexander polynomial and its applications (Japanese). Sūgaku **65**, 360–384 (2013)
169. B. Kleiner, J. Lott, Notes on Perelman's papers. Geom. Topol. **12**, 2587–2855 (2008)

170. H. Kneser, Geschlossen Flachen in dreidimensionalen Mannigfaltigkeiten. Jahresbericht der Deutschen Mathematiker Vereinigung **38**, 248–260 (1929)
171. T. Kobayashi, Structures of the Haken manifolds with Heegaard splittings of genus two. Osaka J. Math. **21**, 437–455 (1984)
172. S. Kojima, Determining knots by branched covers, in *Low-Dimensional Topology and Kleinian Groups* (Coventry/Durham, 1984). London Mathematical Society Lecture Note Series, vol. 112 (Cambridge University Press, Cambridge, 1986), pp. 193–207
173. S. Kojima, Isometry transformations of hyperbolic 3-manifolds. Topol. Appl. **29**, 297–307 (1988)
174. P.B. Kronheimer, T.S. Mrowka, Scalar curvature and the Thurston norm. Math. Res. Lett. **4**, 931–937 (1997)
175. M. Lackenby, Word hyperbolic Dehn surgery. Invent. Math. **140**, 243–282 (2000)
176. M. Lackenby, The volume of hyperbolic alternating link complements, with an appendix by Ian Agol and Dylan Thurston. Proc. Lond. Math. Soc. **88**, 204–224 (2004)
177. M. Lackenby, Every knot has characterising slopes. Math. Ann. **374**, 429–446 (2019)
178. M. Lackenby, R. Meyerhoff, The maximal number of exceptional Dehn surgeries. Invent. Math. **191**, 341–382 (2013)
179. T. Le, Homology torsion growth and Mahler measure. Comment. Math. Helv. **89**, 719–757 (2014)
180. T. Le, Growth of homology torsion in finite coverings and hyperbolic volume. Ann. Inst. Fourier **68**, 611–645 (2018)
181. D. Lee, M. Sakuma, Epimorphisms from 2-bridge link groups onto Heckoid groups (I). Hiroshima Math. J. **43**, 239–264 (2013)
182. W.B.R. Lickorish, *An Introduction to Knot Theory*. Graduate Texts in Mathematics, vol. 175 (Springer, New York, 1997), x+201 pp.
183. X.S. Lin, Representations of knot groups and twisted Alexander polynomials. Acta Math. Sin. (Engl. Ser.) **17**, 361–380 (2001)
184. C. Livingston, *Knot Theory*. Carus Mathematical Monographs, vol. 24 (Mathematical Association of America, Washington, 1993), xviii+240 pp.
185. D.D. Long, A.W. Reid, Grothendieck's problem for 3-manifold groups. Groups Geom. Dyn. **5**, 479–499 (2011)
186. W. Lück, L^2-invariants: theory and applications to geometry and KK-theory, in *Ergebnisse der Mathematik und ihrer Grenzgebiete*. 3. Folge. A Series of Modern Surveys in Mathematics, vol. 44 (Springer, Berlin, 2002)
187. F. Luo, Actions of finite groups on knot complements. Pac. J. Math. **154**(2), 317–329 (1992)
188. F. Luo, S. Schleimer, S. Tillman, Geodesic ideal triangulations virtually exists. Proc. Am. Math. Soc. **136**, 2625–2630 (2008)
189. C. Maclachlan, A.W. Reid, in *The Arithmetic of Hyperbolic 3-Manifolds*. Graduate Texts in Mathematics, vol. 219 (Springer, New York, 2003), xiv+463 pp.
190. A. Marden, Deformations of Kleinian groups, in *Handbook of Teichmüller Theory, VOL. I.* IRMA Lectures in Mathematics and Theoretical Physics, vol. 11 (European Mathematical Society, Zürich, 2007), pp. 411–446
191. A. Marden, *Outer Circles. An Introduction to Hyperbolic 3-Manifolds* (Cambridge University Press, Cambridge, 2007), xviii+427 pp.
192. G.A. Margulis, Discrete groups of isometries of manifolds of nonpositive curvature, in *Proceedings of the International Congress of Mathematicians*, Vancouver, vol. 2 (1974), pp. 21–34
193. T.H. Marshall, G.J. Martin, Minimal co-volume hyperbolic lattices, II: Simple torsion in a Kleinian group. Ann. Math. (2) **176**, 261–301 (2012)
194. B. Martelli, *An Introduction to Geometric Topology* (CreateSpace Independent Publishing Platform, Scotts Valley, 2016)
195. G. Martin, The geometry and arithmetic of Kleinian groups, in *Handbook of Group Actions, Vol. I.* Advanced Lectures in Mathematics, vol. 31 (International Press, Somerville, 2015), pp. 411–494

196. G. Martin, Siegel's problem in three dimensions. Not. Am. Math. Soc. **63**, 1244–1247 (2016)
197. K. Matsuzaki, M. Taniguchi, *Hyperbolic Manifolds and Kleinian Groups*. Oxford Mathematical Monographs (Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1998), x+253 pp.
198. J.P. Mayberry, K. Murasugi, Torsion groups of abelian coverings of links. Trans. Am. Math. Soc. **271**, 143–173 (1982)
199. C.T. McMullen, *Renormalization and 3-Manifolds Which Fiber over the Circle*. Annals of Mathematics Studies, vol. 142 (Princeton University Press, Princeton, 1996), x+253 pp.
200. C.T. McMullen, Polynomial invariants for fibered 3-manifolds and Teichmüller geodesics for foliations. Ann. Sci. École Norm. Sup. **33**, 519–560 (2000)
201. C.T. McMullen, The evolution of geometric structures on 3-manifolds. Bull. Am. Math. Soc. **48**, 259–274 (2011)
202. C.T. McMullen, The evolution of geometric structures on 3-manifolds, in *The Poincaré Conjecture*. Clay Mathematics Proceedings, vol. 19 (American Mathematical Society, Providence, RI, 2014), pp. 31–46
203. M. Mecchia, M. Reni, Hyperbolic 2-fold branched coverings of links and their quotients. Pac. J. Math. **202**, 429–447 (2002)
204. W.H. Meeks III, P. Scott, Finite group actions on 3-manifolds. Invent. Math. **86**, 287–346 (1986)
205. W.H. Meeks III, S.T. Yau, The equivariant Dehn's lemma and loop theorem. Comment. Math. Helv. **56**, 225–239 (1981)
206. W.W. Menasco, Polyhedra representation of link complements, in *Low-Dimensional Topology* (San Francisco, CA, 1981). Contemporary Mathematics, vol. 20 (American Mathematical Society, Providence, 1983), pp. 305–325
207. W.W. Menasco, M. Thistlethwaite, *Handbook of Knot Theory* (Elsevier Science, Amsterdam, 2005), 492 pp.
208. R. Meyerhoff, Density of the Chern-Simons invariant for hyperbolic 3-manifolds, *Low-Dimensional Topology and Kleinian Groups* (Coventry/Durham, 1984). London Mathematical Society Lecture Note Series, vol. 112 (Cambridge University Press, Cambridge, 1986), pp. 217–239
209. C. Millichap, Factorial growth rates for the number of hyperbolic 3-manifolds of a given volume. Proc. Am. Math. Soc. **143**, 2201–2214 (2015)
210. C. Millichap, W. Worden, Hidden symmetries and commensurability of 2-bridge link complements. Pac. J. Math. **285**, 453–484 (2016)
211. J. Milnor, A unique decomposition theorem for 3-manifolds. Am. J. Math. **84**, 1–7 (1962)
212. Y.N. Minsky, S.J. Taylor, Fibered faces, veering triangulations, and the arc complex. Geom. Funct. Anal. **27**, 1450–1496 (2017)
213. M. Mj, Cannon–Thurston maps, in *Proceedings of the International Congress of Mathematicians–Rio de Janeiro 2018*, vol. II. Invited Lectures (World Scientific Publishing, Hackensack, 2018), pp. 885–917
214. J.M. Montesinos, Variedades de Seifert que són recubridores cíclicos de dos hojas. Bol. Soc. Mat. Mexicana **18**, 1–32 (1973)
215. J.M. Montesinos, Surgery on links and double branched covers of S^3, in *Knots, Groups, and 3-Manifolds* (Papers dedicated to the memory of R.H. Fox). Annals of Mathematics Studies, vol. 84 (Princeton University Press, Princeton, 1975), pp. 227–259
216. J.M. Montesinos, *Classical Tessellations and Three-Manifolds*. Universitext (Springer, Berlin, 1987), xviii+230 pp.
217. J.W. Morgan, Λ-trees and their applications. Bull. Am. Math. Soc. **26**, 87–112 (1992)
218. J.W. Morgan, Recent progress on the Poincaré conjecture and the classification of 3-manifolds. Bull. Am. Math. Soc. **42**, 57–78 (2005)
219. J.W. Morgan, H. Bass, in *The Smith Conjecture*. Pure and Applied Mathematics, vol. 112 (Academic, Orlando, 1984)
220. J. Morgan, P. Shalen, Valuations, trees, and degenerations of hyperbolic structures. I. Ann. Math. **120**, 401–476 (1984)

221. J. Morgan, P. Shalen, Degenerations of hyperbolic structures. II. Measured laminations in 3-manifolds. Ann. Math. **127**, 403–456 (1988)
222. J. Morgan, P. Shalen, Degenerations of hyperbolic structures. III. Actions of 3-manifold groups on trees and Thurston's compactness theorem. Ann. Math. **127**, 457–519 (1988)
223. J. Morgan, G. Tian, *Ricci Flow and the Poincaré Conjecture*. Clay Mathematics Monographs, vol. 3 (American Mathematical Society/Clay Mathematics Institute, Providence/Cambridge, 2007), xlii+521 pp.
224. J. Morgan, G. Tian, *The Geometrization Conjecture*. Clay Mathematics Monographs, vol. 5 (American Mathematical Society/Clay Mathematics Institute, Providence/Cambridge, 2014), x+291 pp.
225. G.D. Mostow, Quasi-conformal mappings in n-space and the rigidity of hyperbolic space forms. Inst. Hautes Études Sci. Publ. Math. **34**, 53–104 (1968)
226. K. Motegi, Dehn surgery on knots - tracing the evolution of research. Dehn - Sūgaku **70**(4), 379–405 (2018). English translation will appear in "Sugaku Expositions"
227. D. Mumford, C. Series, D. Wright, *Indra's Pearls. The Vision of Felix Klein* (Cambridge University Press, New York, 2002), xx+396 pp.
228. H. Murakami, An introduction to the volume conjecture, in *Interactions Between Hyperbolic Geometry, Quantum Topology and Number Theory*. Contemporary Mathematics, vol. 541 (American Mathematical Society, Providence, 2011), pp. 1–40
229. H. Murakami, Current status of the volume conjecture. Sugaku Expositions **26**, 181–203 (2013)
230. H. Murakami, J. Murakami, The colored Jones polynomials and the simplicial volume of a knot. Acta Math. **186**, 85–104 (2001)
231. H. Murakami, Y. Yokota, *Volume Conjecture for Knots*. SpringerBriefs in Mathematical Physics, vol. 30 (Springer, Singapore, 2018), ix+120 pp.
232. H. Murakami, J. Murakami, M. Okamoto, T. Takata, Y. Yokota, Kashaev's conjecture and the Chern-Simons invariants of knots and links. Exp. Math. **11**, 427–435 (2002)
233. K. Murasugi, On the genus of the alternating knot. I, II. J. Math. Soc. Jpn. **10**, 94–105, 235–248 (1958)
234. K. Murasugi On a certain subgroup of the group of an alternating link. Am. J. Math. **85**, 544–550 (1963)
235. K. Murasugi, *Knot Theory and Its Applications*. Translated from the 1993 Japanese original by Bohdan Kurpita (Birkhüser Boston, Boston, 1996), viii+341 pp.
236. Y. Nakanishi, Primeness of links. Math. Sem. Notes Kobe Univ. **9**, 415–440 (1981)
237. W.D. Neumann, Hilbert's 3rd problem and invariants of 3-manifolds, in *The Epstein Birthday Schrift*. Geometry & Topology Monographs, vol. 1 (Geometry & Topology Publications, Coventry, 1998), pp. 383–411
238. W.D. Neumann, Notes on geometry and 3-manifolds, with appendix by Paul Norbury, in *Low Dimensional Topology*, ed. by K. Böröczky, W. Neumann, A. Stipsicz. Bolyai Society Mathematical Studies, vol. 8 (1999), pp. 191–267
239. W.D. Neumann, A.W. Reid, Amalgamation and the invariant trace field of a Kleinian group. Math. Proc. Camb. Philos. Soc. **109**, 509–515 (1991)
240. W.D. Neumann, A.W. Reid, Arithmetic of hyperbolic manifolds, in *Topology '90* (Columbus, OH, 1990). Ohio State University Mathematical Research Institute Publications, vol. 1 (de Gruyter, Berlin, 1992), pp. 273–310
241. W.D. Neumann, G. Swarup, Canonical decompositions of 3-manifolds. Geom. Topol. **1**, 21–40 (1997)
242. W.D. Neumann, D. Zagier, Volumes of hyperbolic three-manifolds. Topology **24**, 307–332 (1985)
243. Y. Ni, X. Zhang, Finite Dehn surgeries on knots in S^3. Algebr. Geom. Topol. **18**, 441–492 (2018)
244. N. Nikolov, D. Segal, On finitely generated profinite groups. I. Strong completeness and uniform bounds. Ann. Math. **165**, 171–238 (2007)

245. K. Ohshika, *Discrete Groups*. Translations of in Mathematical Monographs. Iwanami Series in Modern Mathematics, vol. 207. (American Mathematical Society, Providence, 2002), x+193 pp. Translated from the 1998 Japanese original by the author.
246. J. P. Otal, *Thurston's Hyperbolization of Haken Manifolds*. Surveys in Differential Geometry, vol. III (Cambridge, MA, 1996) (International Press, Boston, 1998), pp. 77–194
247. J.P. Otal, *The Hyperbolization Theorem for Fibered 3-Manifolds*. Translated from the 1996 French original by L.D. Kay. SMF/AMS Texts and Monographs, vol. 7. (American Mathematical Society/Société Mathématique de France, Providence/Paris, 2001), xiv+126 pp.
248. P. Ozsváth, Z. Szabó, On knot Floer homology and lens space surgeries. Topology **44**, 1281–1300 (2005)
249. P. Ozsváth, Z. Szabó, An introduction to Heegaard Floer homology, in *Floer Homology, Gauge Theory, and Low-Dimensional Topology*. Clay Mathematics Proceedings, vol. 5 (American Mathematical Society, Providence, RI, 2006), pp. 3–27
250. L. Paoluzzi, On hyperbolic type involutions. Rend. Istit. Mat. Univ. Trieste **32**(suppl. 1), 221–256 (2001)
251. L. Paoluzzi, Three cyclic branched covers suffice to determine hyperbolic knots. J. Knot Theory Ramifications **14**, 641–655 (2005)
252. C.D. Papakyriakopoulos, On Dehn's lemma and the asphericity of knots. Ann. Math. **66**, 1–26 (1957)
253. G. Perelman, The entropy formula for the Ricci flow and its geometric applications. arXiv:math/0211159 [math.DG]
254. G. Perelman, Ricci flow with surgery on three-manifolds. arXiv:math/0303109 [math.DG]
255. G. Perelman, Finite extinction time for the solutions to the Ricci flow on certain three-manifolds. arXiv:math/0307245 [math.DG]
256. C. Petronio, J. Porti, Negatively oriented ideal triangulations and a proof of Thurston's hyperbolic Dehn filling theorem. Expo. Math. **18**, 1–35 (2000)
257. G. Prasad, Strong rigidity of Q-rank 1 lattices. Invent. Math. **21**, 255–286 (1973)
258. V.V. Prasolov, A.B. Sossinsky, *Knots, Links, Braids and 3-Manifolds. An Introduction to the New Invariants in Low-Dimensional Topology*. Translated from the Russian manuscript by Sossinsky. Translations of Mathematical Monographs, vol. 154 (American Mathematical Society, Providence, 1997), viii+239 pp.
259. J. Rasmussen, Lens space surgeries and a conjecture of Goda and Teragaito. Geom. Topol. **8**, 1013–1031 (2004)
260. J.G. Ratcliffe, *Foundations of Hyperbolic Manifolds*. Graduate Texts in Mathematics, vol. 149 (Springer, New York, 1994), xii+747 pp.
261. A.W. Reid, A note on trace-fields of Kleinian groups. Bull. Lond. Math. Soc. **22**, 349–352 (1990)
262. A.W. Reid, Arithmeticity of knot complements. J. Lond. Math. Soc. **43**, 171–184 (1991)
263. A.W. Reid, C. Maclachlan, *The Arithmetic of Hyperbolic 3-Manifolds*. Graduate Texts in Mathematics, vol. 219 (Springer, New York, 2003), xiv+463 pp.
264. A. Reid, G.S. Walsh, Commensurability classes of 2-bridge knot complements. Algebr. Geom. Topol. **8**, 1031–1057 (2008)
265. K. Reidemeister, Homotopieringe und Linsenräume. Abh. Math. Sem. Univ. Hamburg **11**, 102–109 (1935)
266. M. Reni, On π-hyperbolic knots with the same 2-fold branched coverings. Math. Ann. **316**, 681–697 (2000)
267. M. Reni, B. Zimmermann, Hyperbolic 3-manifolds as cyclic branched coverings. Comment. Math. Helv. **76**, 300–313 (2001)
268. L. Ribes, P. Zaresskii, Profinite groups, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd series], vol. 40 (Springer, Berlin, 2000), xiv+435 pp.
269. Y. Rieck, Y. Yamashita, Cosmetic surgery and the link volume of hyperbolic 3-manifolds. Algebr. Geom. Topol. **16**, 3445–3521 (2016)
270. R. Riley, Homomorphisms of knot groups on finite groups. Math. Comp. **25**, 603–619 (1971)

271. R. Riley, A quadratic parabolic group. Math. Proc. Camb. Philos. Soc. **77**, 281–288 (1975)
272. R. Riley, An elliptical path from parabolic representations to hyperbolic structures, in *Topology of Low-Dimensional Manifolds* (Proceedings of the Second Sussex Conference, Chelwood Gate, 1977). Lecture Notes in Mathematics, vol. 722 (Springer, Berlin, 1979), pp. 99–133
273. R. Riley, Growth of order of homology of cyclic branched covers of knots. Bull. Lond. Math. Soc. **22**, 287–297 (1990)
274. R. Riley, Algebra for Heckoid groups. Trans. Am. Math. Soc. **334**, 389–409 (1992)
275. D. Rolfsen, *Knots and Links*. Mathematics Lecture Series, vol. 7 (Publish or Perish, Inc., Berkeley, 1976), ix+439 pp.
276. D. Ruberman, Mutation and volumes of knots in S^3. Invent. Math. **90**, 189–215 (1987)
277. T. Sakai, Geodesic knots in a hyperbolic 3-manifold. Kobe J. Math. **8**, 81–87 (1991)
278. M. Sakuma, Uniqueness of symmetries of knots. Math. Z. **192**, 225–242 (1986)
279. M. Sakuma, The geometries of spherical Montesinos links. Kobe J. Math. **7**, 167–190 (1990)
280. M. Sakuma, Homology of abelian coverings of links and spatial graphs. Can. J. Math. **47**, 201–224 (1995)
281. M. Sakuma, J. Weeks, Examples of canonical decompositions of hyperbolic link complements. Jpn. J. Math. (N.S.) **21**, 393–439 (1995)
282. K. Schmidt, *Dynamical Systems of Algebraic Origin*. Progress in Mathematics, vol. 128 (Birkhäuser Verlag, Basel, 1995)
283. H. Schubert, Die eindeutige Zerlegbarkeit eines Knotens in Primknoten. S.-B. Heidelberger Akad. Wiss. Math.-Nat. Kl. **1949**, 57–104 (1949)
284. H. Schubert, Knoten und Vollringe. Acta Math. **90**, 131–286 (1953)
285. H. Schubert, Knoten mit zwei Brücken. Math. Z. **65**, 133–170 (1956)
286. J. Schultens, *Introduction to 3-Manifolds*. Graduate Studies in Mathematics, vol. 151 (American Mathematical Society, Providence, 2014), x+286 pp.
287. P. Scott, The geometries of 3-manifolds. Bull. Lond. Math. Soc. **15**, 401–487 (1983)
288. H. Seifert, Über das Geschlecht von Knoten. Math. Ann. **110**, 571–592 (1934)
289. J.-P. Serre, *Arbres, Amalgames,* SL_2*. Avec un sommaire anglais. Rédigé avec la collaboration de Hyman Bass*. Astérisque, vol. 46 (Société Mathématique de France, Paris, 1977), 189 pp.
290. P.B. Shalen, Representations of 3-manifold groups, in *Handbook of Geometric Topology* (North-Holland, Amsterdam, 2002), pp. 955–1044
291. C.L. Siegel, Discontinuous groups. Ann. Math. **44**, 674–689 (1943)
292. C.L. Siegel, Some remarks on discontinuous groups. Ann. Math. **46**, 708–718 (1945)
293. D. Silver, S. Williams, Mahler measure, links and homology growth. Topology **41**, 979–991 (2002)
294. T. Soma, The Gromov invariant of links. Invent. Math. **64**, 445–454 (1981)
295. J. Stallings, On fibering certain 3-manifolds, in *Topology of 3-Manifolds and Related Topics* (Proceedings of The University of Georgia Institute, 1961) (Prentice-Hall, Englewood Cliffs, 1962), pp. 95–100
296. J. Stallings, Constructions of fibred knots and links, in *Algebraic and Geometric Topology* (Proc. Sympos. Pure Math., Stanford Univ., Stanford, CA, 1976), Part 2. Proc. Sympos. Pure Math., vol. XXXII (American Mathematical Society, Providence, 1978), pp. 55–60
297. M. Takahashi, On the concrete construction of hyperbolic structures of 3-manifolds. Tsukuba J. Math. **9**, 41–83 (1985)
298. H. Terasaka, The world of non-hyperbolic geometry (in Japanese), Bluebacks, vol. B312 (Kodansha, Tokyo, 1977)
299. M.B. Thistlethwaite, Knot tabulations and related topics, in *Aspects of Topology*. London Mathematical Society Lecture Note Series, vol. 93 (Cambridge University Press, Cambridge, 1985), pp. 1–76
300. W.P. Thurston, *The Geometry and Topology of Three-Manifolds*. Lecture Notes (Princeton University Press, Princeton, 1976–1980)
301. W.P. Thurston, *Three Manifolds with Symmetry* (1981). Preprint

302. W.P. Thurston, Three-dimensional manifolds, Kleinian groups and hyperbolic geometry. Bull. Am. Math. Soc. **6**, 357–381 (1982)
303. W.P. Thurston, Universal Links (1882). Preprint
304. W.P. Thurston, A norm for the homology of 3-manifolds. Mem. Am. Math. Soc. **59**(339), i–vi and 99–130 (1986)
305. W.P. Thurston, Hyperbolic structures on 3-manifolds. I. Deformation of acylindrical manifolds. Ann. Math.**124**, 203–246 (1986)
306. W.P. Thurston, On the geometry and dynamics of diffeomorphisms of surfaces. Bull. Am. Math. Soc. **19**, 417–431 (1988)
307. W.P. Thurston, *Three-Dimensional Geometry and Topology*, Vol. 1, ed. by S. Levy. Princeton Mathematical Series, vol. 35 (Princeton University Press, Princeton, 1997), x+311 pp.
308. W.P. Thurston, Hyperbolic structures on 3-manifolds, II: surface groups and 3-manifolds which fiber over the circle. arXiv:math/9801045
309. H.F. Trotter, Non-invertible knots exist. Topology **2**, 275–280 (1963)
310. V.G. Turaev, Shadow links and face models of statistical mechanics. J. Differ. Geom. **36**, 35–74 (1992)
311. V.G. Turaev, *Quantum Invariants of Knots and 3-Manifolds*. de Gruyter Studies in Mathematics, vol. 18 (Walter de Gruyter & Co., Berlin, 1994)
312. O.Ja. Viro, Nonprojecting isotopies and knots with homeomorphic coverings. Studies in topology, II. Zap. Naucn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **66**, 133–147, 207–208 (1976)
313. M. Wada, Twisted Alexander polynomial for finitely presentable groups. Topology **33**, 241–256 (1994)
314. M. Wada, OPTi. Computer Software available at http://delta-mat.ist.osaka-u.ac.jp/OPTi/
315. M. Wada, Y. Yamashita, H. Yoshida, An inequality for polyhedra and ideal triangulations of cusped hyperbolic 3-manifolds. Proc. Am. Math. Soc. **124**, 3905–3911 (1996)
316. F. Waldhausen, Eine Klasse von 3-dimensionalen Mannigfaltigkeiten. I, II. Invent. Math. **3**, 308–333 (1967); Invent. Math. **4**, 87–117 (1967)
317. F. Waldhausen, On irreducible 3-manifolds which are sufficiently large. Ann. Math. **87**, 56–88 (1968)
318. G.S. Walsh, Great circle links and virtually fibered knots. Topology **44**, 947–958 (2005)
319. H.C. Wang, Topics on totally discontinuous groups, in *Symmetric Spaces* (Short Courses, Washington University, St. Louis, MO, 1969–1970). Pure and Applied Mathematics, vol. 8 (Dekker, New York, 1972), pp. 459–487
320. J. Weeks, SnapPea. Computer Software available at http://www.geometrygames.org/SnapPea/
321. J. Weeks, Convex hulls and isometries of cusped hyperbolic 3-manifolds, Topol. Appl. **52**, 127–149 (1993)
322. J. Weeks, Computation of hyperbolic structures in knot theory, in *Handbook of Knot Theory* (Elsevier B.V., Amsterdam, 2005), pp. 461–480
323. J.H.C. Whitehead, On doubled knots. J. Lond. Math. Soc. **12**, 63–71 (1937)
324. W. Whitten, Knot complements and groups. Topology **26**, 41–44 (1987)
325. T. Yoshida, The η-invariant of hyperbolic 3-manifolds. Invent. Math. **81**, 473–514 (1985)
326. H. Yoshida, Ideal tetrahedral decompositions of hyperbolic 3-manifolds. Osaka J. Math. **33**, 37–46 (1996)
327. K. Yoshida, The minimal volume orientable hyperbolic 3-manifold with 4 cusps. Pac. J. Math. **266**, 457–476 (2013)
328. B. Zimmermann, On hyperbolic knots with homeomorphic cyclic branched coverings. Math. Ann. **311**, 665–673 (1998)

Chapter 4
Thurston's Theory of 3-Manifolds

Sadayoshi Kojima

Abstract This chapter briefly presents Thurston's contribution to the theory of 3-manifolds and recent developments afterwards exclusively. We focus particularly on the 3-manifold theory and will not discuss other contributions to mathematics or science in general by Thurston.

Keywords 3-Manifolds · Hyperbolic geometry · Geometrization conjecture · Virtual Haken conjecture · Virtual fiber conjecture

2010 Mathematics Subject Classification Primary 57M50; Secondary 57M60

4.1 Prologue

The aim of this chapter is to briefly present Thurston's contribution to the theory of 3-manifolds and recent developments afterwards exclusively. We focus particularly on the 3-manifold theory and will not discuss other contributions to mathematics or science in general by Thurston. However, we hope the readers will realize how the study of 3-manifolds was dramatically accelerated by Thurston.

4.2 Pre-Thurston Era

Throughout this chapter, we assume for simplicity that all manifolds discussed are compact and orientable unless otherwise stated.

The study of topology of manifolds has a history of more than 120 years. It was initiated by Poincaré in his paper entitled "Analysis Situs" in 1895 and five

S. Kojima (✉)
Global Center for Science and Engineering, Waseda University, Tokyo, Japan
e-mail: sadayosi@aoni.waseda.jp

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_4

subsequent supplements appearing between 1899 and 1904. In the last supplement [33], Poincaré asked if every closed simply connected 3-manifold is homeomorphic to the 3-sphere. This is a type of question in mathematics which asks if every object with a particular property is very special and which looks too specific at first glance. However, surprisingly enough, this question later called the Poincaré conjecture motivated extensive studies of the topology of manifolds in the last century and led to the development of the basis of algebraic topology, homology theory and homotopy theory. On the other hand, the study of 3-manifolds in particular walked a bit different path from the mainstream of topology while elaborating on this claim.

Among the many trials in 3-manifold topology up to the 1960s, the best success in the structure theory was the theorem by Waldhausen [45] for Haken manifolds. Recall that an essential 2-sphere or a 2-disk in a 3-manifold is one which does not bound a 3-ball. A Haken manifold is, by definition, a 3-manifold without separating essential 2-spheres but with either essential 2-spheres, 2-disks or connected essential (π_1-injective) embedded surfaces of nonpositive Euler characteristic. A 3-manifold without separating essential 2-spheres is called irreducible, and irreducibility was used in those days to avoid a possible counterexample to the Poincaré conjecture. The theorem by Waldhausen asserts that a homotopy equivalence between Haken manifolds preserving peripheral structures is homotopic to a homeomorphism. A Haken manifold admits a hierarchy to a union of 3-balls by cutting along essential embedded surfaces and this property was a key ingredient of Waldhausen's proof.

The Poincaré conjecture can be rephrased as whether a 3-manifold homotopy equivalent to the 3-sphere is homeomorphic to the 3-sphere. An easy extension of the Poincaré conjecture could be stated as: are homotopy equivalent 3-manifolds homeomorphic to each other? This question was known to have counterexamples in lens spaces, but Waldhauen answered the question in the affirmative for Haken 3-manifolds, the class which was expected to occupy a large portion of the world of 3-manifolds.

Waldhausen's structural study was continued by Jaco–Shalen and Johansson for 3-manifolds with essential embedded tori in the later half of the 1970s in [16, 17]. To be more precise, recall that a Seifert fibered space is, by definition, a circle bundle over a 2-dimensional orbifold as an orbifold. Seifert fibered spaces had been well studied by experts. The sphere, lens spaces and many other 3-manifolds with finite π_1 were known to be Seifert fibered, and today we know that so is every 3-manifolds with finite π_1. Those are non-Haken manifolds. There is also a class of non-Haken Seifert fibered spaces with infinite π_1, which in those days were sometimes called small. A neighborhood of an essential embedded torus in a 3-manifold forms a Seifert fibered part. This part may be or may not be a maximal Seifert fibered part in the 3-manifold. What was shown by Jaco–Shalen and Johannson's work is that the maximal Seifert fibered part in a Haken manifold is unique up to isotopy, and its complement (if it exists) is atoroidal, meaning that every essential embedding from the torus is homotopic to some boundary of the maximal Seifert part.

Here we need to define a stronger atoroidality to go further. A 3-manifold is said to be homotopically atoroidal if any $\mathbb{Z} \times \mathbb{Z}$ in π_1 is conjugate to a subgroup of the group generated by the toroidal boundary. Homotopically atoroidal manifolds either have no $\mathbb{Z} \times \mathbb{Z}$ in π_1 or contain an essential torus homotopic to the boundary. In particular, they are atoroidal. Meanwhile small Seifert fibered spaces are atoroidal but not homotopically atoroidal.

We can now state two critical questions about the structure of 3-manifolds, which arose when Jaco–Shalen and Johannson's torus decomposition theory was established:

1. Are there non-Haken 3-manifolds that are not Seifert fibered?
2. What do homotopically atoroidal manifolds look like?

These questions seemed to be very wide open in those days even with the machinery supporting the successful structural study.

4.3 Thurston Era

The period in which Thurston himself proved serious results on 3-manifold theory was probably between the later half of the 1970s and the early 1980s, if we ignore influences by what he said in public and his unseen attempts in succeeding years. After establishing surprising contributions in foliation theory, Thurston moved to 3-manifold theory in the mid-1970s. He changed the research direction of the theory of 3-manifolds during a very short period by introducing completely new geometric aspects. He had special insight into hyperbolic geometry, the geometry of constant negative sectional curvature, and brought his own vision to 3-manifold topology.

The first surprise by Thurston in 3-manifold theory was an elementary but extremely original study of the Dehn surgeries on the figure eight knot discussed in §4 of his seminal lecture notes [39]. The figure eight knot complement can be obtained by appropriately pasting two ideal regular hyperbolic tetrahedra. Through a very naive but ingenious argument using the deformation of tetrahedra, Thurston showed that Dehn surgeries on the figure eight knot produce non-Haken and hyperbolic (and hence non-Seifert) 3-manifolds except in ten cases. This answers the first question of the previous section in a rather generic form. Moreover, Thurston immediately generalized his argument based on pioneering use of character varieties, studied extensively in these days, in §5 of [39] as follows.

Theorem 1 (Hyperbolic Dehn Surgery Theorem) *All but finitely many Dehn surgeries on a hyperbolic knot produce hyperbolic manifolds.*

This theorem should be stated as one consequence of Thurston's results on the flexibility of hyperbolic structures on cusped manifolds. Mostow–Prasad rigidity [27, 34] claims that the hyperbolic structure is rigid if we stay within complete structures in dimension at least 3. However, Thurston found that hyperbolic structures are flexible on 3-manifolds if we allow incomplete metrics with reasonably understandable singularities in their completion. This highly original observation predicts that hyperbolicity would become the crucial concept in classifying the structure of 3-manifolds, rather than Hakenness.

As for the second question in the previous section, one should mention that a hyperbolic 3-manifold is easily shown to be homotopically atoroidal, based on the structure of the isometry group of 3-dimensional hyperbolic space. Thus its answer should be expected to involve hyperbolicity. Thurston proved the following remarkable theorem for Haken manifolds almost at the same time that he proved the Hyperbolic Dehn Surgery Theorem.

Theorem 2 (Hyperbolization Theorem for Haken Manifolds) *Any homotopically atoroidal Haken manifold admits a hyperbolic structure.*

This theorem answers the second question in the previous section for Haken manifolds, including the case without essential embedded tori, The result was strong enough to prove the 40 years standing Smith conjecture in the affirmative, with the help of minimal surface theory developed by Meeks–Yau and several additional arguments by Bass, Gordon–Litherland and Shalen. The Smith conjecture asks if any non-trivial orientation-preserving periodic diffeomorphism on the 3-sphere with nontrivial fixed point set is equivariant to a periodic isometry on the 3-sphere with the standard spherical metric. The solution of the Smith conjecture, which gathered many ideas of different aspects of mathematics such as differential geometry, subgroups of Lie groups, geometric topology and their interplay was presented in the late 1970s in the book [26] that collected articles by the main players for the resolution except for Thurston.

Thurston's proof of Theorem 2 was divided completely but not exclusively into two cases according to whether a surface along which we first cut is a fiber or not. The proof for the second case was based on the hierarchy of Haken manifolds and uses many new and exciting ideas to overcome serious difficulties that appeared in the geometric setting in contrast to Waldhausen's topological setting. The proof for the first case needs a double limiting argument involving Kleinian groups and was based on Thurston's deep study of automorphisms of surfaces in [42]. A fairly detailed outline of the proof was presented in the article by Morgan in [26]. Thurston promised to provide a series of six papers that would cover all details. However this series was never completed. The first paper [41], discussing the deformation space for the acylinderical case, appeared in 1986 in the Annals of Mathematics. But the second one [44] discussing the fibered case, in other words the I-bundle case after cutting along the first essential surface, and the third one [43] discussing the mixed case where the complement of a cutting surface has both acylinder parts and I-bundle parts are still in preprint forms. The remaining papers are unavailable in public. There were several attempts to cover all of the details. Among others, we recommend the articles by Otal [28, 29] and the book by Kapovich [19] for expositions that are close to Thurston's original ideas, and also the articles by Sullivan [38] and McMullen [22] for alternative proofs for the fibered case. Theorem 2 was called the Monster in those days, because of the extreme difficulty of its proof.

In the very late 1970s, Thurston seemed to be convinced that hyperbolic geometry would cover the missing part of the Waldhausen, Jaco–Shalen and Johannson theory. He formulated his vision of the structure theory of 3-manifolds in terms of geometric concepts and announced it as the geometrization conjecture at the AMS symposium on the mathematical heritage of Henri Poincaré held in 1980 at Indiana. His article based on the lecture was published as [40].

To state the geometrization conjecture precisely, we need to give a few definitions. Recall that we are assuming all manifolds to be compact and orientable. A 3-manifold is known to decompose into a finitely many irreducible or $\mathbb{S}^1 \times \mathbb{S}^2$ factors by splitting along separating essential spheres by Kneser [21]. This is called a connected sum decomposition. The uniqueness of this decomposition was proved by Milnor [24] in the early 1960s. As we have already mentioned in the previous section, Jaco–Shalen and Johannson then developed the torus decomposition theory. These topological decompositions provided the basic pieces that, Thurston conjectured, would admit one of eight geometric structures. Six of them are for Seifert fibered spaces and another one is for solvable manifolds. The last one is the hyperbolic geometry that was expected to be the generic case. A more precise and detailed description of the eight 3-dimensional geometries can be found in the article by Scott [37].

The following geometrization conjecture, which includes the Poincaré conjecture as a special case, lasted for 20 years as an undoubted working hypothesis for people involved in 3-dimensional topology.

Conjecture 3 (Geometrization Conjecture) Any closed 3-manifold admits a canonical decomposition into geometric pieces.

In [40], Thurston also listed a set of problems. Among them, one group on 3-manifold theory consists of variants of the geometrization conjecture. Problems in another group on 3-manifold theory are mostly related to the conjecture usually attributed to Waldhausen in [45], though he did not formally state it, asserting that every irreducible 3-manifold with infinite π_1 is virtually Haken, that is, finitely covered by a Haken manifold. This is obviously a question about non-Haken manifolds, and one should remark that non-Haken Seifert manifolds with infinite π_1 were known to admit a Haken finite cover. Also, it was proved in the early 1990s by Casson–Jungries [6] and independently by Gabai [10] based on several previous works that an irreducible 3-manifold with a normal $\mathbb{Z}$ in its π_1 and hence a $\mathbb{Z} \times \mathbb{Z}$ is a Seifert fibered space. Thus the real target of Waldhausen's conjecture became homotopically atoroidal manifolds, for instance hyperbolic manifolds. In anticipation of such progress, Thurston conjectured more wildly,

Conjecture 4 (Virtual Fiber Conjecture) Any hyperbolic 3-manifold is finitely covered by a manifold which fibers over the circle.

Since a hyperbolic 3-manifold fibering over the circle certainly contains an essential embedded surface of negative Euler characteristic as a fiber, it is Haken. Thus the positive resolution of the virtual fiber conjecture implies the resolution for

Waldhausen's conjecture mentioned in the previous paragraph if the geometrization conjecture is true.

These two challenging conjectures of Thurston have greatly encouraged the development of the theory of 3-manifolds, and both were finally resolved early in this century. But before going into their details, let us come back to another of Thurston's contribution to the geometrization conjecture.

Thurston announced the following orbifold version of the geometrization conjecture in 1982.

Theorem 5 (Orbifold Theorem) *Any closed 3-manifold admitting an orientation preserving periodic self-diffeomorphism with a nontrivial fixed point set admits a geometric decomposition such that the diffeomorphism is isometric.*

The orbifold theorem announced by Thurston was planned to be proved by giving a geometric structure on the quotient orbifold. It has many consequences. For example, it provides a new proof of the Smith conjecture. It implies the geometrization conjecture for closed 3-manifold with Heegaard decomposition of genus 2 since these manifolds admit orientation preserving involutions coming from the hyper-elliptic involution on the Heegaard surface.

The proof starts with looking at the quotient orbifold. The complement of the singular set is Haken and hence, by the Monster, admits a geometric decomposition The main case is when this complement admits a complete hyperbolic structure. One can use the hyperbolic Dehn surgery deformation theory from Theorem 1 so that the structure is conically singular along the singular set without changing the underlying space. If the cone angle increasing the Dehn surgery deformation reaches the expected cone angle, then we are done. However, the argument needs much more care when the deformation cannot reach the expected angle. Thurston had analyzed the possible manners of degeneration and claimed that the conclusion holds with suitable care after degeneration. Unfortunately, people needed more time to be convinced of what Thurston claimed compared to the case of the Monster. More than 15 years later, two groups succeeded in giving complete proofs for the orbifold theorem in [5, 8]. At the very end of Thurston's argument, he used an equivariant version of Hamilton's work [11] on Ricci flow of 3-manifolds admitting positive Ricci curvature. The role of Hamilton's work is to deform metrics on 3-orbifolds with positive Ricci curvature to positive constant sectional curvatures. We will come back to the contribution of Hamilton to the geometrization conjecture in the next section.

Thurston was awarded the Fields medal for his contribution to the study of 3-manifolds in 1982. C. T. C. Wall, who presented his work, said that 3-dimensional topology had now firmly rejoined the mainstream of mathematics by the work of Thurston. These are undoubtedly the best words to express Thurston's contribution in 3-dimensional topology.

4.4 Post-Thurston Era

In this century, there were two miracles that independently resolved these extremely wild conjectures of 1982. We will briefly describe them in this section.

4.4.1 Geometrization Conjecture

There had been many attempts to partially answer or approach the geometrization conjecture for the case of infinite π_1 since it was formulated. However the case of finite π_1 had stayed completely outside of the development in the last century except for the orbifold theorem and Hamilton–Perelman's attempts.

Perelman put three papers [30–32] on arXiv in 2002 and 2003 and declared the resolution of the geometrization conjecture in the affirmative. The proof was based on Hamilton's Ricci flow, which is a flow on the space of Riemannian metrics on a manifold satisfying an evolutional PDE,

$$\frac{d}{dt}\, g = -2\,\mathrm{Ric}_g$$

where Ric represents the Ricci curvature tensor. This equation can be regarded as a variational equation that averages metrics in the long term. After Hamilton proved a sort of geometrization for 3-manifolds with positive Ricci curvature in [11], Yau suggested to Hamilton that the solution of the Ricci flow might converge to the metric that the geometrization conjecture expects. Hamilton had been seriously involved in Yau's suggestion and has been establishing a program to resolve the geometrization conjecture through several deep studies on the Ricci flow in a series of papers [12–14]. A survey by Chow [7] discusses Hamilton's work along these lines.

Perelman essentially followed Hamilton's program. However, he had to fill in conceptual and technical details and needed to modify Hamilton's set up. His breakthrough would be to establish a locally non-collapsing theorem which guarantees that the rescaling limit of the solution of the Ricci flow is controllable. This is presented in the first preprint [30]. Then following [13], Perelman introduced Ricci flow with surgery in [31] and claimed the resolution of the geometrization conjecture.

Theorem 6 (Perelman) *The geometrization conjecture is true.*

This theorem was officially recognized at the ICM 2006 in Madrid. Excellent expositions of this innovation were presented by Milnor [25] in its very early stage and by McMullen [23] in its mature stage. The third preprint [32] presents a bypath argument independent from a half of the argument in [31] to prove for instance the Poincaré conjecture, when we start with a connected sum of manifolds with finite π_1.

Perelman's proof of the geometrization conjecture has a similarity with the proof of the orbifold theorem. Both arguments are based on the deformation of metrics. The difference is that Thurston used hyperbolic metrics with cone singularities, while Perelman used nonsingular Riemaniann metrics. Thus Thurston's argument is geometric, while Perelman's involves more analysis. Perelman's arguments brought a big surprise not only to geometers but also to analysts world-wide. His arguments certainly led to a new direction for the study of singular solutions for some PDE. In fact, Perelman's work has promoted the popularity of the term "blow up" in contemporary PDE theory.

4.4.2 Virtual Fiber Conjecture

Once again, the virtual Haken conjecture by Waldhausen states that every irreducible 3-manifold with infinite π_1 is virtually Haken, namely, finitely covered by a Haken manifold. After Perelman's resolution of the geometrization conjecture, this conjecture remained unsolved only for hyperbolic manifolds. Thus, the virtual fiber conjecture came to include the virtual Haken conjecture, and this motivated a significant number of people. There were actually several serious studies by experts including Thurston, but most of them were rather illustrative or experimental before Agol [1]. In this paper, Agol presented a fairly generic criterion for virtual fibering that he called residually finite rational solvability. It is an effective refinement of separability of a subgroup in a group. Agol showed, in particular, that any subgroup of a right-angled Artin group has this useful property.

There was a strong reason for Agol to reach to this definition. Early this century, Wise and his many collaborators developed the theory of special groups. They started with groups acting effectively on non-positively curved (NPC for short) cube complexes. One of their most notable contributions was the discovery of the specialness of a group or an action defined in [15]. Specialness can be defined by a somewhat technical combinatorial structure of an NPC cube complex. Surprisingly, Haglund–Wise [15] showed that specialness can be characterized in a purely group theoretic fashion: that a special group is a subgroup of a right angled Artin group.

An epoch making result was provided by Kahn–Markovic [18] who showed that there are many almost geodesic quasi-Fuchsian surfaces in any closed hyperbolic 3-manifold. Their result, which was presented at the 21st Nevanlinna Colloquium held at Kyoto University in 2009, gave a completely new tool to discuss the virtual fiber conjecture. In fact, Bergeron–Wise [3] immediately noticed that the result could be used to build a wall system of Sageev, as in [36], in the 3-dimensional hyperbolic space invariant under the action of a cocompact Kleinian group. Then the recipe by Sageev [35, 36] leads to a construction of a proper group action on an NPC cube complex by the same group.

Agol then succeeded in showing that the action constructed in this way is special, and concluded that

Theorem 7 (Agol) *The virtual fiber conjecture is true, and so is the virtual Haken conjecture.*

What Agol actually proved in [2] was much stronger. He showed that every compact cube complex with a hyperbolic fundamental group is virtually special. In other words, a hyperbolic group that acts on an NPC cube complex cocompactly is virtually special. The resolution of the virtual fiber conjecture is one of its many consequences when combined with previous results by Sageev, Wise and his collaborators, Kahn–Markovic and Bergeron–Wise. Thus the theorem by Agol pertains not only to the theory of 3-manifolds. However, we would like to note that the idea of the proof was connected to the techniques developed in the theory of 3-manifold, in the Thurston Era, especially hyperbolic Dehn surgery theory in the context of group theory.

We would like to refer for more details of this innovative result to the articles by Klarreich [20] for general audiences, and Bestvina [4] and Friedle [9] for experts.

4.5 Epilogue

Thurston's vision of 3-manifold theory, as presented for instance in the problem set in [40], has been completed by Perelman and Agol. However, it is more important to mention that Thurston presented us simultaneously with many new directions in mathematics even within the theory of 3-manifolds. In fact, we found fine geometric nature of 3-manifolds and many interesting questions. The connection between quantum topology and hyperbolic geometry in 3-manifolds could be one such fascinating topics. For example, there is the volume conjecture, which expects to precisely identify the hyperbolic volume of a knot complement with a certain asymptote of special values of quantum invariants of the knot. This conjecture has motivated and involved many researchers including people in mathematical physics. There are many other germs to create new links from 3-manifold theory within topology, with the other mathematical fields including geometry and science in general. For instance, geometric group theory is now expanding in many new directions, Teichmüller theory is finding interest in a higher dimensional analogue, and so on.

What Thurston did and left to us was extremely incredible and more than dramatic indeed.

References

1. I. Agol, Criteria for virtual fibering. J. Topol. **1**, 269–284 (2008)
2. I. Agol, The virtual Haken conjecture, with appendix by I. Agol, D. Groves and J. Manning. Doc. Math. **18**, 1045–1087 (2013)
3. N. Bergeron, D. Wise, A boundary criterion for cubulation. Am. J. Math. **134**, 843–859 (2012)
4. M. Bestvina, Geometric group theory and 3-manifolds hand in hand: the fulfillment of Thurston's vision. Bull. Am. Math. Soc. **51**, 53–70 (2014)
5. M. Boileau, J. Porti, Geometrization of 3-dimensional orbifolds. Ann. Math. **162**, 195–290 (2005)
6. A. Casson, D. Jungreis, Convergence groups and Seifert fibered 3-manifolds. Invent. Math. **118**, 441–456 (1994)
7. B. Chow, A survey on Hamilton's program for the Ricci flow on 3-manifolds. AMS Contemp. Math. **367**, 63–78 (2005)
8. D. Cooper, C. Hodgson, S. Kerckhoff, The orbifold theorem. MSJ Mem. **5** (2000)
9. S. Friedle, Thurston's vision and the virtual fibering theorem for 3-manifolds, Jahresber. Deutsch. Math.-Verein. **116**, 223–241 (2014)
10. D. Gabai, Convergence groups are Fuchsian groups. Ann. Math. **136**, 447–510 (1992)
11. R. Hamilton, Three-manifolds with positive Ricci curvature. J. Differ. Geom. **17**, 255–306 (1982)
12. R. Hamilton, The formation of singularities in the Ricci flow. Surv. Differ. Geom. **2**, 7–136 (1995)
13. R. Hamilton, Four-manifolds with positive isotoropic curvature. Comm. Anal. Geom. **5**, 1–92 (1997)
14. R. Hamilton, Non-singular solution of the Ricci flow on 3-manifolds. Comm. Anal. Geom. **7**, 695–729 (1999)
15. F. Haglund, D. Wise, Special cube complexes. Geom. Funct. Anal. **17**, 1551–1620 (2008)
16. W. Jaco ad P. Shalen, Seifert fibered spaces in 3-manifolds. Mem. Amer. Math. Soc. **21**(2), 220 (1979)
17. K. Johannson, *Homotopy Equivalence of 3-manifolds with Boundaries*. Lecture Notes in Mathematics, no. 7761 (Springer, Berlin, 1979)
18. J. Kahn, V. Markovic, Immersing almost geodesic surfaces in a closed hyperbolic 3-manifold, Ann. Math. **175**, 1127–1190 (2012)
19. M. Kapovich, Hyperbolic manifolds and discrete groups (Modern Birkhäuser Classics, Birkhäuser, 2009)
20. E. Klarreich, Getting into shapes: from hyperbolic geometry to cube complexes and back. Quanta Magazine. Sci. Am. (2012). https://scientificamerican.com/article/getting-into-shapes/
21. H. Kneser, Geschlossen Flächen in dreidimensionalen Mannigfaltigkeiten. Jber. Deutsch. Math. Verein. **38**, 248–260 (1929)
22. C. McMullen, Renormalization and 3-manifolds which fiber over the circle. Ann. Math. Study **142** (1996)
23. C. McMullen, The evolution of geometric structures on 3-manifolds, Bull. Amer. Math. Soc. **48**, 259–274 (2011)
24. J. Milnor, A unique decomposition theorem for 3-manifolds, Amer. J. Math. **84**, 1–7 (1962)
25. J. Milnor, Towards the Poincaré conjecture and the classification of 3-manifolds. Notices Amer. Math. Soc. **50**, 1226–1233 (2003)
26. J. Morgan, H. Bass, *The Smith Conjecture*. Pure and Applied Mathematics, vol. 112 (Academic Press, New York, 1984)
27. G. Mostow, Quasi-conformal mappings in n-space and the rigidity of the hyperbolic space forms. Publ. IHES **34**, 53–104 (1968)
28. J.P. Otal, Le théorème d'hyperbolisation pour les variétés fibrées de dimension 3. Astérisque **235** (1986)

29. J.P. Otal, Thurston's hyperbolization of Haken manifolds, in *Surveys in Differential Geometry*, vol. III, ed. by C.C. Hsiung, S.T. Yau (International Press, Boston, 1998), pp. 77–194
30. G. Perelman, The entropy formula for the Ricci flow and its geometric applications (2002). arxiv:math/0211159
31. G. Perelman, Ricci flow with surgery on three-manifolds (2003). arXiv:math/0303109
32. G. Perelman, Finite extinction time for the solutions to the Ricci flow on certain three-manifolds (2003). arXiv:math/0309021
33. H. Poincaré, Cinquième élément à l'analysis situs. Rend. Circa. Mat. Palema **18**, 45–110 (1904)
34. G. Prasad, Strong rigidity of $\mathbb{Q}$-rank 1 lattices. Inventiones Math. **21**, 255–286 (1973)
35. M. Sageev, Codimension-1 subgroups and splittings of groups. J. Algebra **189**, 377–389 (1997)
36. M. Sageev, CAT(0) Cube Complexes and Groups. PCMI Lecture Notes **21**, 7–54 (2014)
37. P. Scott, The geometry of 3-manifolds. Bull. London Math. Soc. **15**, 401–487 (1983)
38. D. Sullivan, Travaux Thurston sur les groups quasi-Fuchsiens et les variétés hyperboliques de dimension 3 firbrees sur S^1. Seminaire Bourbaki, 1979/80 no. 554
39. W. Thurston, *The Geometry and Topology of* 3-*manifolds*. Lecture Notes (Princeton University, Princeton, 1979). Available at http://library.msri.org/books/gt3m/
40. W. Thurston, Three dimensional manifolds, Kleinian groups and hyperbolic geometry. Bull. Amer. Math. Soc. **6**, 357–382 (1982)
41. W. Thurston, Hyperbolic structures on 3-manifolds, I: Deformations of acylinderical manifolds. Ann. Math. **124**, 203–246 (1986)
42. W. Thurston, On the geometry and dynamics of diffeomorphisms of surfaces. Bull. Amer. Math. Soc. **19**, 417–431 (1988)
43. W. Thurston, Hyperbolic structures on 3-manifolds III: Deformations of 3-manifolds with incompressible boundary (1998). arXiv:math/9801058
44. W. Thurston, Hyperbolic structures on 3-manifolds II: Surface groups and 3-manifolds which fiber over the circle (1999). arXiv:math/9801049
45. F. Waldhausen, On irreducible 3-manifolds which are sufficiently large. Ann. Math. **87**, 56–88 (1968)

Chapter 5
Combinatorics Encoding Geometry: The Legacy of Bill Thurston in the Story of One Theorem

Philip L. Bowers

Abstract This chapter presents a whirlwind tour of some results surrounding the *Koebe–Andre'ev–Thurston Theorem*, Bill Thurston's seminal circle packing theorem that appears in Chapter 13 of *The Geometry and Topology of Three-Manifolds*.

Keywords Circle packing · Uniformization · Circle domain · Discrete conformal geometry · Conformal type · Inversive distance · Polyhedron

AMS Codes 52C26, 57M20

5.1 Introduction

Bill Thurston was the most original and influential topologist of the last half-century. His impact on the discipline of geometric topology during that time is unsurpassed, and his insights in the topology and geometry of three-manifolds led finally to the resolution of the most celebrated problem of topology over the last century—The Poincaré Conjecture. He made fundamental contributions to many sub-disciplines within geometric topology, from the theory of foliations on manifolds to the combinatorial structure of rational maps on the two-sphere, and from geometric and automatic group theory to classical polyhedral geometry. Of course his foundational work on three-manifolds, first laid out in his courses at Princeton in the late 1970s, compiled initially as a Princeton paper-back monograph inscribed by Bill Floyd and available upon request as *The Geometry and Topology of Three-Manifolds* (GTTM) [73], and maturing as the famous *Thurston Geometrization Conjecture* of the early 1980s, was the driving force behind the development of geometric topology for the next thirty years. The final confirmation of the Geometrization Conjecture

P. L. Bowers (✉)
Department of Mathematics, Florida State University, Tallahassee, FL, USA
e-mail: bowers@math.fsu.edu

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_5

by Giorgi Perelman using the flow of Ricci curvature, following a program that had been introduced by Richard Hamilton, is one of the crown jewels of twentieth century mathematics.

Thurston marks a watershed in the short history of topology,[1] a signpost, demarcating topology before Thurston, and topology after Thurston. This is evidenced not only in the fabulous results he proved, explained, and inspired, but even more so in how he taught us to do mathematics. Topology before Thurston was dominated by the general and the abstract, entrapped in the rarified heights that captured the mathematical world in general, and topology in particular, in the period from the 1930s until the 1970s. Topology after Thurston is dominated by the particular and the geometric, a throwback to the nineteenth-century, having much in common with the highly geometric landscape that inspired Felix Klein and Max Dehn, who walked around and within Riemann surfaces, knew them intimately, and understood them in their particularity. Thurston's vision gave a generation of topologists permission to get their collective hands dirty by examining in great depth specific structures on specific examples.

One of the organizing principles that lies behind Thurston's vision is that geometry informs topology, and that the non-Euclidean geometry of Lobachevski, Bolyai, and Beltrami in particular is systemic to the study of topology. Hyperbolic geometry permeates topology after Thurston, and it is hyperbolic geometry that becomes the common thread of the present chapter. This will be seen in the interrelated studies presented here. All to varying degrees are due to the direct influence of Bill Thurston and his generalization of the earlier results of Koebe and Andre'ev. All involve hyperbolic geometry in some form or influence, and even further all illustrate how combinatorics encodes geometry, another of the principles that underlies Thurston's vision. To my mind, the proposition that *combinatorics encodes geometry, which in turn informs topology* has become a fundamental guiding motif for topology after Thurston. I offer this chapter as a celebration of Bill Thurston's vision and his immense influence over our discipline.

5.1.1 An Introductory Overview

The Koebe–Andre'ev–Thurston Theorem represents a rediscovery and broad generalization of a curiosity of Paul Koebe's from 1936, and has an interpretation that recovers a characterization of certain three-dimensional hyperbolic polyhedra due to E.M. Andre'ev in two papers from 1970. This theorem is the foundation stone of the discipline that has been dubbed as *discrete conformal geometry*, which itself has been developed extensively by many mathematicians in many different directions over the last thirty years. Discrete conformal geometry in its purest form is geometry born of combinatorics, but it has theoretical and practical applications.

[1] I will use the term *topology* henceforth to mean *geometric topology*. By dropping the adjective *geometric* I certainly mean no slight of general, set-theoretic, or algebraic topology.

In the theoretical realm, it produces a discrete analytic function theory that is faithful to its continuous cousin, a quantum theory of complex analysis from which the classical theory emerges in the limit of large scales. In the realm of applications, it has been developed in a variety of directions, for practical applications in areas as diverse as biomedical imaging and 3D print head guidance. This rather large body of work flows from simple insights that Thurston presented in his lecture at Purdue University in 1985 on how to use the most elementary case of his circle packing theorem to provide a practical algorithm for approximating the Riemann mapping from a proper, simply-connected planar domain to the unit disk. A personal accounting of this development can be found in the author's own review [14] of the bible of circle packing theory, Ken Stephenson's *Introduction to Circle Packing: The Theory of Discrete Analytic Functions* [71].

A perusal of the section and subsection headings of this chapter will give the reader a clue as to where I am going in this survey. I primarily stick with the theoretical results for which there are fairly direct lines from the Koebe–Andre'ev–Thurston Theorem to those results. This means in particular that I almost totally ignore the really vast array of practical applications that circle packing has found, especially in the last two decades as discrete differential geometry has become of primary importance in so many applications among computer scientists and computational geometers. A survey of applications will have to wait as space constraints preclude a discussion that does justice to the topic.

5.1.2 Dedication and Appreciation

This chapter is dedicated to the memory of Bill Thurston and his student Oded Schramm, and to an appreciation of Jim Cannon and Ken Stephenson. I have spoken already of Bill Thurston's legacy. Oded Schramm was one of the first to press Thurston's ideas on circle packings to a high level of development and application, and his great originality in approaching these problems has bequeathed to us a treasure trove of beautiful gems of mathematics. Most of Oded's work on circle packing and discrete geometry was accomplished in the decade of the 1990s. As Bill is a demarcation point in the history of topology, Oded is one in the history of probability theory. In the late nineties, Oded became interested in some classical open problems in probability theory generated by physicists, in percolation theory and in random planar triangulations in particular. Physicists had much theoretical and computational evidence for the veracity of their conjectures, but little mathematical proof, or even mathematical tools to approach their verifications. In Oded's hands these venerable conjectures and problems began to yield to mathematical proof, using ingenious tools developed or refined by Oded and his collaborators, chief among which are SLE_κ, originally Stochastic-Loewner Evolution, now renamed as Schramm–Loewner Evolution, and UIPT's, or Uniform Infinite Planar Triangulations. For a wonderful biographical commentary on Oded's contributions to mathematics, see Steffen Rohde's article *Oded Schramm: From Circle Packing to SLE* in [58].

The two individuals who have had the greatest impact on my mathematical work are Jim Cannon and Ken Stephenson, the one a mathematical hero of mine, the other my stalwart collaborator for three decades. Jim's work has influenced mine significantly, and I greatly admire his mathematical tastes and contributions. Pre-Thurston, Jim had made a name for himself in geometric topology in the flavor of Bing and Milnor, having solved the famous double suspension problem and having made seminal contributions to cell-like decomposition theory and the characterization of manifolds. In the beginning of the Thurston era, his influential paper *The combinatorial structure of cocompact discrete hyperbolic groups* [24] anticipated many of the later developments of geometric group theory, presaging Gromov's thin triangle condition and, à la Dehn, the importance of negative curvature in solving the classical word and conjugacy problems of combinatorial group theory. He with Thurston invented automatic group theory and then Jim settled upon the conjecture that bears his name as the work that for three decades has consumed his attention. Ken has been a joy with whom to collaborate over the past three decades. He was inspired upon attending Thurston's Purdue lecture in 1985 to change his mathematical attentions from a successful career as a complex function theorist, to a geometer exploring this new idea of circle packing using both traditional mathematical proof and the power of computations for mathematical experimentation. I began my foray into Thurston-style geometry and topology by answering in [17] a question of Ken and Alan Beardon from one of the first research papers [7] to appear on circle packings after Rodin and Sullivan's 1987 paper [62] confirming the conjecture of Thurston from the Purdue lecture. Ken and I are co-authors on a number of research articles and his down-to-earth approach to the understanding of mathematics has been a constant check on my tendency toward flights of fancy. I have learned from him how to tell a good story of a mathematical topic. For Ken's warm friendship and collaboration I am grateful.

5.2 The Koebe–Andre'ev–Thurston Theorem, Part I

5.2.1 Koebe Uniformization and Circle Packing

In the early years of the twentieth century, rigorous proofs of the Riemann Mapping Theorem and the more general Uniformization Theorem were given by such eminent mathematicians as Osgood, Carathéodory, Poincaré, and Koebe, and refinements and re-workings would continue to be made by others, even up to the present.[2] The generalization of the Riemann Mapping Theorem to multiply-

[2]The author recommends rather highly the article *On the history of the Riemann mapping theorem* [38] by Jeremy Gray and the monograph *Uniformization of Riemann Surfaces: Revisiting a Hundred-Year-Old Theorem* [32]. These two works give insightful historical accountings of the discovery, articulation, understanding, and finally rigorous proofs of the Riemann Mapping

connected domains fell to the hands of Paul Koebe, who in 1920 in [49] proved that every finitely-connected domain in the Riemann sphere is conformally equivalent to a *circle domain*, a connected open set all of whose complementary components are points or closed round disks. Of course for a 1-connected, or simply-connected, domain this is nothing more than the Riemann Mapping Theorem. He proved also a rigidity result, that any conformal homeomorphism between any two circle domains with finitely many complementary components is in fact the restriction of a Möbius transformation.[3] Koebe's real goal was what is known by its German name as *Koebe's Kreisnormierungsproblem* and by its English equivalent as *Koebe's Uniformization Conjecture*, which he posed in 1908.

Koebe Uniformization Conjecture ([48]) *Every domain in the Riemann sphere is conformally homeomorphic to a circle domain.*

This of course includes those domains with infinitely many, whether countably or uncountably many, complementary components. The general Koebe Uniformization Conjecture remains open to this day. More on this later.

In a paper of 1936, Koebe obtained the following circle packing theorem as a limiting case of his uniformization theorem of 1920. This went unnoticed by the circle packing community until sometime in the early 1990s.

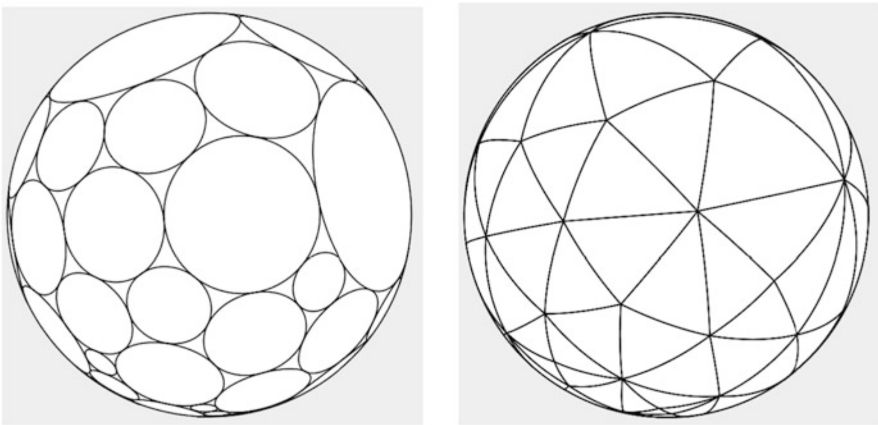
Koebe Circle Packing Theorem ([50]) *Every oriented simplicial triangulation K of the 2-sphere $\mathbb{S}^2$ determines a univalent circle packing $K(\mathcal{C})$ for K, unique up to Möbius transformations of the sphere.*

Here the *circle packing* $K(\mathcal{C})$ is a collection $\mathcal{C} = \{C_v : v \in V(K)\}$ of circles C_v in the sphere $\mathbb{S}^2$ indexed by the vertex set $V(K)$ of K such that C_u and C_v are tangent whenever uv is an edge of K, and for which circles C_u, C_v, and C_w bound a positively oriented interstice whenever uvw is a positively oriented face of K. The circle packing is *univalent* if there is a collection $\mathcal{D} = \{D_v : v \in V(K)\}$ of disks with $C_v = \partial D_v$ whose interiors are pairwise disjoint.[4] Connecting the centers of the adjacent circles by appropriate great circular arcs then produces a geodesic triangulation of $\mathbb{S}^2$ isomorphic to K. Figure 5.1 shows a circle packing of the sphere determined by an abstract triangulation K, and the realization of K as a geodesic triangulation. Of course the circle packings for a fixed K are Möbius equivalent, while the corresponding geodesic triangulations are not, simply because neither circle centers nor great circles are Möbius-invariant. I will look at a proof of the Koebe Circle Packing Theorem later, but first I'll present Thurston's generalization.

Theorem and the Uniformization Theorem. The narratives are at once engaging and perceptive, illustrating wonderfully the fact that mathematics is generally a common endeavor of a community of folks rather than the singular achievement of an enlightened few.

[3]Beware! This is not true in general. Two domains with uncountably many complementary components may be conformally equivalent yet fail to be Möbius equivalent.

[4]Without univalence, packings with branching would be allowed, where the sequence of circles tangent to a single circle C may wrap around C multiple times before closing up. See Sect. 5.3.3.



(a) The circle packing determined by a triangulation K of $\mathbb{S}^2$.

(b) The corresponding geodesic triangulation of $\mathbb{S}^2$.

Fig. 5.1 An abstract triangulation K of the 2-sphere determines (**a**) a circle packing, which in turn determines a realization of K as (**b**) a geodesic triangulation of the 2-sphere.

5.2.2 Koebe–Andre'ev–Thurston, or KAT for Short

In his Princeton course of 1978–79, Thurston greatly generalized the Koebe Circle Packing Theorem, though at the time he was unaware of Koebe's result. He generalized in two ways, first by allowing adjacent circles to overlap and second by extending the theorem to arbitrary compact orientable surfaces. Thurston realized that his version of the theorem on the sphere $\mathbb{S}^2$ in fact encodes information about convex hyperbolic polyhedra, the connection of course through the fact that the sphere $\mathbb{S}^2$ serves as the space at infinity of the Beltrami–Klein and Poincaré ball versions of hyperbolic three-space $\mathbb{H}^3$ with circles on the sphere the ideal boundaries of hyperbolic planes in $\mathbb{H}^3$. These polyhedra had been characterized in two papers of Andre'ev from 1970, whose results can be interpreted in terms of the existence and uniqueness of the circle packings Thurston examined in his generalization of Koebe. Thurston's generalization to overlapping packings on the sphere is now known as the *Koebe–Andre'ev–Thurston Theorem*, honoring its three principle protagonists.

Koebe–Andre'ev–Thurston Theorem I (for the sphere) *Let K be an oriented simplicial triangulation of $\mathbb{S}^2$, different from the tetrahedral triangulation, and let $\Phi : E(K) \to [0, \pi/2]$ be a map assigning angle values to each edge of K. Assume that the following two conditions hold.*

(i) *If e_1, e_2, e_3 form a closed loop of edges from K with $\sum_{i=1}^{3} \Phi(e_i) \geq \pi$, then e_1, e_2, and e_3 form the boundary of a face of K.*

(ii) *If e_1, e_2, e_3, e_4 form a closed loop of edges from K with $\sum_{i=1}^{4} \Phi(e_i) = 2\pi$, then e_1, e_2, e_3, and e_4 form the boundary of the union of two adjacent faces of K.*

Then there is a realization of K as a geodesic triangulation of $\mathbb{S}^2$ and a family $\mathcal{C} = \{C_v : v \in V(K)\}$ of circles centered at the vertices of the triangulation so that the two circles C_v and C_w meet at angle $\Phi(e)$ whenever $e = vw$ is an edge of K. The circle packing $\mathcal{C}$ is unique up to Möbius transformations.

Now I want to point out that exactly what is called the *Koebe–Andre'ev–Thurston Theorem* is not at all settled. Some references use the term to mean the tangency case of the theorem ($\Phi \equiv 0$), which is nothing more than the Koebe Circle Packing Theorem, while others use the term to mean Thurston's full generalization of the theorem to arbitrary closed surfaces that is presented in Sect. 5.3.2. Exactly what Thurston proved in GTTM also often is misreported. In fact my introduction to this section is a bit of a misreporting, so let me take a little time to say exactly what Thurston does in Chapter 13 of GTTM.

In terms of circle packings on the 2-sphere, Thurston does not allow overlaps of adjacent circles, only tangencies. His version of the tangency case appears as Corollary 13.6.2 in Chapter 13 of GTTM, and appears as a corollary of Theorem 13.6.1, which he attributes to Andre'ev. This theorem concerns hyperbolic structures on orbifolds and, as it was Thurston who invented the notion of orbifold in his course at Princeton during 1976–77 as recorded in the footnote on page 13.5 of Chapter 13 of GTTM itself, this theorem is an interpretation of Andre'ev's in the context of orbifolds. Thurston does not give a proof of Theorem 13.6.1, but uses its result ensuring the existence of a hyperbolic structure on a suitable orbifold to prove Koebe's Theorem of 1936, Corollary 13.6.2. He does this by using the triangulation K to define an associated polyhedron P by cutting off vertices by planes that pass through midpoints of edges. He then uses the Andre'ev result to realize P as a right-angled ideal polyhedron in $\mathbb{H}^3$. The faces of this polyhedron then lie in planes whose ideal circular boundaries are the circles of the desired tangency packing complemented by the orthogonal circles through three mutually adjacent points of tangency. He then invokes Mostow rigidity for uniqueness.

It isn't until he presents Theorem 13.7.1 that Thurston allows for adjacent circles to overlap with angle between zero and $\pi/2$, and that only for surfaces other than the sphere, those surfaces with nonpositive Euler characteristic. Thurston proves this by assigning polyhedral metrics with curvature concentrated at the vertices $v_1, \ldots, v_n$ by assigning a radius r_i at vertex v_i. Defining the mapping $c : \mathbb{R}^n \to \mathbb{R}^n$ that measures the curvature via $c(r)_i = 2\pi -$ (the angle sum at vertex v_i), he then argues in nine pages that the origin $\mathbf{0}$ is in the image of c, which implies the desired result. It is the case that the version Thurston presents on the sphere, Corollary 13.6.2, is Koebe's result, and uses Andre'ev's ideas for the proof. It is only with this positive genus version, Theorem 13.7.1, that Thurston puts forth new geometric ideas, fertile enough to spawn an industry dedicated to understanding polyhedral metrics on surfaces and their induced circle packings.

Thurston's approach to circle packing is rather entwined with his overall concern, that of hyperbolic structures on three-dimensional manifolds and orbifolds. Since this work of the 1970s, Thurston's circle packing results have spawned a rather extensive theory that is more combinatorial and geometric, and related more to classical complex function theory and Riemann surfaces, and less to three-manifolds. It is related intimately to hyperbolic polyhedra and their generalizations, this the subject of Sect. 5.7, and has found several scientific applications. In the hands of folks like Ken Stephenson and his students and collaborators, it has spawned a discrete theory of complex analytic functions, laid out ever so elegantly in Stephenson's *Introduction to Circle Packing* [71]. It has yielded beautiful results on, for example, discrete minimal surfaces in the hands of the Berlin school of Bobenko, Hoffman, Springborn, Pinkall, and Polthier; see for example [11] and [56]. Though the theory now is rather mature, it continues to interact in new and interesting ways with new areas, for instance lying in the background in conformal tilings [21, 22], or in the foreground with its interaction with the classical rigidity theory of Euclidean frameworks [23]. There is an immense literature here, and so much of it owes a great debt of gratitude to the insights of Bill Thurston.

5.2.3 A Proof of the Koebe Circle Packing Theorem

Rather than proving the whole of KAT I, I will address the case where Φ is identically zero and prove Koebe's result. The proof presented here can be modified to give a complete proof of KAT I, which is done in [18] in proving a generalization.[5] Koebe's original proof of his namesake theorem uses a limiting process on circle domains and classical analytic arguments on convergence of analytic families of maps, very much in the flavor of what we now teach as classical techniques in our complex analysis courses. There are now many proofs of the Koebe Circle Packing Theorem. To name a few, besides Koebe's, there is Thurston's in GTTM already outlined above based on Andre'ev's results on hyperbolic polyhedra, Al Marden and Burt Rodin's using piecewise flat polyhedral metrics, Alan Beardon and Ken Stephenson's [7] that adapts the classical Perron method for constructing harmonic maps as an upper envelope of subharmonic maps, Colin de Verdière's [33] based on a variational principle, Igor Rivins's hidden in his paper [60] on Euclidean structures on triangulated surfaces, the author's [12] that turns the Beardon–Stephenson proof upside down to address packings on punctured surfaces, and Alexandre Bobenko and Boris Springborn's [10] that uses a minimal principle on integrable systems. Here I present a geometric and combinatorial proof where hyperbolic geometry is the crucial ingredient. The proof is a twist on the Perron method used by Beardon and Stephenson in [7] and is specialized from a more general version that applies to arbitrary surfaces of finite conformal type that appears in [12]. We will see that it has the advantage of generalizing in interesting ways.

Proof of the Koebe Circle Packing Theorem By removing one vertex v_0 from K and its adjacent edges and faces, one obtains a triangulation T of a closed disk.

[5] See Sect. 5.3.3.

Place a piecewise hyperbolic metric on T as follows. For any positive function $r : V(T) \to (0, \infty)$, let $|T(r)|$ be the metric space obtained by identifying the face $v_1 v_2 v_3$ of T with the hyperbolic triangle of side lengths $r(v_i) + r(v_j)$ for $i \neq j \in \{1, 2, 3\}$. This places a piecewise hyperbolic metric on T with cone-like singularities at the interior vertices. This structure often now is called a piecewise hyperbolic *polyhedral metric*, and the function r is called variously a *radius vector* or *label*. For any vertex v, one can measure the angle sum $\theta_r(v)$ of the angles at v in all the faces incident to v. I will say that r is a *superpacking label* for T if the angle sums of all interior vertices are at most 2π, and a *packing label*[6] if all are equal to 2π.

Now modify this a little by allowing r to take infinite values at the boundary vertices. This causes some ambiguity only if there is a separating edge in T that disconnects T when removed. This will be taken care of later, so for now assume no separating edge exists. The goal is to find a packing label $\mathfrak{r}$ with $\mathfrak{r}(w) = \infty$ whenever w is a boundary vertex. Assuming that such an $\mathfrak{r}$ exists, we may glue on hyperbolic half planes along the faces with two boundary vertices to give a complete hyperbolic metric on a topological disk, which must be isometric to the hyperbolic plane. This implies that the metric space $|T(\mathfrak{r})|$ is isometric to an ideal polygon in the hyperbolic plane whose sides are hyperbolic lines connecting adjacent ideal vertices that correspond to the boundary vertices of T. Now placing hyperbolic circles of radii $\mathfrak{r}(v)$ centered at interior vertices v and horocycles centered at ideal vertices determined by the boundary vertices gives a univalent circle packing of the hyperbolic plane realized as, say, the Poincaré disk, the unit disk $\mathbb{D}$ in the complex plane with Poincaré metric $ds = 2|dz|/(1-|z|^2)$. The boundary circles are horocycles in the hyperbolic metric on the disk and are therefore circles internally tangent to the unit circle. Stereographic projection to the sphere $\mathbb{S}^2$ and addition of the equator as the circle corresponding to the vertex v_0 removed initially produces a univalent circle packing of the sphere in the pattern of K as desired. Uniqueness follows from uniqueness of the packing label $\mathfrak{r}$ with infinite boundary values, which follows from the construction of $\mathfrak{r}$ explained next.

Define the function $\mathfrak{r}$ as

$$\mathfrak{r}(v) = \inf\{r(v) : r \in \mathfrak{R}\} \tag{5.2.1}$$

where

$$\mathfrak{R} = \{r : V(T) \to (0, \infty] : r \text{ is a superpacking label for } T \text{ with infinite boundary values}\}.$$

The claim is that this is the desired packing label. The first observation is that $\mathfrak{R} \neq \emptyset$ so that we are not taking the infimum of the empty set. This is because one may choose label values so large on the interior vertices that all of the faces become hyperbolic triangles whose interior angles are no more than $2\pi/d$, where d is the

[6]For emphasis one sometimes calls this a *hyperbolic packing label* to distinguish it from *flat* or *Euclidean packing labels* that also find their use in this discipline.

maximum degree of all the vertices of T. It follows that $\mathfrak{r}$ is a non-negative function with infinite boundary values. To verify that $\mathfrak{r}$ is a packing label, I show that

(i) $\mathfrak{r}$ cannot take a zero value on any interior vertex, which then implies that $\mathfrak{r} \in \mathfrak{R}$, and,
(ii) the angle sum at any interior vertex is 2π, meaning further that $\mathfrak{r}$ is a packing label.

We need two preliminary observations.

(iii) *Hyperbolic area is bounded away from zero.* The hyperbolic area of the singular hyperbolic surface $|T(r)|$ is $\geq \pi$ for all superpacking labels $r \in \mathfrak{R}$.
(iv) *Monotonicity of angles.* For a face $f = v_0v_1v_2$ of T, let $\alpha_r(i)$, for $i = 0, 1, 2$, be the angle that the label $r \in \mathfrak{R}$ gives to f at vertex v_i. Then $\alpha_r(0) \uparrow \pi$, $\alpha_r(1) \downarrow 0$, and $\alpha_r(2) \downarrow 0$ monotonically as $r(v_0) \downarrow 0$ when $r(v_1)$ and $r(v_2)$ are held fixed.

In calculating the hyperbolic area to confirm item (iii), let $V(T)$ and $F(T)$ be the respective vertex and face sets of T of respective cardinalities $\mathtt{V}$ and $\mathtt{F}$. The sum of the angles of a face when given its metric by r is denoted $\alpha_r(f)$ so that its hyperbolic area is $A_r(f) = \pi - \alpha_r(f)$. Finally, with $\mathtt{V}_{\text{int}}$ and $\mathtt{V}_{\text{bd}}$ denoting the numbers of interior and boundary vertices of T so that $\mathtt{V} = \mathtt{V}_{\text{int}} + \mathtt{V}_{\text{bd}}$, one has

$$\text{hyp-area}\,(|T(r)|) = \pi\,\mathtt{F} - \sum_{f \in F(T)} \alpha_r(f) = \pi\,\mathtt{F} - \sum_{v \in V(T)} \theta_r(v) \geq \pi(\mathtt{F} - 2\mathtt{V}_{\text{int}}), \tag{5.2.2}$$

since $\theta_r(v) \leq 2\pi$ at interior vertices and $\theta_r(v) = 0$ at boundary ones. An Euler characteristic exercise then shows that $\mathtt{F} - 2\mathtt{V}_{\text{int}} = \mathtt{V}_{\text{bd}} - 2 \geq 1$, the inequality holding since K is simplicial. It follows that every superpacking label with infinite boundary values produces a metric on T with hyperbolic area at least π. Item (iv) is almost obvious from drawing examples, but can be given a rigorous proof using the hyperbolic law of cosines from hyperbolic trigonometry.

I now address item (i). First the claim is that the label $\mathfrak{r}$ cannot be identically zero on the set of interior vertices. Indeed, if $\mathfrak{r}$ is identically zero, one may choose a sequence of superpacking labels r_i with infinite boundary values such that, for each interior vertex v, $r_i(v) \to 0$ as $i \to \infty$. This latter fact in turn follows from the observation that the minimum label $\min\{r_1, r_2\}$ is in $\mathfrak{R}$ whenever r_1 and r_2 are labels in $\mathfrak{R}$, which in turn is a consequence of the monotonicity of angles (iv). Recall that we are under the assumption that there are no separating edges so that at least one vertex of any face f of T is interior. Any such interior vertex has r_i-values converging to zero, and any boundary one is fixed at infinity, and with this it is easy to see that the hyperbolic area $A_{r_i}(f) \to 0$ as $i \to \infty$. But this implies that the hyperbolic area of $|T(r_i)|$ converges to zero as $i \to \infty$, which contradicts item (iii).

Now could it be that $\mathfrak{r}$ takes a zero value at some interior vertex, but not at all? The argument that this in fact does not happen is a generalization of what I have argued thus far. I will but give an indication of how it goes, referring the reader to [12] for details. Let T' be the subcomplex of T determined by those faces of T that have a vertex in $\mathfrak{r}^{-1}(0)$. An argument using Euler characteristic similar to that already given implies that the hyperbolic area of $|T'(r)|$ is positive and bounded away from zero for every superpacking label r with fixed non-negative boundary values. But an argument as in the preceding paragraph shows that the hyperbolic areas of $|T'(r_i)|$ converge to zero for a sequence of superpacking labels with fixed boundary values and interior vertex values converging to zero. This contradiction implies that $\mathfrak{r}$ is a positive function on the interior vertex set, and continuity of angles of a triangle with respect to edge lengths implies that $\theta_{\mathfrak{r}}(v) = \lim_{i\to\infty} \theta_{r_i}(v) \leq 2\pi$ at any interior vertex, since $\theta_{r_i}(v) \leq 2\pi$ for all i. This shows that $\mathfrak{r} \in \mathfrak{R}$ and completes the verification of item (i).

Item (ii) follows quickly from item (iv). Indeed, if (ii) fails, then there is an interior vertex v of T such that $\theta_{\mathfrak{r}}(v) < 2\pi$. By the monotonicity properties (iv), varying $\mathfrak{r}$ by slightly decreasing its value at v without changing any other values increases $\theta_{\mathfrak{r}}(v)$ while decreasing $\theta_{\mathfrak{r}}(w)$ for any vertex w incident to v. By making that decrease of $\mathfrak{r}(v)$ small enough to keep the angle sum at v below 2π, we obtain a superpacking label r with infinite boundary values that satisfies $r(v) < \mathfrak{r}(v)$, which contradicts the definition of $\mathfrak{r}$ in Eq. 5.2.1.

At this point I have shown that $\mathfrak{r}$ is a packing label with infinite boundary values, and I now claim that it is the only one. Suppose there is a packing label r in $\mathfrak{R}$ that differs from the infimum label $\mathfrak{r}$ defined in Eq. 5.2.1. Then $\mathfrak{r}(v) \leq r(v)$ for all vertices v, but there must be some interior vertex w with $\mathfrak{r}(w) < r(w)$. This implies that the hyperbolic area of the surface $|T(\mathfrak{r})|$ is strictly less than that of $|T(r)|$. But this is impossible since $\mathfrak{r}$ and r are packing labels with infinite boundary values, and as argued above, both $|T(\mathfrak{r})|$ and $|T(r)|$ are ideal hyperbolic polygons with $\mathrm{V_{Bd}}$ sides. An easy exercise shows that the hyperbolic area of any such hyperbolic polygon is equal to $(\mathrm{V_{Bd}} - 2)\pi$.

This completes the proof modulo the assumption that T has no separating edge. This is handled by induction on the number of such edges. If there is one separating edge uv, cut T into T_1 and T_2 along that edge, circle pack each in the unit disk with horocyclic boundary circles, and then using Möbius transformations, place the T_1 packing in the upper half disk with the horocycles for u and v circles tangent at the origin and centered on the real axis, and place the T_2 packing in the lower half of the disk with those same horocylic circles for u and v. This is possible since T is oriented, and this gives an appropriate packing label of T with infinite boundary values. □

5.2.4 *Maximal Packings and the Boundary Value Problem*

This proof actually proves the following extremely useful fact, which Beardon and Stephenson [7] exploited to give the first extension of the Koebe Circle Packing Theorem to infinite packings of the disk and the plane. The infinite theory is presented in Sect. 5.4.

Maximal Disk Packing Theorem *Every oriented simplicial triangulation T of a closed disk determines a univalent circle packing $T(\mathcal{C})$ for T in the unit disk $\mathbb{D}$ in the complex plane $\mathbb{C}$, unique up to Möbius transformations of the disk, with the circles corresponding to boundary vertices of T internally tangent to the unit circle boundary $\partial\mathbb{D} = \mathbb{S}^1$. Moreover, when given its canonical hyperbolic metric making $\mathbb{D}$ into the Poincaré disk model of the hyperbolic plane $\mathbb{H}^2$, the circle radii of the packing are uniquely determined by T.*

The circle packing guaranteed by this theorem is called the *maximal packing* for T. This theorem is in fact a special case of the more general result of Beardon and Stephenson [8] that solves the discrete version of the classical Dirichlet boundary value problem of harmonic analysis. In that paper, the authors also prove a discrete version of the classical Schwarz-Pick Lemma of complex analysis. These two theorems finish up the present section.

Discrete Boundary Value Theorem (Beardon and Stephenson [8]) *Let T be an oriented simplicial triangulation of a closed disk and $f : V_{\mathrm{Bd}}(T) \to (0, \infty]$ a function assigning positive or infinite values to the boundary vertices. Then there exists a unique hyperbolic packing label $\mathfrak{r} : V(T) \to (0, \infty]$ extending f. The resulting circle packing $T(\mathcal{C}_{\mathfrak{r}})$ of the unit disk $\mathbb{D}$ is unique up to Möbius transformations of $\mathbb{D}$.*

Proof The proof is a straightforward modification of that of the Koebe Circle Packing Theorem already presented. Again $\mathfrak{r} = \inf \mathfrak{R}$ is the desired packing label, provided that

$$\mathfrak{R} = \{r : r \text{ is a superpacking label for } T \text{ with } r(w) \geq f(w) \text{ when } w \in V_{\mathrm{Bd}}(T)\}.$$

Of course, $f \equiv \infty$ gives the maximal packing of the preceding theorem. □

This proof is a modification of the Beardon–Stephenson proof, which uses *subpacking* rather than superpacking labels. In a subpacking label, the interior angle sums are greater than or equal to 2π and one obtains the packing label as an upper envelope of subpacking labels, with the packing label given by $\mathfrak{r} = \sup \mathfrak{R}'$ where $\mathfrak{R}'$ is the set of subpacking labels with boundary values given by f. The advantage of approaching the desired packing label $\mathfrak{r}$ from above using superpackings ($\inf \mathfrak{R}$) rather than from below using subpackings ($\sup \mathfrak{R}'$) is that this *upper Perron method* readily generalizes to include cusp type singularities and cone type singularities at interior vertices.[7] This is presented in Sect. 5.3.4.

A word of warning here. When the boundary values are allowed to be finite, the resulting packing, though locally univalent, may not be globally univalent. This means that the disks bounded by the circles of the packing may overlap non-trivially, though ones neighboring the same interior vertex never do; this is the meaning of

[7] Another not insignificant advantage is that it is easy to show that $\mathfrak{R} \neq \emptyset$ while proving that $\mathfrak{R}' \neq \emptyset$ generally is difficult.

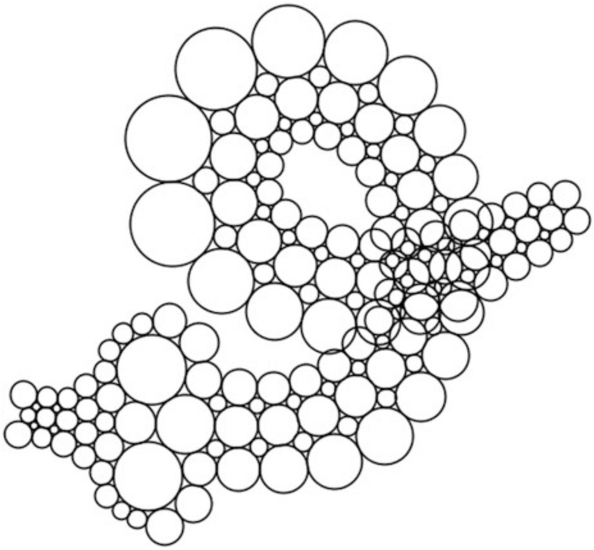


Fig. 5.2 A locally univalent circle packing that is not globally univalent

locally univalent. Figure 5.2 shows a locally univalent packing that is not globally univalent.

The second theorem of Beardon and Stephenson follows partly from the proof of the first (item (i)), and the rest of the theorem follows from a careful analysis of paths and angles in piecewise hyperbolic surfaces. The details of course appear in [8]. The theorem I state here is the generalization of what the reference [8] calls the Discrete Schwarz-Pick Lemma, which in turn is obtained by setting the boundary values of r identically to infinity.

Discrete Schwarz-Pick Lemma (Beardon and Stephenson [8]) *Let* $r, r'; V(T) \to (0, \infty]$ *be packing labels for the oriented simplicial triangulation* T *of a closed disk. Suppose that* $r(w) \geq r'(w)$ *at every boundary vertex* $w \in V_{\text{Bd}}(T)$*. Then*

(i) $r \geq r'$*; i.e.,* $r(v) \geq r'(v)$ *at every vertex* v *of* T*;*
(ii) $\rho_r(u, v) \geq \rho_{r'}(u, v)$ *for any two vertices* u *and* v*, where* ρ_r *is the distance function on the metric surface* $|T(r)|$*, and similarly for* $\rho_{r'}$*;*
(iii) $A_r(f) \geq A_{r'}(f)$ *for any face* f *of* T*.* (*Recall that* $A_r(f)$ *is the hyperbolic area of the face* f*.*)

Moreover, if a single instance of finite equality occurs at an interior vertex in (i)*, or at vertices* u *and* v *at least one of which is interior in* (ii)*, or at any face in* (iii)*, then* $r = r'$*.*

5.3 The Koebe–Andre'ev–Thurston Theorem, Part II

5.3.1 Circle Packings of Compact Surfaces

Thurston's other avenue of generalization of Koebe, indeed the more far-reaching one, is his extension of KAT to arbitrary orientable closed surfaces. Here there are striking illustrations of how purely combinatorial information encodes precise geometry. I will start with Thurston's tangency case of packings before presenting his version with overlaps.

Theorem 5.3.1 *Let K be an oriented simplicial triangulation of a closed surface S_g of positive genus. Then there is a metric of constant curvature, unique up to scaling, on S_g that supports a univalent, tangency circle packing $\mathcal{C} = \{C_v : v \in V(K)\}$ modeled on K. In particular, C_u and C_v are tangent whenever uv is an edge of K. The packing $\mathcal{C}$ is unique up to isometries of S_g in this metric when $g \geq 2$, and up to scaling when $g = 1$. Connecting the centers of adjacent circles by geodesic shortest paths produces a geodesic triangulation of the surface in the pattern of K. The metric is locally Euclidean when $g = 1$ and locally hyperbolic otherwise.*

Just in case the reader blinked and missed it, I aim to emphasize the extent to which combinatorics determines geometry in this theorem. The simplicial complex K provides purely combinatorial data with topological overtones. Yet hidden inside of the combinatorics is precise geometry. For example in the hyperbolic case where $g > 1$, among the uncountably many possible pairwise distinct hyperbolic metrics of constant curvature -1 as tabulated in the $(6g-6)$-dimensional moduli space $\mathcal{M}(S_g) \cong \mathbb{R}^{6g-6}$, the complex K chooses exactly one of these metrics, and in that metric, determines a univalent circle packing unique up to isometry! For none of the other metrics that S_g supports is there a univalent tangency packing of circles in the pattern of K! Since there are only countably many pairwise distinct simplicial triangulations of the fixed surface S_g, only countably many of the metrics parameterized by $\mathcal{M}(S_g)$ support any univalent tangency packing at all, though the set of metrics that do support such circle packings does form a dense subset of the moduli space.

I present a Proof of Theorem 5.3.1 based on the upper Perron method used to prove the Koebe Circle Packing Theorem.

Proof Let $\mathfrak{R} = \{r : V(K) \to (0, \infty) : \theta_r(v) \leq 2\pi$ for all $v \in V(K)\}$, the set of superpacking labels for K. Here again, exactly as in the proof of the Koebe Circle Packing Theorem, the label r determines a hyperbolic polyhedral metric surface $|K(r)|$. A unique packing label for which the angle sum at every vertex is equal to 2π would give all the claims of the theorem in the hyperbolic case. My claim is that when $g \geq 2$, the function $\mathfrak{r} = \inf \mathfrak{R}$ is the unique packing label for K, and when $g = 1$, then $\mathfrak{r} = \inf \mathfrak{R}$ is identically zero, but provides a way to place a flat polyhedral metric on K that meets the packing condition.

Exactly the calculation of Inequality 5.2.2 gives hyp-area($|K(r)|$) $\geq (\mathtt{F} - 2\mathtt{V})\pi$ for any superpacking label $r \in \mathfrak{R}$, and an Euler characteristic argument gives

$$\mathtt{F} - 2\mathtt{V} = -2\chi(S_g) = 4g - 4. \tag{5.3.1}$$

When $g \geq 2$ so that $\mathtt{F} - 2\mathtt{V}$ is positive and hence hyp-area($|K(r)|$) is positive, the same argument used in the proof of the Koebe Circle Packing Theorem shows that items (i) and (ii) of that proof hold, so that $\mathfrak{r}$ is a packing label. Uniqueness follows exactly as in that proof.

The remaining case is when $g = 1$ so that S_g is a topological torus. Here are the steps in proving that S_1 supports a flat metric that supports a univalent circle packing in the pattern of K, both the packing and the metric unique up to scaling.

(i) When $g = 1$, $\mathtt{F} - 2\mathtt{V} = 0$ and this implies that $\mathfrak{r} = \inf \mathfrak{R}$ is identically zero on $V(K)$.
(ii) Fix a vertex $v^\dagger$ in K and let $\mathfrak{R}^\dagger = \{r^\dagger : r \in \mathfrak{R}\}$, where $r^\dagger$ is the normalized label defined by $r^\dagger(v) = r(v)/r(v^\dagger)$.
(iii) Show that $\mathfrak{r}^\dagger = \inf \mathfrak{R}^\dagger$ takes only positive values.
(iv) Let $|K(\mathfrak{r}^\dagger)|_{\text{flat}}$ be the flat polyhedral surface with cone type singularities obtained by identifying a face $v_1v_2v_3$ with the Euclidean triangle of side-lengths $\mathfrak{r}^\dagger(v_i) + \mathfrak{r}^\dagger(v_j)$ for $i \neq j \in \{1, 2, 3\}$.
(v) Show that $|K(\mathfrak{r}^\dagger)|_{\text{flat}}$ is non-singular; i.e., $\mathfrak{r}^\dagger$ is a flat packing label with Euclidean angle sums $\theta^{\text{flat}}_{\mathfrak{r}^\dagger}(v) = 2\pi$ at every vertex v.
(vi) Show that $\mathfrak{r}^\dagger$ is the unique flat packing label with value 1 at $v^\dagger$.

The details of the argument appear in [7], but I will give an indication of why this outline works to prove the desired result. Let $A(r)$ be the hyperbolic area of the singular hyperbolic surface $|K(r)|$ when $r \in \mathfrak{R}$ and observe that

$$A(r) - s(r) = (\mathtt{F} - 2\mathtt{V})\pi, \quad \text{where} \quad s(r) = \sum_{v \in V(K)} (2\pi - \theta_r(v)). \tag{5.3.2}$$

Here $s(r)$ is the total *angle shortage*.[8] In the genus 1 case, $\mathtt{F} - 2\mathtt{V} = 0$ so $A(r) = s(r)$ for all superpacking labels $r \in \mathfrak{R}$. Now assuming that item (i) has been verified, any superpacking label r that is close to the infimum $\inf \mathfrak{R} = 0$ has area $A(r)$ close to zero and hence so too is the shortage $s(r)$ close to zero. In the limit as $r \to \inf \mathfrak{R} = 0$, the shortages $s(r) \to 0$ and this implies that the singular hyperbolic surfaces $|K(r)|$ have angle sums $\theta_r(v) \to 2\pi$ for every vertex v. Since Euclidean geometry is the small scale limit of hyperbolic geometry, this implies that the Euclidean angle sums $\theta^{\text{flat}}_r(v) \to 2\pi$ as $r \to 0$. Thus the collection $\{|K(r)|_{\text{flat}}\}_{r \in \mathfrak{R}}$ is a collection of singular flat surfaces whose singularities are removed in the limit as $r \to 0$. Of course there is no limiting surface since

[8] Also called the *discrete curvature*.

$r \to 0$. Whereas this cannot be remedied in hyperbolic geometry, it can be remedied in Euclidean geometry by rescaling the labels r as described in item (ii). With item (iii) confirmed so that the flat polyhedral surface $|K(\mathfrak{r}^\dagger)|_{\text{flat}}$ of item (iv) exists, since similarity transformations exist in Euclidean geometry, these rescalings preserve the Euclidean angles and imply that the limit surface $|K(\mathfrak{r}^\dagger)|_{\text{flat}}$ is non-singular. Items (v) and (vi) just state formally the result of making this imprecise but rather accurate discussion rigorous. □

5.3.2 KAT for Compact Surfaces

Thurston's Theorem 13.7.1 of GTTM combines the introduction of surfaces of genus greater than zero in Theorem 5.3.1 with the overlap conditions of the KAT Circle Packing Theorem.

Koebe–Andre'ev–Thurston Theorem II (for compact surfaces) (Theorem 13.7.1, GTTM) *Let K be an oriented simplicial triangulation of a surface S_g of genus $g \geq 1$, and let $\Phi : E(K) \to [0, \pi/2]$ be a map assigning angle values to each edge of K. Assume that the following two conditions hold.*

(i) *If e_1, e_2, e_3 form a closed loop of edges from K with $\sum_{i=1}^{3} \Phi(e_i) \geq \pi$, then e_1, e_2, and e_3 form the boundary of a face of K.*
(ii) *If e_1, e_2, e_3, e_4 form a closed loop of edges from K with $\sum_{i=1}^{4} \Phi(e_i) = 2\pi$, then e_1, e_2, e_3, and e_4 form the boundary of the union of two adjacent faces of K.*

Then there is a metric of constant curvature on S_g, unique up to scaling, and a realization of K as a geodesic triangulation in that metric, as well as a family $\mathcal{C} = \{C_v : v \in V(K)\}$ of circles centered at the vertices of the triangulation so that the two circles C_v and C_w meet at angle $\Phi(e)$ whenever $e = vw$ is an edge of K. The circle packing $\mathcal{C}$ is unique up to isometry.

I already have discussed the proof in GTTM. Let me say further that it was in this proof that Thurston introduced the idea of using labels, or radii assignments to vertices, to build a polyhedral surface with cone type singularities, and then to vary the labels until the packing condition is met. This is still the basic idea for proving many packing results, though the way in which one varies the labels and the choice of initial labels changes from researcher to researcher and from application to application. The Perron method used in this chapter is a modification of the method of Beardon and Stephenson [7]. This idea also led to a practical algorithm for producing the packing labels that was the starting point for Ken Stephenson's `CirclePack`. This sophisticated software package for computing circle packings has enjoyed extensive development over the past 30 years and is freely available at Ken's webpage.

Before I introduce infinite circle packings and their really interesting and novel features in Sect. 5.4, I'll discuss two generalizations of the KAT Theorems. The first is presented in Sect. 5.3.3 and generalizes KAT I to certain branched packings of the 2-sphere where circles tangent to a given one wrap around that one more than once. These packings of course fail to be univalent, but provide a rich family of packings that model the behavior of polynomial mappings of the Riemann sphere. The ultimate goal is to model arbitrary rational mappings of the sphere, which would require the theory to extend to more general branch structures, this a topic of current research; see for example [5]. The second is presented in Sect. 5.3.4 and examines how to include both cusps with ideal vertices as well as prescribed discrete curvature at pre-chosen vertices.

5.3.3 A Branched KAT Theorem and Polynomial Branching

Ken Stephenson and I generalized KAT I by allowing for *polynomial branching* to occur in the circle packing. *Branching* means that we allow for the angle sums at predetermined vertices to be a positive integer multiple of 2π rather than just 2π, or stated differently, we allow the circles tangent to a given one to wrap around that given circle multiple times before closing up; see Fig. 5.3. *Polynomial* means that half the branching is concentrated at one vertex. The terminology comes from the classical theory of rational maps. Indeed, rational mappings may be thought of as

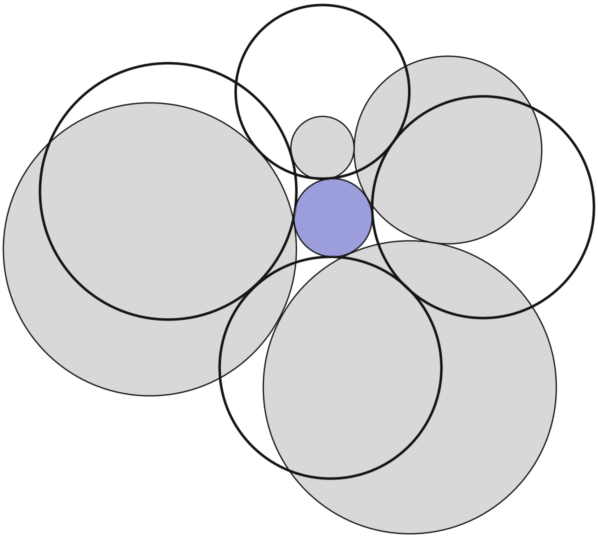


Fig. 5.3 Branching of multiplicity $m = 2$ or order $\mathfrak{o} = 1$. Starting with the grey disk on the left and moving counterclockwise, four sequentially tangent grey disks wrap around the blue central disk nearly one full turn, at which point the sequentially tangent transparent (or white) disks take over to wrap around slightly more than one full turn to close up the flower of circles with angle sum $\theta = 4\pi$

branched self-mappings of the 2-sphere, and the polynomial mappings are precisely those in which there is an even amount of branching with half the branching occurring at a single point. Taken together, a circle packing promised by the next theorem mimics the behavior of a polynomial mapping of the Riemann sphere.

Our proof of the theorem as presented in [18] offers an independent proof of KAT I, which the branched version reduces to when the *branch structure* β is empty. In fact as far as I know, it was the first full direct proof of KAT I given that Thurston proves only the tangency case (the Koebe Circle Packing Theorem) and Marden-Rodin [54], though allowing overlapping circles, has more restrictive hypotheses. KAT I is implied by Igor Rivin's earlier work, which bears the same resemblance to KAT I as does Andre'ev's in that it is a result on the existence of hyperbolic polyhedra.

I state the result and then backtrack to fill in definitions and discuss the proof.

Polynomially Branched KAT Theorem (Bowers and Stephenson [18]) *Let K be an oriented simplicial triangulation of $\mathbb{S}^2$, different from the tetrahedral triangulation, and let $\Phi : E(K) \to [0, \pi/2]$ be a map assigning angle values to each edge of K. Assume that the following two conditions hold.*

(i) *If e_1, e_2, e_3 form a closed loop of edges from K with $\sum_{i=1}^{3} \Phi(e_i) \geq \pi$, then e_1, e_2, and e_3 form the boundary of a face of K.*
(ii) *If e_1, e_2, e_3, e_4 form a closed loop of edges from K with $\sum_{i=1}^{4} \Phi(e_i) = 2\pi$, then e_1, e_2, e_3, and e_4 form the boundary of the union of two adjacent faces of K.*

If β is a polynomial branch structure for the edge-labeled triangulation (K, Φ), then there exists a circle packing $\mathcal{C} = \{C_v : v \in V(K)\}$ for (K, Φ), a family of circles in $\mathbb{S}^2$ so that the two circles C_v and C_w meet at angle $\Phi(e)$ whenever $e = vw$ is an edge of K, with $\text{br}(\mathcal{C}) = \beta$. *The circle packing $\mathcal{C}$ is unique up to Möbius transformations.*

A branch structure essentially is a listing of some of the vertices of K, each paired with an integer ≥ 2 that indicates how many times the circles adjacent to the ones corresponding to the selected vertices wrap around before closing up. Before making this precise, let's observe that there must be further combinatorial conditions to ensure that a branched circle packing exists for the branch structure. Indeed, note that when there is no branching, the fact that K is a simplicial triangulation implies that the degree of each vertex is at least three, and this local condition guarantees that there are enough circles adjacent to a given circle to wrap around once, with angle sum 2π, at least in the tangency case. A moment's thought will show that if the desire is that there be branching of *multiplicity* $m \geq 2$ at a circle C_v, meaning that the circles adjacent to C_v wrap around m times before closing up, there had better be at least $1 + 2m$ adjacent ones to achieve the angle sum of $2\pi m$. This may not be sufficient but certainly is necessary, and the definition of a polynomial branch structure includes enough combinatorial conditions to ensure sufficiency.

To clothe this discussion in a bit of flesh, suppose that $\mathcal{C} = \{C_v : v \in V(K)\}$ is a circle packing for the pair (K, Φ). For each vertex $v \in V(K)$, identify v with the center of its corresponding circle C_v. Fixing a vertex v, let $v_1, \ldots, v_n$ be the list of neighbors of v forming the consecutive vertices in a walk around the boundary of the star st(v) of v, and let α_i be the measure of the spherical angle $\angle v_i v v_{i+1}$. Then v is said to be a *branch point* of *order* $\mathfrak{o} = m - 1$, or of *multiplicity* m, if $\theta(v) = 2\pi m$ for some integer $m \geq 2$, where $\theta(v) = \alpha_1 + \cdots + \alpha_n$ is the *angle sum* at v; again, see Fig. 5.3. The *branch set* br($\mathcal{C}$) of the circle packing is the set of ordered pairs $(v, \mathfrak{o}(v))$ as v ranges over the branch points and $\mathfrak{o}(v)$ is the order of v. It is clear that the combinatorics of K as well as the values of Φ restrict the branch orders.

My aim is to construct circle packings of $\mathbb{S}^2$ in the pattern of K with overlaps given by Φ with a given, predetermined branch set. Toward this end, I will define a branch structure on the complex $T = K \setminus \mathrm{Int}[\mathrm{st}(v_\infty)]$ that triangulates the closed disk one obtains by deleting one vertex, v_∞, and its incident open cells from K. I will use Φ_T to mean the restriction of Φ to the vertices of T.

Definition (Branch Structure) A set $\beta = \{(v_1, \mathfrak{o}_1,), \ldots, (v_\ell, \mathfrak{o}_\ell)\}$, where $v_i, \ldots, v_\ell$ is a pairwise distinct list of interior vertices of T and each $\mathfrak{o}_i$ is a positive integer, is a *branch structure* for the pair (T, Φ_T) if the following condition holds: for each simple closed edge path $\gamma = e_1 \cdots e_n$ in T that bounds a combinatorial disk D that contains at least one of the vertices v_i, the inequality

$$\sum_{i=1}^{n} [\pi - \Phi_T(e_i)] > 2\pi(\mathfrak{o}(D) + 1) \tag{5.3.3}$$

holds, where $\mathfrak{o}(D) = \sum \mathfrak{o}_i$, the sum taken over all indices i for which $v_i \in \mathrm{Int}(D)$.

We will see that this condition on the combinatorics of T and the values of Φ_T ensures that there are no local obstructions to the existence of a circle packing for (T, Φ_T) whose branch set is β, and in fact is enough to ensure that there are no global ones.

Definition (Polynomial Branch Structure) Let K be a simplicial triangulation of $\mathbb{S}^2$ with edge function $\Phi : V(K) \to [0, \pi/2]$. A collection

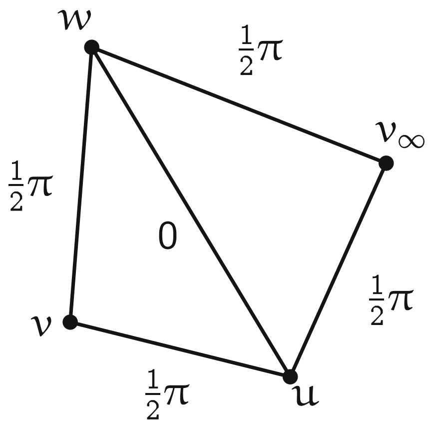
$$\beta = \{(v_\infty, \mathfrak{o}_\infty), (v_1, \mathfrak{o}_1), \ldots (v_\ell, \mathfrak{o}_\ell)\}$$

is a *polynomial branch structure* for (K, Φ) if the following conditions prevail.

(1) $\mathfrak{o}_\infty = \mathfrak{o}_1 + \cdots + \mathfrak{o}_\ell$.
(2) The vertices $v_1, \ldots, v_\ell$ are all interior vertices of the complex $T = K \setminus \mathrm{Int}[\mathrm{st}(v_\infty)]$.
(3) $\beta_T = \{(v_1, \mathfrak{o}_1), \ldots (v_\ell, \mathfrak{o}_\ell)\}$ is a branch structure for (T, Φ_T).
(4) No Φ-edge labeled subgraph of the type given in Fig. 5.4 occurs in K where v is one of the branch vertices $v_1, \ldots, v_\ell$.

A few comments concerning this definition are in order. Item (1) says that there is an even amount of branching and half of it occurs at vertex v_∞; item (2) says that no branch vertex from the list $v_1, \ldots, v_\ell$ is adjacent to the vertex v_∞; item (3) in

Fig. 5.4 A forbidden edge-labeled subgraph for a polynomial branch structure



particular says that Inequality Eq. 5.3.3 holds for (T, Φ_T); item (4) is a technical condition that avoids impossible configurations.

Discussion of Proof How do we put all of this together to prove the Polynomially Branched KAT Theorem? Letting $\beta = \{(v_\infty, \mathfrak{o}_\infty), (v_1, \mathfrak{o}_1), \dots (v_\ell, \mathfrak{o}_\ell)\}$ be a polynomial branch structure for (K, Φ), we remove the vertex v_∞ and work with hyperbolic polyhedral metrics on the disk triangulation T as in the proof of the Koebe Circle Packing Theorem. The idea is the same as there in that we want to use vertex labels on T to describe hyperbolic triangles that then are identified with faces to form a singular hyperbolic surface, and then vary the labels to meet angle targets at the vertices. There are three new difficulties that appear.

(i) Target overlap angles are given by Φ_T for adjacent circles rather than tangencies.
(ii) Rather that 2π, the target angle sums at branch vertices are $2\pi m_i$ for integers $m_i = \mathfrak{o}_i + 1 \geq 2$.
(iii) As the boundary $\partial\mathbb{D}$ ultimately will serve as the circle corresponding to v_∞ in the desired circle packing, the overlaps of the boundary circles of the packing for T must intersect the unit circle at the angles demanded by Φ.

Now items (i) and (ii) are really no problem as superpacking labels can be described that allow for prescribed overlap angles for adjacent circles and target angles prescribed by the branch structure. The real difficulty is item (iii). If we use radius labels, the best we can do is, as in the proof of the Koebe Theorem, get boundary circles that meet the unit circle at single points with intersection angle zero. The hint for resolving this difficulty is found in thinking a bit more about the role of circles in hyperbolic geometry, and in particular in the Poincaré disk model where $\mathbb{H}^2$ is identified with the unit disk $\mathbb{D}$, and the ideal boundary of $\mathbb{H}^2$ is identified with $\partial\mathbb{D} = \mathbb{S}^1$. Euclidean circles that meet the Poincaré disk $\mathbb{D}$ not only serve as hyperbolic circles, but also as horocycles and hypercycles. Those that lie entirely within $\mathbb{D}$ are hyperbolic circles, those internally tangent to the ideal boundary $\mathbb{S}^1$ are horocycles, and those that meet the boundary in two points a and b are hypercycles whose points in $\mathbb{D}$ lie equidistant to the hyperbolic line with ideal endpoints a and b. This latter case includes the hyperbolic geodesic lines. What proves fruitful here is the fact that, when oriented, these Euclidean circles and circular arcs are precisely the curves of constant geodetic curvature in the hyperbolic plane. This is implied immediately by the fact that these are the flow lines of 1-parameter groups of hyperbolic isometries, the hyperbolic circles the flow lines of elliptic flows, horocycles of parabolic flows, and hypercycles of hyperbolic flows.

Here are the salient facts about the geodetic curvature κ of an arc of an oriented Euclidean circle that lies in the Poincaré disk $\mathbb{D}$. Call an arc $c = C \cap \mathbb{D}$, where C is a Euclidean circle that meets $\mathbb{D}$, a *cycle* with *parent circle* C. There is a normalized setting in which the curvature can be read off easily. Apply a conformal automorphism of the disk so that c passes through the origin and its parent circle C is centered on the positive real axis. Orient c counterclockwise and let t, $0 < t \leq \infty$ denote the point of intersection of C with the interval $(0, \infty]$. Then the curvature satisfies $\kappa = \kappa(c) = 1/t$. In terms of intrinsic parameters, for counterclockwise-oriented hypercycles when $t > 1$, $\kappa(c) = \cos\alpha$ where α is the acute angle of intersection of C with the unit circle. This includes the case of a hyperbolic geodesic where $\alpha = \pi/2$ and $\kappa = 0$. Assuming still the counterclockwise orientation, when $t = 1$, c is a horocycle with $\kappa(c) = 1$, and when $t < 1$, $c = C$ is a hyperbolic circle of some hyperbolic radius r with $\kappa(c) = \coth r$.

For our purposes it is quite fortuitous that monotone curvature parameters for cycles can be used as vertex labels on T in place of radii labels to encode a singular hyperbolic metric on a disk that T triangulates. The curvature is inversely related to the radii, but the really important feature is that, unlike radii labels, the curvature label can be used to identify faces of T, not only with hyperbolic triangles with both finite and ideal vertices, but also triangles with "hyperideal vertices."[9] This means that when curvatures κ_1, κ_2, and κ_3 label the vertices of the face f and values $\Phi_T(e_i)$ for $i = 1, 2, 3$ label the opposite edges, the face f may be identified with the region of the hyperbolic plane determined by cycles of curvatures κ_1, κ_2, and κ_3 overlapping with angles $\Phi_T(e_i)$ for $i = 1, 2, 3$. This accomplishes two things. First, the overlaps of cycles are given by the edge function Φ_T. Second, and very importantly, if the vertex w of f is a boundary vertex and the value $\kappa = \cos\Phi_T(wv_\infty)$ is used for the curvature, then the boundary cycle corresponding to w overlaps with the unit circle by an angle of $\Phi_T(wv_\infty)$.

The important point is that the set $\mathfrak{K}$ of curvature labels, ones whose boundary values are given by $g(w) = \cos\Phi(wv_\infty)$ for the boundary vertex w, and that produce superpackings where the angle sums at interior vertices are no more than 2π at non-branch points and no more than $2\pi m_i$ at branch point v_i, may be varied to obtain a β-packing label, this time as $\sup\mathfrak{K}$, the supremum instead of the infimum

[9]When the Klein disk is used as the model for the hyperbolic plane these are in fact Euclidean triangles that meet the disk, but whose vertices may lie within the disk, on the ideal boundary, or outside the closed disk. The hyperideal vertices are the latter ones.

since curvatures are inversely related to radii. Of course by *β-packing label* I mean that the angle sum at any interior vertex that is not a branch vertex is 2π, and at v_i is $2\pi m_i$. The argument is akin to that of the proof of the Koebe Circle Packing Theorem, but, though still elementary, is much more intricate and involved. The full detailed proof appears in [18] where the key proposition, stated below, generalizes the Discrete Boundary Value Theorem of Beardon and Stephenson. Setting up this result with appropriate definitions and analysis of hyperideal hyperbolic triangles, as well as the proof itself, takes up most of the content of the paper.

Proposition 5.3.2 (Bowers and Stephenson [18]) *Let g be a proper boundary label for T and β a branch structure for (T, Φ_T). Then there exists a unique β-packing label $\mathfrak{k}$ for (T, Φ_T) such that $\mathfrak{k}(w) = g(w)$ for every boundary vertex of T.*

This then is used to complete the proof of the Polynomially Branched KAT Theorem by using the circle packing produced by Proposition 5.3.2, augmented by the unit circle corresponding to the removed vertex v_∞ to define $\mathcal{C}$. Much of this becomes routine at this point, except that one still must confirm that half the branching occurs at v_∞. This turns out to be nontrivial. Again the details are rather involved and can be found in [18]. □

5.3.4 Cusps and Cone Type Singularities

In this section I offer a generalization of KAT II where prescribed target angle sums at vertices are assigned, and necessary and sufficient conditions are sought to guarantee existence of such packings. This is the discrete version of the classical Schwarz-Picard problem of the existence of hyperbolic metrics on Riemann surfaces with prescribed cone type singularities. For simplicity I am going to restrict to the tangency case where Φ is identically zero.

To set up the problem, let K be a simplicial triangulation of a compact surface, possibly with boundary, with $\mathtt{F}$ faces, $\mathtt{E}$ edges, and $\mathtt{V}$ vertices. The vertex set $V(K)$ is partitioned into three sets: two disjoint subsets of interior vertices denoted as V_{Int} and V_{cusps}, and the set V_{Bd} of boundary vertices, with respective cardinalities $\mathtt{V}_{\text{Int}}$, $\mathtt{V}_{\text{cusps}}$, and $\mathtt{V}_{\text{Bd}}$. Elements of V_{Int} are called *interior vertices* and of V_{cusps} are called *cusp vertices*. Two functions are given, the first $f : V_{\text{Bd}} \to (0, \infty]$ giving target radii for the boundary vertices and the second $\theta : V_{\text{Int}} \to (0, \infty)$ giving target angle sums at interior vertices. The target angle sums at the *cusp vertices* in V_{cusps} are zero. The task is to give necessary and sufficient conditions on K to guarantee the existence of a packing label $r : V(K) \to (0, \infty]$ for this data such that $r = f$ on V_{Bd}, $r = \infty$ on V_{cusps}, and $\theta_r(v) = \theta(v)$ for every interior vertex $v \in V_{\text{Int}}$.

To describe a solution to this problem, for any set V of vertices, let $\mathtt{F}_V$ denote the number of faces of K that meet V, and let $\theta(V) = \sum_{v \in V} \theta(v)$ denote the total angle sum of the vertices of V. Let

$$\mathfrak{R} = \{r : V(K) \to (0, \infty] : r = f \text{ on } V_{\text{Bd}},\ r = \infty \text{ on } V_{\text{cusps}},\ \theta_r(v) \le \theta(v) \text{ for all } v \in V_{\text{int}}\}.$$

This describes the set of *superpacking labels* for the data θ with boundary values given by f and cusp set V_{cusps}. A *packing label* for this data is a superpacking label where, in addition, the target angle sums given by θ are met, so that $\theta_r(v) = \theta(v)$ for all $v \in V_{\text{Int}}$. For any superpacking label r and vertex set V, let $\theta_r(V) = \sum_{v \in V} \theta_r(v)$. The next theorem gives necessary and sufficient conditions for a solution to the discrete Schwarz-Picard boundary value problem. The proof is a generalization of the proof presented herein for the Koebe Circle Packing Theorem. There the important invariant is $\mathtt{F} - 2\mathtt{V}_{\text{Int}}$. In the borderless case of Theorem 5.3.1, the important invariant is $\mathtt{F} - 2\mathtt{V}$. These arise from writing the hyperbolic area of the surface determined by a packing label, provided one exits, in terms of combinatorial invariants. The corresponding fact in this setting is that, for any packing label $\mathfrak{r}$ for the data f, θ, and V_{cusps},

$$\text{hyp-area}(K(\mathfrak{r})) + \theta_{\mathfrak{r}}(V_{\text{Bd}}) = \pi\,\mathtt{F} - \theta_{\mathfrak{r}}(V_{\text{Int}}) = \pi\,\mathtt{F} - \theta(V_{\text{Int}}).$$

The right hand side of this equation is an invariant of K and θ and must be positive since the left hand side is positive. Also, for every interior vertex v,

$$\theta(v) = \theta_{\mathfrak{r}}(v) < \pi \deg v$$

These give two necessary conditions for a desired label to exist, but these are not sufficient. Nonetheless, these two conditions are the extreme cases of the sufficient condition that appears as item (i) of the theorem.

Discrete Schwarz-Picard Boundary Value Theorem (Bowers [12]) *The following are equivalent.*

(i) *For every edge-path connected set $V \subset V_{\text{Int}}$ of interior vertices, the invariant $\pi\,\mathtt{F}_V - \theta(V)$ is positive.*
(ii) *The function $\mathfrak{r} = \inf \mathfrak{R}$ does not take a zero value at any vertex.*
(iii) *The function $\mathfrak{r} = \inf \mathfrak{R}$ is the unique packing label for K with data f, θ, and V_{cusps}.*
(iv) *There exist a packing label for K for the data f, θ, and V_{cusps}.*

A word of caution is in order. Though this does solve the discrete Schwarz-Picard problem, the combinatorial condition of item (i), that $\pi\,\mathtt{F}_V - \theta(V) > 0$ *for every* path connected subset V of interior vertices, is a very difficult condition to check once the size of K becomes in any way substantial. This pure mathematician has learnt to appreciate the difficulties our computational geometer cousins face when trying to make the elegant output of our theorems practical tools for performing geometric computations. This difficulty often is unrecognized or left unacknowledged by my pure mathematician siblings.

5.4 Infinite Packings of Non-compact Surfaces

I now turn our attention to infinite packings of non-compact surfaces. Here new and interesting phenomena arise, fraught with their own peculiar difficulties. To keep the conversation manageable, I am restricting attention to tangency circle packings of simply connected domains and will concentrate on one very interesting problem that arises in this setting—the *type problem*—and one great success in attacking the Koebe Uniformization Conjecture.

5.4.1 The Discrete Uniformization Theorem

Does every simplicial triangulation K of every topological surface S, compact or not, admit a circle packing in some geometric structure on S? By passing to the universal covering surface $\widetilde{S}$ and lifting the triangulation to a triangulation $\widetilde{K}$ of $\widetilde{S}$, the question may be approached by asking whether any G-invariant simplicial triangulation of a simply connected surface admits a G-invariant circle packing in some geometric structure, where G is a group of symmetries of the complex. There are only two simply connected surfaces up to homeomorphism, the sphere and the plane. The former case is addressed by the Koebe Circle Packing Theorem. In this section I will address the latter case.

Let $\mathcal{T}$ be a *plane triangulation graph*, by which I mean that $\mathcal{T}$ is the 1-skeleton of a simplicial triangulation K of the topological plane. There are precisely two inequivalent conformal structures on the plane, the one conformally equivalent to the complex plane $\mathbb{C}$ and the other to the open unit disk $\mathbb{D}$. There are precisely two complete metrics of constant curvature up to scaling on the plane, the one isometric to Euclidean 2-space $\mathbb{E}^2$ and of constant zero curvature, the other isometric to the hyperbolic plane $\mathbb{H}^2$ and of constant negative curvature. Fortunately, the conformal and the geometric structures mesh nicely in that the complex plane $\mathbb{C}$ is a conformal model of plane Euclidean geometry via its standard Euclidean metric $ds_{\mathbb{C}} = |dz|$, and the disk $\mathbb{D}$ is a conformal model of plane hyperbolic geometry via the Poincaré metric $ds_{\mathbb{D}} = 2|dz|/(1-|z|^2)$. Metric circles in these two geometries are precisely the Euclidean circles contained in their point sets, so circle packings in these geometric surfaces can be identified with Euclidean circle packings of $\mathbb{C}$ and $\mathbb{D}$. I will use $\mathbb{G}$[10] to mean either $\mathbb{C}$ or $\mathbb{D}$ with the intrinsic Euclidean or hyperbolic geometry determined by either $ds_{\mathbb{C}}$ or $ds_{\mathbb{D}}$ when referring to geometric quantities like geodesics and angles, etc. Here is the foundational result in this setting.

Discrete Uniformization Theorem (Beardon and Stephenson [7], He and Schramm [42]) *Every plane triangulation graph $\mathcal{T}$ can be realized as the contacts graph of a univalent circle packing $\mathcal{T}(\mathcal{C})$ that fills exactly one of the complex plane*

[10] $\mathbb{G}$ means $\mathbb{G}$eometry.

$\mathbb{C}$ or the disk $\mathbb{D}$. The packing is unique up to conformal automorphisms of either $\mathbb{C}$ or $\mathbb{D}$.

The *contacts* graph of a collection is a graph with a vertex for each element of the collection and an edge between two vertices if an only if the corresponding elements meet. The *carrier* of the circle packing $\mathcal{C}$ in the geometry $\mathbb{G}$ is the union of the geodesic triangles formed by connecting centers of triples of mutually adjacent circles with geodesic segments, and $\mathcal{C}$ *fills* $\mathbb{G}$ whenever its carrier is all of $\mathbb{G}$. When $\mathcal{C}$ is univalent and fills $\mathbb{G}$, $\mathcal{C}$ is said to be a *maximal packing* for $\mathcal{T}$ or K, and K may be realized as a geodesic triangulation of $\mathbb{G}$ whose vertices are the centers of the circles of $\mathcal{C}$ with geodesic edges connecting adjacent centers.

Once this theorem is in place, the whole of the theory of tangency circle packings on non-compact surfaces comes into play. As already indicated, in a thoroughly classical way packing questions on surfaces can be transferred to questions of packings on simply connected surfaces, this by passing to covering spaces acted upon by groups of deck transformations. Any combinatorial symmetries of the complex K are realized as automorphic symmetries of $\mathbb{G}$, this from the uniqueness of the Discrete Uniformization Theorem, and this offers an alternate Proof of Theorem 5.3.1, and an extension of that theorem to triangulations of arbitrary, non-compact surfaces.

Beardon and Stephenson [7] proved the Discrete Uniformization Theorem when $\mathcal{T}$ has *bounded degree*, a global bound on the degrees of all the vertices of $\mathcal{T}$. In this foundational paper as well as in their subsequent one [8], Beardon and Stephenson laid out a beautiful theory of circle packings on arbitrary surfaces, gave a blueprint for developing a theory of discrete analytic functions, and articulated one of the most interesting problems in the discipline, that of the *circle packing type problem* for non-compact surfaces, this latter the subject of the section following. The bounded degree assumption was needed both to verify that the packing fills $\mathbb{G}$ and for the uniqueness, and He and Schramm [42] removed the bounded degree hypothesis and proved the general case where there is no global bound on the degrees of vertices. Earlier, Schramm [65] had proved a very general rigidity theorem for infinite packings of planar domains whose complementary domains are a countable collection of points, and He and Schramm [42] extended this to general countably connected domains.

Discussion of Proof The full proof is scattered throughout several articles published in the 1990s. In what constitutes a significant service to the discipline, Ken Stephenson has laid out a complete proof in roughly fifty pages of his wonderful text *Introduction to Circle Packing* [71]. I have not the space here to do justice to the argument, but I will make some comments.

Beardon and Stephenson's proof of existence relies on the Maximal Disk Packing Theorem and uses a diagonal argument on a sequence of finite subcomplexes of K that exhausts K. It does not depend on any bounded degree assumption and is quite straightforward. The proof of existence goes like this. Write $K = \cup_{i=1}^{\infty} K_i$ as a nested, increasing union of finite subcomplexes K_i, each a simplicial triangulation of a closed disk. Apply the Maximal Disk Packing Theorem to obtain a sequence

$\mathcal{C}_i$ of univalent, maximal circle packings for the complexes K_i in the unit disk $\mathbb{D}$ realized as the Poincaré disk model of hyperbolic geometry. Fix a base vertex v_0 of K_1 and let C_i be the circle of $\mathcal{C}_i$ that corresponds to v_0. By applying an automorphism of the disk if needed, assume that C_i is centered at the origin and of hyperbolic radius $\mathfrak{r}_i(v_0)$. Now the Discrete Schwarz-Pick Lemma implies that the sequence $\mathfrak{r}_i(v_0)$ of hyperbolic radii is non-increasing, hence has a limit, say $\mathfrak{r}(v_0) \geq 0$, as $i \to \infty$. There are two cases.

(I) The limit radius $\mathfrak{r}(v_0) \neq 0$;
(II) The limit radius $\mathfrak{r}(v_0) = 0$.

The first claim is that if v is any other vertex of K whose corresponding circle of $\mathcal{C}_i$, for large enough i, has hyperbolic radius $\mathfrak{r}_i(v)$, then $\lim_{i\to\infty} \mathfrak{r}_i(v)$ is not zero when case (I) occurs and is equal to zero when case (II) occurs. This means that the limit radius function $\mathfrak{r} : V(K) \to [0, \infty)$ never takes a zero value in case (I) and is identically zero in case (II). The proof of this claim uses the *Ring Lemma* of Burt Rodin and Dennis Sullivan that was crucial in [62] in their confirmation of Thurston's outlined proof of the *Discrete Riemann Mapping Theorem* presented in his 1985 Purdue lecture; see Sect. 5.5.1. The Ring Lemma guarantees the existence of a sequence of positive constants $\mathfrak{c}_d$ such that, when $d \geq 3$ disks form a cycle of sequentially tangent disks all tangent to a central disk of Euclidean radius R, and the disks have pairwise disjoint interiors, then the smallest disk has Euclidean radius $\geq \mathfrak{c}_d R$. The Ring Lemma is applied as follows. Let $v_0\, v_1 \cdots v_n = v$ be a path of vertices in K from v_0 to v and choose N so large that this path of vertices is contained in the interior of K_i, for all $i \geq N$. The Ring Lemma applied sequentially to the chain of pairwise tangent circles in $\mathcal{C}_i$ corresponding to the path $v_0\, v_1 \cdots v_n = v$ implies that there is a positive constant $\mathfrak{c}$ such that $R_i(v) \geq \mathfrak{c}\, R_i(v_0)$, where R_i is the Euclidean radius function on $\mathcal{C}_i$. This holds for all $i \geq N$ and the constant $\mathfrak{c}$ is independent of i. As hyperbolic and Euclidean radii of circles in the disk are comparable in the small, this implies the claim.

Now order the vertex set $V(K)$ as $v_0, v_1, \ldots$. In case (I), choose a subsequence i_j so that the hyperbolic centers of the circles of the sequence $\mathcal{C}_{i_j}$ all corresponding to the vertex v_1 converge in the closed disk $\overline{\mathbb{D}}$ to a point c_1. An application of item (ii) of the Discrete Schwarz Pick Lemma implies that c_1 is contained in the open disk $\mathbb{D}$. Repeat to find a subsequence of i_j for which the hyperbolic centers of the circles corresponding to v_2 converge to a point c_2 in $\mathbb{D}$. Iterating and applying a diagonal argument gives a subsequence of the sequence of circle packings $\mathcal{C}_i$ for which the hyperbolic centers of the circles corresponding to the vertex v_n of K converges to a point c_n in $\mathbb{D}$ for all positive integers n. Centering a circle of hyperbolic radius $\mathfrak{r}(v_n)$ at the point c_n produces a circle packing in the Poincaré disk in the pattern of K. In case (II) when $\mathfrak{r}$ is identically zero, a diagonal argument applied to the scaled packing $\frac{1}{R_i}\mathcal{C}_i$, where R_i is the Euclidean radius of C_i, produces a circle packing in the plane $\mathbb{C}$ in the pattern of K. Call the limit circle packing in either case $\mathcal{C}$.

There are three facts left to prove: first, that $\mathcal{C}$ is univalent; second, that $\mathcal{C}$ fills the disk in case (I) and the plane in case (II); third, that $\mathcal{C}$ is unique up to automorphisms. The first claim of univalence follows from the fact that each circle packing $\mathcal{C}_i$ is univalent and the convergent subsequence of radii and centers described above essentially describes geometric convergence of circle packings. Beardon and Stephenson's original proof of the second claim that the packing fills $\mathbb{G}$ relied critically on the bounded degree assumption. It was used to ensure that piecewise linear maps from the complexes K_i into the geometry $\mathbb{G}$ defined using the convergent sequence of circle packings are uniformly quasiconformal so that the Carathéodory Kernel Theorem [27] applies to ensure that the image of the limit function is the kernel of the image sets, which is the whole of $\mathbb{G}$. The third claim of uniqueness in the hyperbolic case (I) follows from the uniqueness of the limiting radius function, but in the Euclidean case (II), uniqueness uses the bounded degree assumption. Later He and Schramm removed the bounded degree assumption. Their proof of uniqueness in case (II) is particularly elegant. It is a topological proof based on the winding numbers of mappings defined on the boundaries of corresponding intersticial regions in two circle packings for the same complex K, both of which fill $\mathbb{C}$. All of this is rather nicely laid out in Stephenson's *Introduction to Circle Packing* [71]. □

5.4.2 *Types of Type*

The dichotomy between hyperbolic and Euclidean behavior is evident in the Discrete Uniformization Theorem. Indeed, the combinatorial complex K, or its 1-skeleton $\mathcal{T}$, determines uniquely its geometry in that the maximal circle packing $\mathcal{T}(\mathcal{C})$ fills either the disk $\mathbb{D}$ or the complex plane $\mathbb{C}$, but forbids two packings where one fills the disk and the other the plane. This leads to the next definition.

Definition (CP-Type) A simplicial triagulation K of the plane, and its 1-skeleton plane triangulation graph $\mathcal{T} = K^{(1)}$, are said to *CP-parabolic* or *CP-hyperbolic* when the maximal circle packing $\mathcal{T}(\mathcal{C})$ fills respectively the complex plane $\mathbb{C}$ or the disk $\mathbb{D}$. The *CP-type problem* is the problem of determining whether a given complex K or plane triangulation graph $\mathcal{T}$ is CP-parabolic or CP-hyperbolic. One seeks conditions or invariants on the complex K or the graph $\mathcal{T}$, reasonably checked or computed, that can determine which of the two CP-types adheres. See Fig. 5.5.

This is a discrete version of the classical *conformal type problem*, or just *type problem* for short, that of determining whether, à la classical Uniformization Theorem, a given non-compact simply connected Riemann surface is *parabolic* and conformally equivalent to the complex plane $\mathbb{C}$, or *hyperbolic* and conformally equivalent to the disk $\mathbb{D}$.

Historically this is not the first discrete type problem. That honor probably goes to the problem of determining the *random walk type*, or *RW-type* for short, of an infinite graph. My aim in this section is to review this and several other species of discrete type problems and explore their interactions in the context of plane

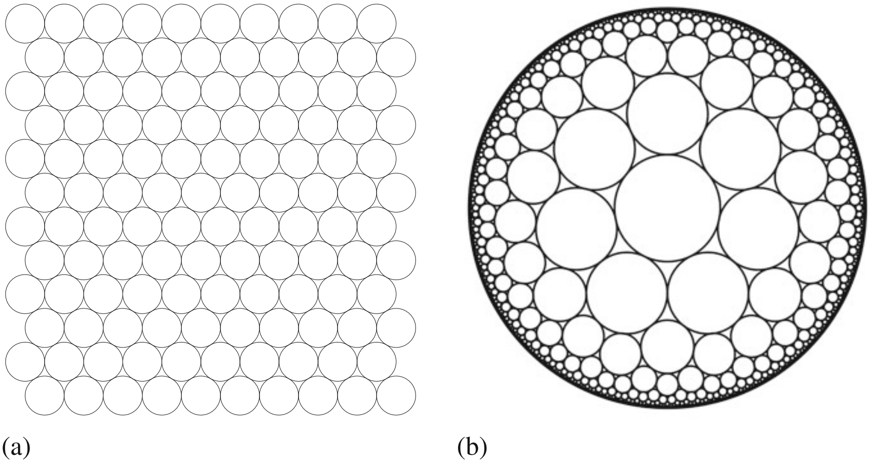


Fig. 5.5 The CP-type of a plane triangulation graph is determined by the corresponding maximal circle packing and whether it fills the plane or the disk. (**a**) The penny packing, the maximal circle packing for the constant 6° plane triangulation graph $\mathcal{G}_6$, fills the plane $\mathbb{C}$. The graph $\mathcal{G}_6$ is parabolic. (**b**) The maximal circle packing for the constant 7° plane triangulation graph $\mathcal{G}_7$ fills the disk $\mathbb{D}$. The graph $\mathcal{G}_7$ is hyperbolic, as are the graphs $\mathcal{G}_d$ for all $d \geq 7$

triangulation graphs. In all I will examine six different species of discrete type that go under the abbreviations CP, RW, EL, EEL, VEL, and EQ-type.

Consider the standard simple random walk on a simple, connected, locally finite graph $\mathcal{G}$ where the probability of walking across a particular edge uv from vertex u to vertex v is $1/\deg u$. The graph $\mathcal{G}$ is said to be *random walk parabolic*, or *RW-parabolic*, if a walker almost surely returns to a fixed base vertex, and *random walk hyperbolic* or *RW-hyperbolic* otherwise. More common terminology is that the graph is *recurrent* when RW-parabolic and *transient* when RW-hyperbolic. In a transient graph, a random walker has a positive probability for escaping to infinity whereas in a recurrent one, the escape probability vanishes and, in fact, the walker almost surely returns to every vertex infinitely often. Woess [74] is a fantastic reference for the classical theory of random walks on graphs and Lawler and Limic [51] is an up-to-date reference with many recent results.

Early on in the development of circle packing theory, Ken Stephenson made a connection between the CP- and RW-type problems. The intuition for the connection arises from the close connection in classical complex function theory between the conformal type problem and brownian motion on a Riemann surface. Stephenson [70] proved that the CP- and RW-types of bounded degree plane triangulation graphs always coincide. Later in [43], He and Schramm gave an example of a plane triangulation graph, necessarily of unbounded degree, that is CP-parabolic but RW-hyperbolic. There the authors focused more sharply on the distinction between these two species of type and recalled Duffin's EEL-type from [35] and developed Cannon's VEL-type inspired by Cannon [25] in articulating the distinction.

Before continuing with the discussion of CP-type, let's review a bit of history. The story of discrete type really begins in the 1920s with Póyla's study [57] of the RW-type of the integer grid in $\mathbb{R}^d$ where he proved that the integer grid in $\mathbb{R}^2$ is RW-parabolic while the grid in higher-dimensional Euclidean spaces is RW-hyperbolic. In 1959, Nash–Williams in [55] used a method of Lord Rayleigh to link the RW-type of a locally finite graph with its resistance to electric flow when each edge is thought of as a wire with a unit of electrical resistance, giving rise to EL-type. To be a bit less cryptic, when an infinite graph is thought of as an electric network with each edge representing a wire of unit resistance, the question is whether electricity will flow from a base vertex to infinity when a unit potential is applied to the base vertex and infinity is grounded. This is made a bit more precise by asking what the effective resistance is from the base vertex to infinity for the network. When the effective resistance to infinity is infinite, no current flows and the network is *EL-parabolic*, and when the resistance is finite so that current does flow, the network is *EL-hyperbolic*. In the beautiful 1984 Carus Mathematical Monograph [34] entitled *Random Walks and Electric networks*, Peter Doyle and J. Laurie Snell present an accessible proof that the RW- and EL-type of an infinite graph coincide. In 1962, Duffin [35] gave a combinatorial invariant of a graph, the *edge extremal length*, that characterizes the RW- or EL-type according to whether the edge extremal length of the set of transient edge-paths is infinite or not.

What is the edge extremal length of a path family? It is a discrete version of the classical conformal extremal length of a path family in a Riemann surface in quasiconformal analysis. For a graph $\mathcal{G}$, let Γ be any family of edge-paths, infinite or not. The edge extremal length is obtained by measuring the minimal length-squared of the curves in Γ divided by the area, this maximized over all metric assignments. This is the same as the classical definition, only what changes is how the admissible metrics are assigned. Here are the details. An *edge-path* in $\mathcal{G}$ is a finite or infinite sequence $\mathbf{e} = e_1, e_2, \dots$ of directed edges of $\mathcal{G}$ with the terminal vertex of e_i equal to the initial vertex of e_{i+1}. An *edge metric* on $\mathcal{G}$ is a function $m : E(\mathcal{G}) \to [0, \infty]$ that assigns a non-negative value to each edge, and the *area of* m is defined as $\text{area}(m) = \sum_{e \in E(\mathcal{G})} m(e)^2$. An edge metric is *admissible* if its area is finite and I will let $M_E(\mathcal{G})$ denote the collection of admissible edge metrics. The *m-length* of the edge-path $\mathbf{e}$ is $\ell_m(\mathbf{e}) = \sum_{i=1} m(e_i)$. Finally, the *edge extremal length* of the family Γ of edge-paths is

$$\text{EEL}(\Gamma) = \sup_{m \in M_E(\mathcal{G})} \frac{\inf_{\mathbf{e} \in \Gamma} \ell_m(\mathbf{e})^2}{\text{area}(m)}.$$

The notation $\text{EEL}(\mathcal{G})$ is reserved for the case where Γ is the set of paths to infinity that start at a given base vertex v_0. These are called the *transient* edge-paths in $\mathcal{G}$ based at v_0, and any such transient edge-path $\mathbf{e} \in \Gamma$ has initial vertex v_0 at its first edge e_1 and is not contained in any finite collection of edges. One says that the graph $\mathcal{G}$ is *EEL-parabolic* if $\text{EEL}(\mathcal{G}) = \infty$ and *EEL-hyperbolic* otherwise. It is an easy exercise to confirm that EEL-type does not depend on which base vertex is chosen. Duffin's result of [35] already mentioned is that both the RW- and EL-type of a graph coincides with the EEL-type. This was the state of the art in discrete type in the early 1990s when Stephenson connected CP-type with RW-type for bounded degree plane triangulation graphs.

In 1995, He and Schramm [43] in a remarkable article clarified the role of the bounded degree assumption. There, after constructing a plane triangulation graph that, though CP-parabolic, is RW-hyperbolic, they applied Cannon's vertex extremal length to characterize CP-type combinatorially in the way that edge extremal length characterizes RW-type. Cannon [25] introduced the vertex extremal length of a discrete curve family made of shinglings and used it as a tool for assigning combinatorial moduli to ring domains in the space at infinity of a negatively curved group. He and Schramm adapted Cannon's vertex extremal length to Duffin's development of EEL-type to create VEL-type. The adjustment merely replaces edge-paths by vertex-paths and edge metrics by vertex metrics. The *vertex extremal length* of a family Δ of vertex paths is

$$\mathrm{VEL}(\Delta) = \sup_{m \in M_V(\mathcal{G})} \frac{\inf_{\mathbf{v} \in \Delta} \ell_m(\mathbf{v})^2}{\mathrm{area}(m)}. \tag{5.4.1}$$

Here, a *vertex-path* is a sequence $\mathbf{v} = v_1, v_2, \ldots$ where each v_i is incident with its successor v_{i+1}, and a *vertex metric* is a non-negative function $m : V(\mathcal{G}) \to [0, \infty]$ with *area* $\mathrm{area}(m) = \sum_{v \in V(\mathcal{G})} m(v)^2$. The m-length of the vertex-path $\mathbf{v}$ is $\ell_m(\mathbf{v}) = \sum_{i=1} m(v_i)$ and the set of *admissible metrics*, the ones of finite area, is denoted as $M_V(\mathcal{G})$. The VEL-type of $\mathcal{G}$ now is defined analogously to EEL-type. Indeed, $\mathrm{VEL}(\mathcal{G})$ means $\mathrm{VEL}(\Delta)$, where Δ is the set of *transient* vertex-paths based at v_0, those that meet infinitely many vertices. The graph $\mathcal{G}$ is *VEL-parabolic* if $\mathrm{VEL}(\mathcal{G}) = \infty$ and *VEL-hyperbolic* otherwise, and again it is an easy exercise to confirm that VEL-type does not depend on which base vertex is chosen. This seemingly innocuous adjustment to the definition of EEL-type turns out to be precisely the tool needed to characterize CP-type.

Though, easily, the EEL- and VEL-types of a bounded degree graph coincide, they may differ for a graph of unbounded degree. The relationships between the four types—RW, EL, EEL, VEL—are summarized in the next theorem.

Discrete Type Theorem for Graphs *Let $\mathcal{G}$ be a connected, infinite, locally finite graph.*

(i) [Nash–Willliams [55], Duffin [35]] *The three types—RW, EL, EEL—coincide for $\mathcal{G}$.*
(ii) [He–Schramm [43]] *If $\mathcal{G}$ is EEL-parabolic then it is VEL-parabolic. If $\mathcal{G}$ has bounded degree and is VEL-parabolic, then it is EEL-parabolic.*
(iii) [He–Schramm [43]] *There is a VEL-parabolic plane triangulation graph that is EEL-hyperbolic, necessarily of unbounded degree.*

For a plane triangulation graph $\mathcal{T}$, all five types—RW, EL, EEL, VEL, CP—coincide provided $\mathcal{T}$ has bounded degree. As stated above, it was Stephenson who first proved this for RW- and CP-types. He and Schramm clarified the need for the bounded degree hypothesis, and the relationship between discrete types for plane triangulation graphs is summarized next.

Discrete Type Theorem for Plane Triangulation Graphs (He–Schramm [43]) *Let $\mathcal{T}$ be a plane triangulation graph. Then $\mathcal{T}$ is CP-parabolic if and only if it is VEL-parabolic.*

The proofs of these theorems are quite difficult and involved, though still elementary, and space forbids any sort of discussion of the proofs that would do justice to the subject. Suffice it to say that the interested reader can do no better than to consult the references cited in this section to fill in gaps in the desired detail of proofs.

The Discrete Type Theorem for Plane Triangulation Graphs reduces the very difficult problem of determining whether the maximal circle packing for $\mathcal{T}$ is parabolic or hyperbolic to a combinatorial computation on the graph $\mathcal{T}$. The disappointment comes when one actually tries to do the computation of VEL($\mathcal{T}$) from Eq. 5.4.1 for almost any given plane triangulation graph. One then finds out just how difficult it is to perform this computation; nonetheless, this development is useful for some theoretical considerations. For example, He and Schramm use the theorem to extend Stephenson's result on RW- and CP-type. Here is an interesting result of the author that uses the computation of Eq. 5.4.1 for the proof of item (ii) of the theorem.

Theorem 5.4.1 (Bowers [13]) *Let $\mathcal{G}$ be a connected, infinite, locally finite graph and $\mathcal{T}$ a plane triangulation graph.*

(i) *If $\mathcal{G}$ is Gromov negatively curved and its Gromov boundary contains a nontrivial continuum, then $\mathcal{G}$ is RW-hyperbolic.*
(ii) *If $\mathcal{T}$ is Gromov negatively curved, then $\mathcal{T}$ is CP-parabolic if and only if its Gromov boundary is a singleton; alternately, it is CP-hyperbolic if and only if its Gromov boundary is a topological circle.*

I refer the reader to the appendix of the article [13] for definitions and basic theorems on Gromov negatively curved graphs and metric spaces. To show how the computation from Eq. 5.4.1 may proceed, I'll prove the lemma used in [13] to prove the first assertion of item (ii) of Theorem 5.4.1.

Lemma 5.4.2 *Let v_0 be a vertex in the connected, infinite, locally finite graph $\mathcal{G}$ and let $\{V_n\}$ be a sequence of pairwise disjoint sets of vertices, each of which separates v_0 from infinity. Suppose there exist positive constants C and ε such that, for $n \geq N$,*

$$\mathrm{Card}(V_n) \leq Cn.$$

Then the graph $\mathcal{G}$ is VEL-parabolic.

Proof Define the vertex metric m by $m(v) = 1/(n \log n)$ for any $v \in V_n$ when $n \geq N$, and $m(v) = 0$ otherwise. Then m is admissible since

$$\text{area}(m) = \sum_{n=N}^{\infty} \frac{\text{Card}(V_n)}{(n \log n)^2} \leq \sum_{n=N}^{\infty} \frac{C}{n(\log n)^2} < \infty.$$

For any transient vertex-path $\mathbf{v}$, the m-length satisfies $\ell_m(\mathbf{v}) \geq \sum_{n=N}^{\infty} 1/(n \log n) = \infty$, hence every transient vertex-path has infinite m-length. This implies that $\text{VEL}(\mathcal{G}) = \infty$ and $\mathcal{G}$ is VEL-parabolic. □

I'll end this section with a sixth version of discrete type that is of recent interest in several settings. It arose first for me when Ken Stephenson and I constructed expansion complexes of finite subdivision rules, for the first time in [19] when examining the pentagonal subdivision rule of Cannon, Floyd, and Parry [26]. More recently it arises in our examination of hierarchical conformal tilings [21, 22], and in Gill and Rohde's [37] examination of random planar maps. I name this version of discrete type *EQ-type* with *EQ* an abbreviation for *equilateral*. A plane triangulation graph $\mathcal{T} = K^{(1)}$ can be used to build a piecewise equilateral surface by setting each edge to unit length and isometrically gluing unit-sided equilateral triangles along their boundaries to the boundaries of the faces of K. This produces a piecewise flat surface $|\mathcal{T}|_{\text{eq}}$ that has a natural conformal atlas obtained as follows. Each edge e of $\mathcal{T}$ indexes a chart map φ_e defined on the interior of the union of the faces incident with e. These have been identified with unit equilateral triangles and the chart map φ_e is an orientation-preserving isometry to the plane $\mathbb{C}$. Each vertex v also indexes a chart map φ_v defined on the metric neighborhood of v in $|\mathcal{T}|_{\text{eq}}$ of radius $1/2$, and uses an appropriate complex power mapping to flatten that neighborhood to a disk in the plane $\mathbb{C}$. The overlap maps are conformal homeomorphisms between the appropriate domains. The chart family $\mathcal{A} = \{\varphi_x : x \in V(\mathcal{T}) \cup E(\mathcal{T})\}$ forms a complex atlas making $|\mathcal{T}|_{\text{eq}}$ into a non-compact simply connected Riemann surface $S(\mathcal{T})$. The type problem now is manifest. Is $S(\mathcal{T})$ conformally the plane $\mathbb{C}$ or the disk $\mathbb{D}$? In the former case, $\mathcal{T}$ and K are said to be *EQ-parabolic*, in the latter *EQ-hyperbolic*.

Notice that the question of the EQ-type of a plane triangulation graph is the classical question of the conformal type of a simply connected Riemann surface. It bares the moniker *discrete* because of how the surface is built—using discrete building blocks, the equilateral triangles, glued in a combinatorial pattern encoded in $\mathcal{T}$. The desire is for a combinatorial invariant of $\mathcal{T}$ or K that will determine its EQ-type. So, what relationship exists between the discrete types already discussed and EQ-type? For plane triangulation graphs of bounded degree, easy arguments using quasiconformal mappings show that EQ-type coincides with CP-type—just map the equilateral triangle in $|\mathcal{T}|_{\text{eq}}$ at face f to the corresponding geodesic triangle in $\mathbb{G}$. When $\mathcal{T}$ has bounded degree, this map is uniformly quasiconformal and so the EQ-type agrees with the conformal type of $\mathbb{G}$. For unbounded degree plane triangulation graphs, it remains an open question as to whether the EQ-type coincides with, say, the EEL- or the VEL-type, or perhaps neither. I am bold enough to offer the following conjecture.

Conjecture 5.4.3 *For any plane triangulation graph, EQ-type coincides with VEL-type, and therefore with CP-type.*

A great reference for various expressions of discrete type and their stability under subdivision is Bill Wood's doctoral thesis [75] and the subsequent article [76]. I now turn our attention to Koebe's original inspiration for his circle packing theorem, his interest in circle domains, uniformization, and the Kreisnormierungsproblem.

5.4.3 Koebe Uniformization for Countably-Connected Domains

Zheng-Xu He and Oded Schramm's work on circle packing in the late 1980s and early 1990s led them to a study of Koebe's Uniformization Conjecture. Though the discrete circle packing tools they developed and used did not directly apply to Koebe's problem, the perspective they had gained turned out to be useful. By 1992–1993, they had made the greatest advance on Koebe's problem since its articulation and had proved a circle packing version that greatly generalized the Discrete Uniformization Theorem. Their work is detailed in the *Annals of Mathematics* article *Fixed points, Koebe uniformization, and circle packings*. The proofs are rather intricate and so I am content to state the two main results without any indication of the proofs, leaving it to the interested reader to peruse [42] for details.

He–Schramm Uniformization Theorem (He and Schramm [42], Schramm [67]) *Every countably connected domain in the Riemann sphere is conformally homeomorphic to a circle domain. Moreover, the circle domain is unique up to Möbius transformations and every conformal automorphism of the circle domain is the restriction of a Möbius transformation.*

A *domain triangulation graph* is the 1-skeleton of a simplicial triangulation of a planar domain.

He–Schramm Discrete Uniformization Theorem (He and Schramm [42]) *Every domain triangulation graph with at most countably many ends has a univalent circle packing in the plane $\mathbb{C}$ whose carrier is a circle domain. Moreover, the circle packing is unique up to Möbius transformations.*

He and Schramm prove a theorem that generalizes their Uniformization Theorem to *generaized domains* and *generalized circle domains*. This more general unifomization theorem then is used to give a quick proof of their Discrete Uniformization Theorem.

I'll close this section by mentioning that Schramm in a 1995 paper [67] introduced the notion of *transboundary extremal length* that generalizes the classical extremal length of curve families. Transboundary extremal length is more suited to path families in multiply connected domains that allow for the curves of the family to pass through the complementary components of the domain. Using this tool, Schramm gives a short proof of Koebe uniformization of countably connected domains and generalizes it in two ways. First, he shows that circle domains as the target of uniformization may be replaced by more general domains, namely, those where the complementary components are what he calls τ*-fat sets*. Second, he shows that some domains with uncountably many complementary components may be uniformized to circle domains, namely those where the complementary components are uniformly fat. This includes for example domains whose boundary components are points and μ-quasicircles for a fixed constant $\mu \geq 1$.

5.5 Some Theoretical Applications

The theoretical work in circle packing has grown up hand-in-hand with various applications. In the past score of years, the needs of computer imaging have added a practical bent to the applications with the use of the theory for everything from medical imaging to 3D-printer head guidance. This has been one of the impetuses for the development of the discipline of discrete differential geometry with discrete conformal geometry as but one of its chapters. Circle packing theory à la Thurston as described in this chapter is one flavor of this, but several groups of computational geometers and computer scientists have developed discrete conformal geometry in a great variety of ways, with new techniques designed to solve both practical and theoretical problems. The discipline has grown to a vast enterprise too large and complicated for a review of this type. Rather than attempt a thorough discussion of these applications, I'll only mention a couple of the theoretical applications. The first stands as one of the linchpins of the discipline, and the second generalizes the first. I'll leave it for the interested reader to peruse the many resources available to learn of the state of the art today in practical applications.

5.5.1 Approximating the Riemann Mapping

The event that really got circle packing launched, piquing the interest of a small group of research mathematicians from as diverse fields as complex function theory, combinatorial and computational geometry, geometric topology, and the classical theory of polyhedra, was Bill Thurston's address entitled *The Finite Riemann Mapping Theorem* at Purdue University in 1985. He presented there an algorithm for computing discrete versions of the Riemann mapping of a fixed, proper, simply connected domain in the complex plane $\mathbb{C}$ to the unit disk $\mathbb{D}$, with an indication of why the discrete mappings should converge to a conformal homeomorphism of the domain onto $\mathbb{D}$. Burt Rodin and Dennis Sullivan published in [62] a proof of Thurston's claims in 1987, and this began a steady output of published research on circle packings that continues today. Here I review the content of Thurston's 1985 talk and explain the Rodin-Sullivan verification of Thurston's claims.

Thurston's algorithm is illustrated nicely in the graphics of Fig. 5.6. The scheme is rather simple. Overlay a domain D with a hexagonal circle packing $\mathcal{H}_\varepsilon$ of constant circle radii ε, a 'penny packing.' Use the domain D as a cookie cutter to cut out a portion of the packing, say $\mathcal{P}_\varepsilon$, whose combinatorics are given by the simplicial complex T_ε. Apply the Maximal Disk Packing Theorem to obtain a maximal circle packing $\mathcal{Q}_\varepsilon$ of the disk $\mathbb{D}$. Choosing two points x and y in the domain D, let u_ε and v_ε be the vertices of T_ε whose corresponding circles are closest to the respective points x and y. From the uniqueness of the Maximal Disk Packing Theorem, one may assume that the packings $\mathcal{Q}_\varepsilon$ have been normalized so that the circle corresponding to u_ε is centered at the origin and the one corresponding to v_ε is centered on the positive real axis. Define the *discrete Riemann mapping* $f_\varepsilon : \text{carr}(\mathcal{P}_\varepsilon) \to \text{carr}(\mathcal{Q}_\varepsilon)$ as the piecewise linear mapping that takes centers of circles of $\mathcal{P}_\varepsilon$ to corresponding centers of circles of $\mathcal{Q}_\varepsilon$. Thurston's claim of his 1985 lecture that Rodin and Sullivan verified in 1987 is the content of the next theorem.

Discrete Riemann Mapping Theorem (Rodin and Sullivan [62]) *The mappings f_ε converge as $\varepsilon \downarrow 0$, uniformly on compact subsets of D, to the Riemann mapping f of D onto $\mathbb{D}$ with $f(x) = 0$ and $f(y) > 0$.*

Before I discuss the proof, I should say that there is nothing special about the hexagonal combinatorics. He and Schramm [44] verified that the particular combinatorics of the overlay packings are irrelevant as long as the maximum circle radii approach zero.

Sketch of Proof The proof applies classical tools from quasiconformal analysis to confirm convergence of the discrete mappings to the Riemann mapping. There are three parts. First, the Ring Lemma, already used on page 198 in the proof of the Discrete Uniformization Theorem, is used to observe that the discrete Riemann mappings f_ε for $\varepsilon > 0$ form a family of uniformly quasiconformal mappings with, say, dilatation of all maps bounded by $\mu \geq 1$. Second, standard results of quasiconformal analysis imply that the mappings converge to a μ-quasiconformal mapping f of D onto $\mathbb{D}$. Third, the limit mapping is proved to be 1-quasiconformal, or just conformal, so that it is a Riemann mapping of the domain D onto the disk $\mathbb{D}$. Allow me to fill in each of the three parts of the argument a bit.

The first part, that the discrete Riemann mappings have quasiconformal distortion uniformly bounded, uses the fact that simplicial homeomorphisms are μ-quasiconformal with the distortion constant μ depending only on the shapes of the triangles involved. In particular, because the complexes T_ε have constant degree six on interior vertices, the Ring Lemma implies that there is a minimum possible angle $\omega > 0$ for any of the triangles in the Euclidean carrier $\text{carr}(\mathcal{Q}_\varepsilon)$, this independent of ε. This implies that the discrete maps f_ε are uniformly μ-quasiconformal since the

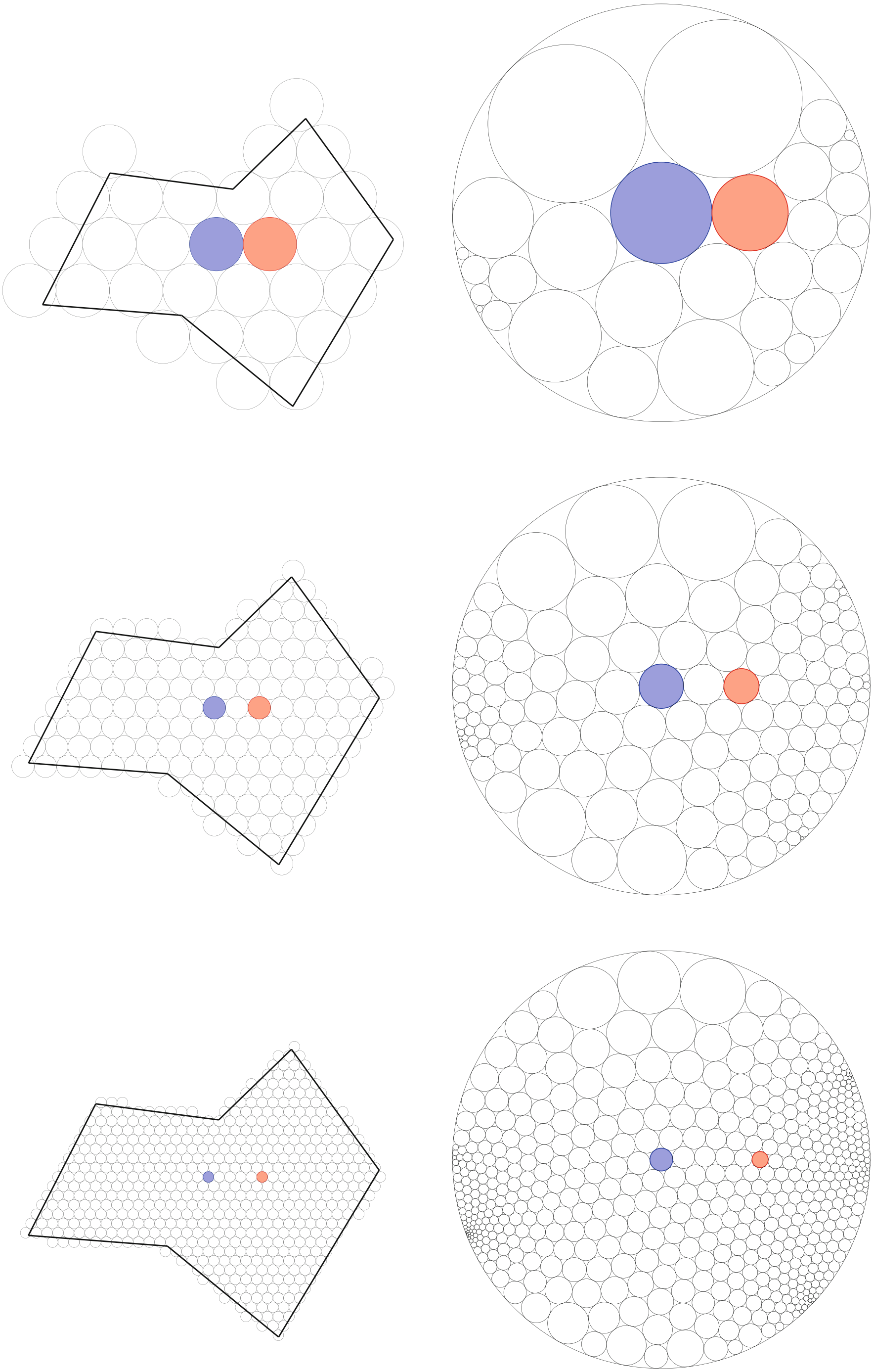


Fig. 5.6 Discrete Riemann mappings with finer and finer hexagonal circle packings

images of the equilateral triangles of carr($\mathcal{P}_\varepsilon$) are triangles of carr($\mathcal{Q}_\varepsilon$) of uniformly bounded distortion.

The second part now follows from standard tools of quasiconformal analysis. The uniformly quasiconformal maps f_ε are equicontinuous on compact subsets of D, as are the maps f_ε^{-1} on compact subsets of $\mathbb{D}$. It follows that the family $\{f_\varepsilon\}_{\varepsilon>0}$ is a normal family and any limit mapping f is bijective between D and $\mathbb{D}$. This latter claim uses the fact that any limit mapping is necessarily μ-quasiconformal, and the Carathéodory Kernel Theorem implies that f takes D onto $\mathbb{D}$.

Finally, that any limit mapping f is conformal follows from the *Hexagonal Packing Lemma*. This says that in a packing with hexagonal combinatorics, any two adjacent circles buried deeply within the packing have nearly equal radii. Here is the exact statement.

Hexagonal Packing Lemma (Rodin and Sullivan [62]) *There is a sequence c_n decreasing to zero as $n \to \infty$ such that in any packing with n generations of the regular hexagonal combinatorics surrounding circle C, the ratio of radii of C and any adjacent circle differs from unity by less than c_n.*

This lemma shows that as $\varepsilon \downarrow 0$, the mappings f_ε restricted to a fixed compact subset of D maps equilateral triangles to triangles of carr($\mathcal{Q}_\varepsilon$) that become arbitrarily close to equilateral, and this implies that any limit mapping is conformal.

This completes the proof modulo the proof of the Hexagonal Packing Lemma. This is proved as follows. Let H_n be any packing of circles in the plane with combinatorics given by greater than or equal to n generations of the hexagonal packing and whose central circle is the circle C_0 of unit radius centered at the origin. The Ring Lemma implies that the radii of the circles n generations removed from C_0 in the packings H_m for $m \geq n$ are bounded away from zero and infinity. A diagonal argument implies that there is a subsequence H_{n_i} that geometrically converges to a packing H, which necessarily has hexagonal combinatorics. But the uniqueness of the Discrete Uniformization Theorem implies that $H = \mathcal{H}_1$, the penny packing of unit radius. If the lemma were not true, one could choose the sequence H_n in such a way that the ratio of the center circle of H_n to at least one of its neighbors differs from unity by at least a fixed constant $\delta > 0$. This would imply that the limit packing H has a circle adjacent to C_0 of non-unit radius, contradicting uniqueness. □

I should mention that Rodin and Sullivan did not have access to the Discrete Uniformization Theorem in 1987 as it was published only in 1990. They had to prove uniqueness of the penny packing of the plane, which they did by invoking results of Dennis Sullivan [72] extending the Mostow Rigidity Theorem to non-compact three-manifolds whose volumes grow slowly enough. This initiated an attempt to prove the Hexagonal Packing Lemma using only elementary means, which ultimately led to a better understanding of the rigidity of infinite circle packings over the next decade. This paper of Rodin and Sullivan was highly influential and can claim to be the genesis of the serious study of circle packings that now includes in its accomplishments hundreds of articles, thousands of citations, and a huge reservoir of applications in a great variety of different settings.

5.5.2 Uniformizing Equilateral Surfaces

I already have defined piecewise equilateral metrics determined by plane triangulation graphs in the context of the type problem. Of course there is nothing special about plane triangulation graphs. Any triangulation T of a surface may be endowed with a piecewise equilateral metric by identifying faces with unit equilateral triangles. Exactly as explained in Sect. 5.4.2, this endows the surface with a complex atlas of conformal charts indexed by the vertices and edges of the triangulation. Equilateral surfaces have become important in several different areas of mathematics. They arise for example in Grothendieck's theory of dessins d'enfants and their corresponding Belyĭ maps, see [20], in Angel and Schramm's theory of uniform infinite planar triangulations [4], in Gill and Rohde's study of random planar maps [37], in Bowers and Stephenson's theory of conformal tilings and especially those that arise from expansion complexes [21, 22], and in discrete conformal flattening of surfaces in $\mathbb{R}^3$ [16]. In this section I introduce a method of uniformizing these surfaces using the tools of Rodin-Sullivan [62] and basic surface theory.

Let T be a triangulation of the topological surface S. The notation $|T|_{\text{eq}}$ is used to denote the piecewise equilateral metric space determined by the triangulation T and $\mathcal{S}_T$ to denote the Riemann surface determined by the atlas $\mathcal{A} = \{\varphi_x : x \in V(T) \cup E(T)\}$. Note that T need not be a simplicial triangulation for this to make sense. A face f of T first is identified as an equilateral triangle in $|T|_{\text{eq}}$ and then as a curvilinear triangle in the canonical metric of constant curvature on the surface $\mathcal{S}_T$. What is the shape of f in $\mathcal{S}_T$? One fact about the shape of this curvilinear triangle is that the angle that two of its sides makes that emanate from the same vertex is $2\pi/d$, where d is the degree of the vertex. Another fact is that the sides are analytic arcs, and in fact any such arc is the fixed point set of an anti-conformal reflection that exchanges the two triangles incident with that arc. In the case $\mathcal{S}_T$ is parabolic or hyperbolic, f can be lifted to the plane $\mathbb{C}$ or the Poincaré disk $\mathbb{D}$ and so this shape may be displayed as a curvilinear triangle in the plane. In case $\mathcal{S}_T$ is elliptic, this shape may be stereographically projected from the 2-sphere to the plane. How does one get at this shape? The answer Ken Stephenson and I supplied in [20] is the content of this section.

For simplicity, let's restrict our attention to closed surfaces. The scheme for approximating a uniformizing map is to use the triangulation T as a pattern for a circle packing, and then refine iteratively using so-called *hex-refinement* to obtain a sequence $\mathcal{P}_n$ of finer and finer circle packings, after an initial barycentric subdivision. Hex-refinement applied to a triangular face just adds a vertex to each existing edge and then connects the three new vertices on the three edges of the face by a 3-cycle of edges, thus subdividing the face into four smaller triangles. Thus barycentric subdivision followed by hex-refinement produces T_1, and iteration of hex-refinement then produces the sequence T_n with $\mathcal{P}_n$ the corresponding circle packing in the surface $\mathcal{S}_n$ in the pattern of T_n. There is an added layer of difficulty here in that, unlike with the use of the hexagonal packing in the Discrete Riemann

Mapping Theorem, the circle packings in this setting do not occupy the same surface. The surfaces $\mathcal{S}_n$ are determined by the triangulations T_n according to Theorem 5.3.1, and these need not be conformally equivalent to one another. Also, any face f of T with nth hex-subdivision f_n in T_n determines a sequence $\mathcal{P}_n(f)$ of circle packings, those circles in $\mathcal{P}_n$ corresponding to the vertices of f_n.

Discrete Uniformization Theorem for Equilateral Surfaces (Bowers and Stephenson [20]) *The surfaces $\mathcal{S}_n$ converge in moduli as $n \to \infty$ to a surface $\mathcal{S}$ that is conformally homeomorphic to the surface $\mathcal{S}_T$, the Riemann surface determined by the equilateral surface $|T|_{eq}$. For any face f of T, the carriers of $\mathcal{P}_n(f)$ converge geometrically to the shape of f in $\mathcal{S}_T$ when given its canonical constant curvature metric.*

The latter statement of the theorem may be understood to mean that when one lifts the carriers to the universal cover, the sphere $\mathbb{S}^2$, the plane $\mathbb{C}$, or the disk $\mathbb{D}$, and normalizes appropriately, the carriers converge in the Hausdorff metric on compacta to the appropriate lift of f in $\mathcal{S}_T$.

Sketch of Proof Note that the realizations of the triangulation T in the metric surface $|T|_{\text{eq}}$ and in the Riemann surface $\mathcal{S}_T$ are reflective, meaning that each edge e is the fixed point set of an anti-conformal reflection that exchanges the two faces contiguous to e.[11] Rather than the canonical constant curvature metric, I shall use the piecewise equilateral metric ρ_T on $\mathcal{S}_T$ throughout the proof. Here is a key observation. Hex-subdivision may be performed metrically in $\mathcal{S}_T$ by adding new vertices $v(e)$ as the mid-points of the edges $e \in E(T)$ and connecting $v(e)$ to $v(e')$ by a Euclidean straight line segment in the metric ρ_T in the face bounded by edges e, e' and e''. This realizes the hex-refined triangulation T_1 as a reflective triangulation in $\mathcal{S}_T$.[12] Iterating, T_n may be realized as a reflective triangulation of $\mathcal{S}_T$ that metrically hex-subdivides T_{n-1}.

Define homeomorphisms $h_n : \mathcal{S}_T \to \mathcal{S}_n$ so that the image of vertex v of T_n under h_n is the center of the circle that corresponds to v in the circle packing $\mathcal{P}_n$, extend linearly along edges and then with minimum quasiconformal distortion across faces. By the Ring Lemma, each mapping h_n is quasiconformal, and since hex-refinement does not increase degree, any bound ≥ 6 on the degrees of the vertices of T also bounds the degrees of the vertices of T_n, for all $n \geq 1$. This implies that the homeomorphisms h_n have uniformly bounded dilatations, and this implies that a subsequence of the surfaces $\mathcal{S}_n$ converges in moduli to a Riemann surface $\mathcal{S}$.

My claim is that $\mathcal{S}$ is conformally equivalent to $\mathcal{S}_T$. This would be confirmed were the maximum dilatations of the homeomorphisms h_n shown to limit to unity as $n \to \infty$, but unfortunately this does not occur. In fact these dilatations are bounded away from unity with large dilatations concentrated near the original vertices of T.

[11] To be clear, the reflection is anti-conformal on the *interior* of the union of the two faces incident at e, but not at the vertices.

[12] Technically, this is after the initial barycentric subdivision, which also is performed in the metric ρ_T and yields a reflective triangulation.

To get around this, let D be a compact domain in $\mathbb{S}_T$ disjoint from the vertex set $V(T)$. Note that the combinatorics of T_n away from the vertices of T is hexagonal, and this implies that as $n \to \infty$, the compact set D is surrounded by a number of generations of the hexagonal combinatorics that increases without bound. The Hexagonal Packing Lemma applies to confirm that the maximum dilatations of the restrictions of the homeomorphisms h_n to D converge to unity. This works for every compact domain that misses the vertex set $V(T)$, and this implies that the limit mapping $h : \mathbb{S}_T \to \mathbb{S}$ is conformal on the complement of the vertex set $V(T)$. Now the removability of isolated singularities comes into play and implies that the homeomorphism h is conformal at the vertices, and so is a conformal homeomorphism of $\mathbb{S}_T$ onto $\mathbb{S}$. □

Figure 5.7 shows an example of an approximation to a portion of an equilateral surface uniformized in the plane. In this figure each edge is the fixed point set of an anti-conformal reflection that exchanges the grey-white pair of triangles sharing that edge. This is an approximation of the conformally correct shapes of the equilateral triangles forming the equilateral surface being imaged.

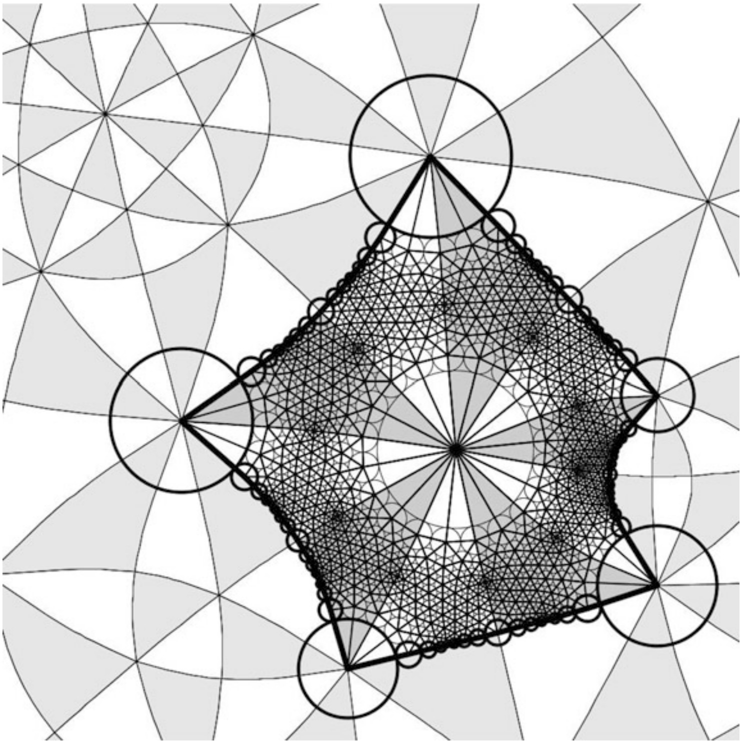


Fig. 5.7 Conformal shapes of equilateral triangles in a planar equilateral surface approximated with the circle packing of the twice hex-refined barycentric subdivision of the original triangulation

5.6 Inversive Distance Circle Packings

Around 2001, Ken Stephenson and I began thinking about *inversive distance circle packings* and how they could be used to uniformize piecewise flat surfaces, those surfaces in which each face is identified with a flat Euclidean triangle, not necessarily equilateral. There is a tentative discussion of this in [20] and further discussion in [16] of the difficulties in proving convergence of discrete mappings to the uniformization mapping, though the method does seem to work well in practice; again see [16]. The first theoretical questions concern (1) the existence of circle packings with prescribed inversive distances between adjacent circles and (2) the rigidity and uniqueness of these packings.

When all inversive distances lie in the unit interval, adjacent circles overlap with specified angle $0 \leq \theta \leq \pi/2$. This is covered by the Koebe–Andre'ev–Thurston Theorems. When inversive distance is greater than unity, the circles do not overlap and the inversive distance is a Möbius-invariant measure of how separated the circles are. In this case Problems (1) and (2) seem much more difficult to approach. Problem (1) is especially difficult in that there are local assignments of inversive distances that must be avoided as there are no circle configurations that realize those distances. These are difficult to catalogue, but even if there are no local obstructions to the existence of a packing, it is not at all clear whether still there may be global obstructions. Little progress has been made on Problem (1), but the situation for Problem (2) has enjoyed some progress, initially in 2011 and more recently in the past couple of years. It is these recent successes in approaching Problem (2) that occupies this section. My contention is that a change of viewpoint can be effective in approaching inversive distance circle packings, and a hint as to how to proceed comes from the classical rigidity theory of bar-and-joint linkages. After a brief review of inversive distance, I will explore this new framework for circle packings and discuss some recent successes.

5.6.1 A Quick Introduction to Inversive Distance

There are a number of ways to define the inversive distance between two circles in the Riemann sphere. I will present several of these below, starting with the most mundane that gives a Euclidean formula for the inversive distance between two planar circles.[13] Let C_1 and C_2 be distinct circles in the complex plane $\mathbb{C}$ centered at the respective points p_1 and p_2, of respective radii r_1 and r_2, and bounding the respective *companion disks* D_1 and D_2.

[13]This easily can be extended to the inversive distance between a circle and a line, or two lines. I will forgo this development since the next definition is completely general.

Definition (Inversive Distance in the Euclidean Metric) The *inversive distance* $\langle C_1, C_2\rangle$ between C_1 and C_2 is

$$\langle C_1, C_2\rangle = \frac{|p_1 - p_2|^2 - r_1^2 - r_2^2}{2r_1 r_2}. \tag{5.6.1}$$

The *absolute inversive distance* between distinct circles is the absolute value of the inversive distance.

The absolute inversive distance is a Möbius invariant of the placement of two circles in the plane. This means that there is a Möbius transformation of $\mathbb{C}$ taking one circle pair to another if and only if the absolute inversive distances of the two pairs agree. The important geometric facts that make the inversive distance useful in inversive geometry and circle packing are as follows. When $\langle C_1, C_2\rangle > 1$, $D_1 \cap D_2 = \emptyset$ and $\langle C_1, C_2\rangle = \cosh\delta$, where δ is the hyperbolic distance between the totally geodesic hyperbolic planes in the upper-half-space model $\mathbb{C} \times (0, \infty)$ of $\mathbb{H}^3$ whose ideal boundaries are C_1 and C_2. When $\langle C_1, C_2\rangle = 1$, D_1 and D_2 are tangent at their single point of intersection. When $1 > \langle C_1, C_2\rangle \geq 0$, D_1 and D_2 overlap with angle $0 < \theta \leq \pi/2$ with $\langle C_1, C_2\rangle = \cos\theta$. In particular, $\langle C_1, C_2\rangle = 0$ precisely when $\theta = \pi/2$. When $\langle C_1, C_2\rangle < 0$, then D_1 and D_2 overlap by an angle greater than $\pi/2$. This includes the case where one of D_1 or D_2 is contained in the other, this when $\langle C_1, C_2\rangle \leq -1$. In fact, when $\langle C_1, C_2\rangle < -1$ then $\langle C_1, C_2\rangle = -\cosh\delta$ where δ has the same meaning as above, and when $\langle C_1, C_2\rangle = -1$ then C_1 and C_2 are 'internally' tangent. When $-1 < \langle C_1, C_2\rangle < 0$, then the overlap angle of D_1 and D_2 satisfies $\pi > \theta > \pi/2$ and again $\langle C_1, C_2\rangle = \cos\theta$.

The more general definition measures the inversive distance between oriented circles. Note that an oriented circle determines a unique closed *companion* or *spanning disk* that the circle bounds. Indeed, assuming fixed orientations for $\mathbb{S}^2$ and $\widehat{\mathbb{C}}$ that are compatible via stereographic projection, the companion disk determined by the oriented circle C is the closed complementary disk D (of the two available) whose positively oriented boundary $\partial^+ D = C$, where of course the orientation of D is inherited from that of $\mathbb{S}^2$ or $\widehat{\mathbb{C}}$. This is described colloquially by saying that D lies to the left of C as one traverses C along the direction of its orientation.

Definition (General Inversive Distance) Let C_1 and C_2 be oriented circles in the extended plane $\widehat{\mathbb{C}}$ bounding their respective companion disks D_1 and D_2, and let C be any oriented circle mutually orthogonal to C_1 and C_2. Denote the points of intersection of C with C_1 as z_1, z_2 ordered so that the oriented sub-arc of C from z_1 to z_2 lies in the disk D_1. Similarly denote the ordered points of intersection of C with D_2 as w_1, w_2. The *general inversive distance* between C_1 and C_2, denoted as $\langle C_1, C_2\rangle$, is defined in terms of the cross ratio

$$[z_1, z_2; w_1, w_2] = \frac{(z_1 - w_1)(z_2 - w_2)}{(z_1 - z_2)(w_1 - w_2)}$$

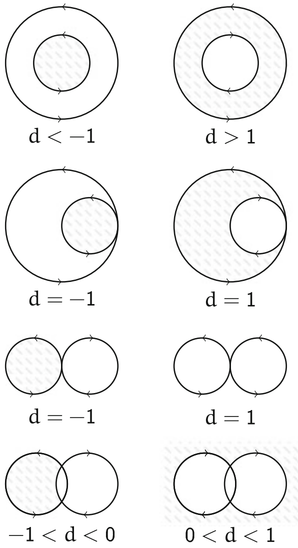


Fig. 5.8 Inversive distances $d = \langle C_1, C_2\rangle$. The shaded regions are the intersections $D_1 \cap D_2$, the points common to the spanning disks D_1 and D_2 for both circles C_1 and C_2

by

$$\langle C_1, C_2\rangle = 2[z_1, z_2; w_1, w_2] - 1.$$

Subsequently, I'll drop the adjective *general* and refer to the inversive distance $\langle C_1, C_2\rangle$ with its absolute value $|\langle C_1, C_2\rangle|$ the *absolute inversive distance*.[14]

Recall that cross ratios of ordered 4-tuples of points in $\widehat{\mathbb{C}}$ are invariant under Möbius transformations and that there is a Möbius transformation taking an ordered set of four points of $\widehat{\mathbb{C}}$ to another ordered set of four if and only if the cross ratios of the sets agree. This implies that which circle C orthogonal to both C_1 and C_2 is used in the definition is irrelevant as a Möbius transformation that set-wise fixes C_1 and C_2 can be used to move any one orthogonal circle to another. Which one of the two orientations on the orthogonal circle C is used is irrelevant as the cross ratio satisfies $[z_1, z_2; w_1, w_2] = [z_2, z_1; w_2, w_1]$. This equation also shows that the inversive distance is preserved when the orientation of both circles is reversed so that it is only the relative orientation of the two circles that is important for the definition. In fact, the general inversive distance is a relative conformal measure of the placement of an oriented circle pair on the Riemann sphere. By this I mean that two oriented circle pairs are inversive equivalent if and only if their inversive distances agree. All of this should cause one to pause to develop some intuition about how companion disks may overlap with various values of inversive distances. See Fig. 5.8 for some

[14]The author first learned of defining inversive distance in this way from his student, Roger Vogeler. He has looked for this in the literature and, unable to find it, can only surmise that it is original with Prof. Vogeler. The definition appeared in [16] in 2003.

corrections to possible misconceptions. Finally, the inversive distance is symmetric with $\langle C_1, C_2\rangle = \langle C_2, C_1\rangle$ since $[z_1, z_2; w_1, w_2] = [w_1, w_2; z_1, z_2]$.

The inversive distance is real since the cross ratio of points lying on a common circle is real and, in fact, every real value is realized as the inversive distance of some oriented circle pair. Notice that if the orientation of only one member of a circle pair is reversed, the inversive distance merely changes sign. This follows from the immediate relation $[z_1, z_2; w_2, w_1] = 1 - [z_1, z_2; w_1, w_2]$. Despite its name, the inversive distance is not a metric as it fails to be non-negative and fails to satisfy the triangle inequality.[15]

The third definition is entirely in terms of the spherical metric.

Definition (Inversive Distance in the Spherical Metric) In the 2-sphere $\mathbb{S}^2$, the inversive distance may be expressed as

$$\langle C_1, C_2\rangle = \frac{-\cos\sphericalangle(p_1, p_2) + \cos(r_1)\cos(r_2)}{\sin(r_1)\sin(r_2)} = \frac{-p_1\cdot p_2 + \cos(r_1)\cos(r_2)}{\sin(r_1)\sin(r_2)}. \tag{5.6.2}$$

Here, $\sphericalangle(p_1, p_2) = \cos^{-1}(p_1 \cdot p_2)$ denotes the spherical distance between the centers, p_1 and p_2, of the respective companion disks, $p_1 \cdot p_2$ the usual Euclidean inner product between the unit vectors p_1 and p_2, and r_1 and r_2 the respective spherical radii of the companion disks. Note that $r_i = \cos^{-1}(p_i \cdot q_i)$ for any point q_i on the circle C_i, for $i = 1, 2$.

Verifying the equivalence of this with the general definition is an exercise in the use of trigonometric identities after a standard placement of C_1 and C_2 on $\mathbb{S}^2$ followed by stereographic projection. This standard placement is obtained by finding the unique great circle C orthogonal to both C_1 and C_2 and then rotating the sphere so that this great circle is the equator, which then stereographically projects to the unit circle in the complex plane. The details are left to the reader.

Here are two more quick descriptions of inversive distance. For those conversant with the representation of circles in $\mathbb{S}^2$ by vectors in de Sitter space, the inversive distance is the Minkowski inner product between the two points of de Sitter space that represent the two oriented circles. This is, perhaps, the most elegant formulation of the product. The final way I'll describe the inversive distance is a neat little curiosity. Let $C_1 = \partial\mathbb{D} = \mathbb{S}^1$ be the unit circle oriented clockwise and C_2 a circle oriented counterclockwise that meets the open unit disk non-trivially. Then, as explained on page 193, the intersection c_2 of C_2 with the open disk is a curve of constant geodetic curvature in the Poincaré disk $\mathbb{D} \cong \mathbb{H}^2$. The inversive distance is $\langle C_1, C_2\rangle = \mathrm{curv}(c_2)$, the geodetic curvature of the cycle c_2 in the Poincaré metric on $\mathbb{D}$. This includes all three cases for the cycle c_2—a hyperbolic circle in $\mathbb{D}$, a

[15] Some authors, perhaps more aptly, call the inversive *distance* the inversive *product* of C_1 and C_2.

horocycle that meets $\partial\mathbb{D}$ at a single point, or a hypercycle that meets $\partial\mathbb{D}$ at two points.[16]

5.6.2 Some Advances on the Rigidity Question

In [20], *inversive distance circle packings* were introduced. Rather than preassigned overlap angles labeling edges of a triangulation of a surface as in the Koebe–Andre'ev–Thurston Theorems, preassigned inversive distances label the edges. As stated already, questions of interest are of the existence and uniqueness of circle configurations in geometric structures on surfaces that realize the inversive distance data. Though the existence question is wide open, in 2011–2012 there were three advances on the uniqueness question for inversive distance packings. First, Guo [41] proved that inversive distance packings of closed surfaces of positive genus, ones supporting flat or hyperbolic metrics, are locally rigid whenever the inversive distances are non-negative. Shortly after that, Luo [52] improved this to global rigidity, or uniqueness of the packings in the cases considered by Guo. Then in a surprising result of the year following, Ma and Schlenker [53] produced a counterexample to global uniqueness for packings of the 2-sphere. They gave examples of pairs of circle packings of $\mathbb{S}^2$ in the pattern of the octahedral triangulation with six circles that satisfy the same inversive distance data, but that are not Möbius equivalent.

The ingredients of Ma and Schlenker's example are Schönhardt's twisted octahedron, which is an infinitesimally flexible polyhedron in Euclidean space $\mathbb{E}^3$, embeddings in de Sitter space $\mathbb{S}^3_1$, and special properties of the Pogorelov map between different geometries. In 2017, John Bowers and I [15] constructed a large family of Ma–Schlenker–like examples using only inversive geometry, producing many counterexamples to the uniqueness of inversive distance circle packings in the 2-sphere.

The Schönhardt octahedron is an example of a bar-and-joint linkage important in the rigidity theory of Euclidean frameworks, and its use in the Ma–Schlenker example hinted at a way forward in understanding the rigidity theory of inversive distance circle packings in the 2-sphere. This led to a fruitful change in viewpoint and a reformatting of the question of uniqueness of inversive distance circle packings to the question of the rigidity—local, global, and infinitesimal—of more general *circle frameworks*. These are analogues in Möbius geometry of the Euclidean frameworks in Euclidean geometry with point configurations in $\mathbb{E}^3$ replaced by circle configurations in $\mathbb{S}^2$ and the Euclidean metric replaced by the non-metric inversive distance. The analogy is not exact, but the theory of linkages in $\mathbb{E}^3$ has been found to be a good guide for understanding some of the rigidity theory of circle

[16]My student, Opal Graham, noticed, then proved this when I was lecturing on the curves of constant geodetic curvature in the hyperbolic plane.

frameworks. Part of why this works so well is because the space of circles in the 2-sphere is a three-dimensional incidence geometry that has much in common with the space of points in Euclidean 3-space. The lines of this geometry are coaxial circle families and the planes are what Carathéodory in [28] called bundles of circles. This allows one to define what is meant by a convex collection of circles, planar collections of circles, circle polyhedra, bounded circle configurations, etc. Space constraints in this chapter interfere with even a cursory account of these issues, so I am content with listing a couple of recent successes of the theory without all the definitions needed for a precise understanding, and then taking some time to set up the language of this change of viewpoint.

The two theorems following are the result, both the statements and the proofs, of an engagement between circle packing theory and the classical rigidity theory of Euclidean frameworks in $\mathbb{E}^3$.

Theorem 5.6.1 (Bowers et al. [23]) *Let $\mathcal{C}$ and $\mathcal{C}'$ be two non-unitary, inversive distance circle packings with ortho-circles for the same oriented edge-labeled triangulation of the 2-sphere $\mathbb{S}^2$. If $\mathcal{C}$ and $\mathcal{C}'$ are convex and proper, then there is a Möbius transformation $T : \mathbb{S}^2 \to \mathbb{S}^2$ such that $T(\mathcal{C}) = \mathcal{C}'$.*

The *edge-label* refers to prescribed inversive distances labeling each edge. *Non-unitary* means that the inversive distance between any pair of adjacent circles is not unity; in fact, these inversive distances are in the set $(-1, 1) \cup (1, \infty)$. Having *ortho-circles* means that each triple of mutually adjacent circles have an orthogonal circle. This generalizes to a global rigidity theorem about *circle polyhedra*, circle configurations in the pattern of three-dimensional polyhedra whose faces correspond to circle configurations that are planar in the incidence geometry of circle space; see [23] for details.

Theorem 5.6.2 (Bowers et al. [23]) *Any two convex and proper non-unitary circle polyhedra with Möbius-congruent faces that are based on the same oriented abstract spherical polyhedron and are consistently oriented are Möbius-congruent.*

Theorem 5.6.1 coupled with the Ma–Schlenker example of [53] and the examples of [15] show that the uniqueness of inversive distance circle packings, and more generally, of circle polyhedra is exactly analogous to that of Euclidean polyhedra—convex and bounded polyhedra in $\mathbb{E}^3$ are prescribed uniquely by their edge lengths and face angles whereas non-convex or unbounded polyhedra are not. The proof of this for convex and bounded Euclidean polyhedra is Cauchy's celebrated rigidity theorem [29], which is reviewed in Sect. 5.7.5. The Proof of Theorem 5.6.2 follows Cauchy's original argument, which splits the proof into two components—a combinatorial lemma and a geometric lemma. Cauchy's combinatorial lemma deals with a certain labeling of the edges of any graph on a sphere, and applies to the present setting. The geometric lemma, known as *Cauchy's Arm Lemma*, requires that a polygon with certain properties be defined for each vertex of the polyhedron, and fails to apply here. The main work of the proof is in describing and analyzing a family of hyperbolic polygons called *green-black polygons* that are defined for each vertex of a circle polyhedron in a Möbius-invariant manner. An analogue of

Cauchy's Arm Lemma for convex green-black polygons is developed and used to prove these theorems.

5.6.3 *Circle Frameworks and Möbius Rigidity*

I'll close out this section with a description of the change in viewpoint from circle packings to circle frameworks. This can be done using only absolute inversive distance, but I find it advantageous to remain as general as possible in setting up the viewpoint. The goal is to generalize the language of circle packings and patterns of triangulations and quadrangulations of the 2-sphere to that of circle realizations of oriented circle frameworks. Let G be a graph, by which I mean a set of vertices $V = V(G)$ and simple edges $E = E(G)$. Both loops and multiple edges are disallowed. An oriented edge incident to the initial vertex u and terminal vertex v is denoted as uv, and $-uv$ means the oppositely oriented edge vu. I will use the same notation, uv, to denote an un-oriented edge, context making the meaning clear. A *circle framework with adjacency graph* G, or *c-framework* for short, is a collection $\mathcal{C} = \{C_u : u \in V(G)\}$ of oriented circles in $\mathbb{S}^2$ indexed by the vertex set of G. This is denoted by $G(\mathcal{C})$. Two c-frameworks $G(\mathcal{C})$ and $G(\mathcal{C}')$ are *equivalent* if $\langle C_u, C_v\rangle = \langle C'_u, C'_v\rangle$ whenever uv is an edge of G. Let H be a subgroup of the inversive group Inv($\mathbb{S}^2$) of the 2-sphere. Two collections $\mathcal{C}$ and $\mathcal{C}'$ of oriented circles indexed by the same set are *H-equivalent* or *H-congruent* provided there is a mapping $T \in H$ such that $T(\mathcal{C}) = \mathcal{C}'$, respecting the common indexing and the orientations of the circles. When H is not so important they are *inversive-equivalent* or *inversive-congruent*, and when T can be chosen to be a Möbius transformation, they are *Möbius-equivalent* or *Möbius-congruent*. The global rigidity theory of c-frameworks concerns conditions on G or $G(\mathcal{C})$ that ensure that the equivalence of the c-frameworks $G(\mathcal{C})$ and $G(\mathcal{C}')$ guarantees their H-equivalence. Often one restricts attention to c-frameworks in a restricted collection $\mathscr{F}$ of c-frameworks. In Theorem 5.6.2, $\mathscr{F}$ is the collection of non-unitary, convex and proper c-polyhedra and the interest is in Möbius equivalence.

Definition (Labeled Graph and Circle Realization) An *edge-label* is a real-valued function $\beta : E(G) \to \mathbb{R}$ defined on the edge set of G, and G together with an edge-label β is denoted as G_β and called an *edge-labeled graph*. The c-framework $G(\mathcal{C})$ is a *circle realization* of the edge-labeled graph G_β provided $\langle C_u, C_v\rangle = \beta(uv)$ for every edge uv of G, which henceforth is denoted as $G_\beta(\mathcal{C})$. See Fig. 5.9.

Circle packings are circle realizations of edge-labeled graphs that arise as the 1-skeletons of oriented triangulations of the 2-sphere that also satisfy certain properties that ensure that the realizations of the triangular boundaries of faces respect orientation. The general definition allows for branch vertices and configurations of circles in which the open geodesic triangles cut out by connecting centers of adjacent circles overlap. There are subtleties in which I have no interest, so I am

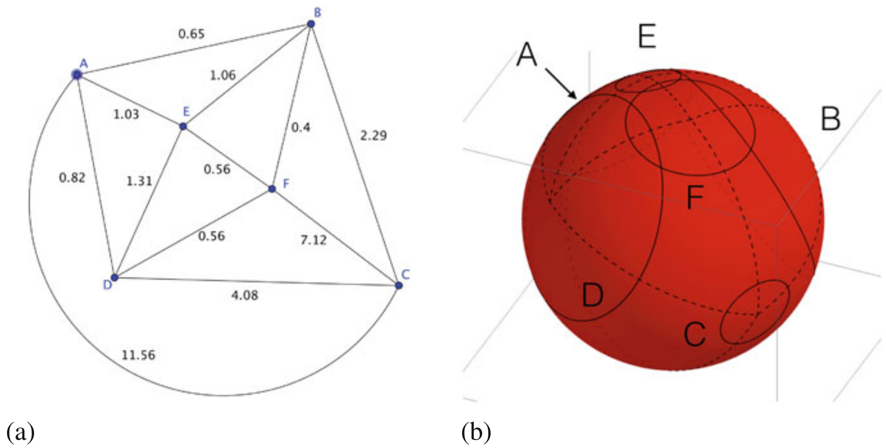


Fig. 5.9 An edge-labeled octahedral graph and its circle realization. The labels are proposed inversive distances between the circles corresponding to the vertices. (**a**) An edge-labeled octahedral graph $\mathcal{O}_\beta$. Labels <1 imply overlapping circles, >1 separated ones. (**b**) A corresponding c-framework realizing $\mathcal{O}_\beta$. Circle A is hidden on the back side of the sphere

going to adapt a restricted definition that corresponds to the circle packings that arise from spherical polyhedral metrics on triangulated surfaces. These are circle realizations of the edge-labeled 1-skeleton $G_\beta = K_\beta^{(1)}$ of an oriented triangulation K of $\mathbb{S}^2$ that produce oriented geodesic triangulations[17] of the 2-sphere when adjacent circle centers are connected by geodesic arcs. The assumption here is that the centers of no two adjacent circles are antipodal, so that there is a unique geodesic arc connecting them, and that the centers of three circles corresponding to the vertices of a face of K do not lie on a great circle. Now this causes no particular problems when all adjacent circles overlap nontrivially, the traditional playing field of circle packing, but does cause some real concern when adjacent circles may have inversive distance greater than unity. For example, a circle realization may produce a geodesic triangulation of the sphere by connecting adjacent centers while its Möbius image may not. This is traced directly to the fact that neither circle centers nor radii, nor geodesic arcs, are Möbius invariants in the inversive geometry of the sphere. This behavior does not occur for inversive distance circle packings of the Euclidean or hyperbolic planes (and surfaces), precisely because circle centers and geodesics are invariant under automorphisms and radii are invariant up to scale in Euclidean geometry and invariant in hyperbolic geometry. My belief is that using centers and radii of circles in inversive geometry should be avoided except where these can be used to simplify computations (as in the use of the spherical definition of inversive distance). The shift then is from inversive distance circle packings to inversive

[17]By this I mean that the orientation of the geodesic triangulation determined by the packing is consistent with the orientation on K.

distance circle realizations. One is less concerned with possible underlying geodesic triangulations and more concerned with Möbius-invariant quantities. For example, rather than working with a geodesic face formed by connecting the centers of three mutually adjacent circles, one is more interested in the existence of an ortho-circle, a circle mutually orthogonal to the three, which is a Möbius invariant. Though the initial motivation was circle packing as reflected in Theorem 5.6.1, the real interest has evolved to circle realizations as reflected in the more general version represented by Theorem 5.6.2.

It turns out that Theorem 5.6.2 has implications for the rigidity of generalized hyperbolic polyhedra in $\mathbb{H}^3$. Thurston was the first to exploit this connection between circle configurations on $\mathbb{S}^2$ and hyperbolic polyhedra in $\mathbb{H}^3$ in really significant ways, and his observations inspired several avenues of clarification and generalization. It is to this that I turn in the penultimate section of this chapter.

5.7 Polyhedra—From Steiner (1832) to Rivin (1996), and Beyond

In this section I survey the rich mathematical vein that has been mined in the geometric theory of polyhedra, particularly of three-dimensional hyperbolic polyhedra, that has its origins in Thurston's insights on using his circle packing theorem to characterize certain hyperbolic polyhedra. The initial observation of Thurston was that the study of polyhedra in hyperbolic three-space can be transferred to the study of overlapping circle packings in the two-sphere by realizing the Riemann sphere as the boundary of the Beltrami–Klein model of $\mathbb{H}^3$ sitting as the unit ball $\mathbb{B}^3$ in the real projective three-space. Theorems in one of these venues correspond to theorems in the other. Later Thurston's students, Oded Schramm and Igor Rivin, made great strides in the theory of both three-dimensional Euclidean and hyperbolic polyhedra, not so much using the techniques of circle packing but instead using very intricate and clever geometric arguments, often times in this classical setting of $\mathbb{H}^3 \cong \mathbb{B}^3 \subset \mathbb{E}^3 \subset \mathbb{RP}^3$. There is here a beautiful interplay among the classical geometries illustrating Arthur Cayley's aphorism that "All geometry is projective geometry." Here one sees the Beltrami–Klein model of hyperbolic three-space as a sub-geometry of the real projective three-space, with its orientation-preserving isometry group naturally identified with the Lorentz group of Minkowski space-time, which itself restricts to the two-sphere boundary of hyperbolic space as the group of circle-preserving transformations of the two-sphere, the group of Möbius transformations. This one geometry, the real projective geometry of dimension three, presents a playing field for studying three-dimensional polyhedra—classical Euclidean polyhedra, hyperbolic polyhedra of various types and generalizations, projective polyhedra, and circle polyhedra of Möbius geometry.

I will begin with an application of Thurston's circle packing theorem on using polyhedra to cage a sphere, and move then to Schramm's generalization. From there

I will discuss the characterization of certain hyperbolic polyhedra—compact by Hodgson and Rivin, ideal by Rivin, and hyper-ideal by Bao and Bonahon—and will finish with very recent work by Chen and Schlenker that characterizes those convex projective polyhedra all of whose vertices lie on the ideal boundary of hyperbolic space. I include a bonus final section on Cauchy's 1813 Rigidity Theorem for the reader who is approaching this subject as a novice. This is the fundamental theorem of rigidity theory, and the techniques and tools Cauchy developed have been used time and again in proofs of rigidity in the past 200 years. Both Schramm and Rivin make use of Cauchy's toolbox in their theorems on convex hyperbolic and Euclidean polyhedra, as do Bao and Bonahon as well as Bowers, Pratt and the author. Before these recent developments, previous generations of mathematicians who delved into the study of polyhedra made use of Cauchy's toolbox—Dehn in his proof of infinitesimal rigidity, Aleksandrov in his rigidity results, Gluck in his examination of generic rigidity, and Connelly in various of his contributions.

5.7.1 Caging Eggs—Thurston and Schramm

In 1832, Jakob Steiner [69] asked

> In which cases does a convex polyhedron have a combinatorial equivalent which is inscribed in, or circumscribed about, a sphere?

When a convex polyhedron P is *inscribed* in the sphere S so that its vertices lie on S, then its polar dual *circumscribes* the sphere S, so that each face of the dual P^* meets S in a single point. It wasn't until 1928 that Ernst Steinitz found families of non-inscribable polyhedral types with the example of a cube truncated at one vertex being the simplest. Marcel Berger [9, p. 532] takes this long duration of time between Steiner and Steinitz as evidence that the theory of polyhedra in the years intervening had fallen into disrespect among mathematicians, being a subject of the old-fashioned mathematics of synthetic geometry.[18] One would be hard pressed to say that the study of polyhedra in the time between Steinitz and Thurston was anything but a curiosity to many a mathematician schooled in the rarified heights of abstraction that had captured the mathematical mind of the time. The sort of "pedestrian geometry" offered by the study of polyhedra captured the imagination of a select few. There has been a healthy development of the rigidity theory of polyhedra, notably by Aleksandrov in the 1950s, and Gluck and Connelly in the 1970s. Aleksandrov's work was largely ignored in the West until the 1980s. Coxeter had done truly foundational work in the combinatorial structure of polyhedra in the 1940s and 1950s, and Victor Klee and Branko Grünbaum began their foundational studies a bit later. Coxeter's work in geometry was routinely dismissed by much of mainstream mathematics as old-fashioned nineteenth century

[18] Berger [9] uses the word *disdain* to describe the prevailing opinion of the study of polyhedra.

mathematics, uninteresting and pedestrian. Both Aleksandrov and Coxeter were "rehabilitated" by the larger community of geometers and topologists when their work of the forties and fifties—Aleksandrov's on metric geometry and Coxeter's on reflection groups—became important to the development of geometric group theory after Gromov's publication of his hyperbolic groups essay [39] in 1987. With apologies to Aleksandrov, Coxeter, Klee, and Grünbaum, it has taken the attention of Thurston and his students Schramm and especially Rivin to resurrect more intense interest among topologists in this venerable old subject of classical geometry.[19]

Steinitz's basic tool for attacking the Steiner question is the following observation. Suppose the polyhedron P circumscribes the sphere S. Let $e = uv$ be an edge of P with adjacent faces f and g. Since P circumscribes S, the face f is tangent to S at a point p and g is tangent at a point q. Then the angle $\angle upv = \angle uqv$ in measure and we let $\Theta(e)$ denote this common value. It is immediate that summing these edge labels for the edges of any face yields an angle sum of 2π. The reader might want to use this observation to see why a dodecahedron truncated at every vertex admits no inscribed sphere as there is no edge labeling Θ for this polyhedron that satisfies this property.

According to Steinitz then, the condition that an edge label $\Theta : E(P) \to (0, \pi)$ exists for the polyhedron P whose sum for the edges of each face is 2π is a necessary condition that P have a combinatorially equivalent realization that circumscribes a sphere, but it is not sufficient. It was not until Rivin's study of hyperbolic polyhedra in the late 1980s and early—1990s that a characterization of polyhedra of *circumscribable type*, ones combinatorially equivalent to polyhedra that may circumscribe a sphere, was found. The definitive result is due to Rivin and reported in Hodgson, Rivin, and Smith [47], and follows from his characterization of ideal convex hyperbolic polyhedra that is presented in a later section.

Circumscribable Type Characterization (Rivin) *A polyhedron P is of circumscribable type if and only if there exists a label $\Theta : E(P) \to (0, \pi)$ such that the sum of the labels $\Theta(e)$ as e ranges over any circuit bounding a face is 2π, while the sum as e ranges over any simple circuit not bounding a face is strictly greater than 2π.*

A polyhedron is of *inscribable type* if it is combinatorially equivalent to one that may be inscribed in a sphere.

Incsribable Type Characterization (Rivin) *A polyhedron P is of inscribable type if and only if its dual P^* is of circumscribable type.*

The proofs will be discussed later, but first I want to generalize this discussion a bit. Inscription and circumscription are the respective cases, $m = 0$ and $m = d - 1$, of the question of whether a d-dimensional convex polytope has a realization in

---

[19] Grünbaum [40] addresses the disinterest of the mathematical community in the combinatorial theory of polytopes in the preface to his book.

$\mathbb{E}^d$ each of whose m-dimensional faces meets a fixed $(d-1)$-dimensional sphere in a single point. One says that the polytope is (m, d)-*scribable* in this case. Egon Shulte [68] proved in the mid-1980s that when $0 \le m < d$ and $d > 2$, then there are combinatorial types of d-dimensional polytopes that are not (m, d)-scribable, except for the single exceptional case when $(m, d) = (1, 3)$. The exceptional case then is when a convex polyhedron in $\mathbb{E}^3$ *midscribes* a sphere S, so that each edge of P is tangent to S, meeting S in exactly one point.

In light of Shulte's result it perhaps is surprising that in his exceptional case, every convex polyhedron in $\mathbb{E}^3$ has a combinatorially equivalent realization that is midscribable about, say, the unit sphere $\mathbb{S}^2$. Thurston in Chapter 13 of GTTM states that this is a consequence of Andre'ev's theorems in [2, 3]. The proof I give merely applies the Koebe–Andre'ev–Thurston Theorem to an appropriately edge-labeled graph.

Midscribability of Convex Polyhedra (Thurston [73]) *Every convex polyhedron in $\mathbb{E}^3$ has a combinatorially equivalent realization that is midscribable about the unit sphere $\mathbb{S}^2$. Considering $\mathbb{E}^3 \subset \mathbb{RP}^3$, any such realization is unique up to projective transformations of $\mathbb{RP}^3$ that set-wise fix the unit sphere $\mathbb{S}^2$.*[20]

Proof Let P be a convex polyhedron in $\mathbb{E}^3$ and let K be the simplicial 2-complex obtained by adding a vertex to each open face of P and starring to the vertices. Precisely, the vertices of K are those of P along with a new vertex v_f for each face f of P. The edges are the edges of P along with edges of the form vv_f, where v is a vertex of f. The faces are the 2-simplices of the form uvv_f where uv is an edge of f. Write the edge-set of K as $E(K) = E(P) \cup E'$, where E' are the new edges of the form vv_f. Define an angle map $\Phi : E(K) \to [0, \pi/2]$ by $\Phi(e) = 0$ when $e \in E(P)$ and $\Phi(e) = \pi/2$ when $e \in E'$. An application of the Koebe–Andre'ev–Thurston Theorem I produces a circle packing $K(\mathcal{C})$ on the 2-sphere $\mathbb{S}^2$ and a geodesic triangulation in the pattern of K with overlap angles of adjacent circles given by Φ. For each face f of P, let H_f be the half-space in $\mathbb{E}^3$ that meets all the circles of $K(\mathcal{C})$ and whose bounding plane ∂H_f contains C_{v_f}. My claim is that the convex polyhedron $Q = \cap_{f \in F(P)} H_f$ midscribes $\mathbb{S}^2$ and is combinatorially equivalent to P.

For any vertex v of P, let v^* be the apex of the cone in $\mathbb{E}^3$ that is tangent to $\mathbb{S}^2$ along the circle C_v, and when $e = uv$ is an edge of P, let $e^* = u^*v^*$ be the segment with endpoints u^* and v^*. Let f be a face of the polyhedron P with vertices $v_1, \dots, v_n$ written in cyclic order. Since the circle C_{v_f} is orthogonal to the circles C_{v_i}, the apexes v_i^* all lie on the bounding plane ∂H_f, for $i = 1, \dots, n$. Let f^* denote the convex hull of the points $v_1^*, \dots, v_n^*$ in ∂H_f. A moment's thought should convince the reader that the convex polyhedron Q may be described as the convex hull of the set $V(Q) = \{v^* : v \in V(P)\}$. It follows that the vertex set of Q is $V(Q)$, edge set is $E(Q) = \{e^* : e \in E(P)\}$, and face set is $F(Q) = \{f^* : f \in F(P)\}$. This verifies that P and Q are combinatorially equivalent. Moreover, the edge $e^* = u^*v^*$

[20] The projective transformations that fix $\mathbb{S}^2$ act as Möbius transformations on $\mathbb{S}^2$.

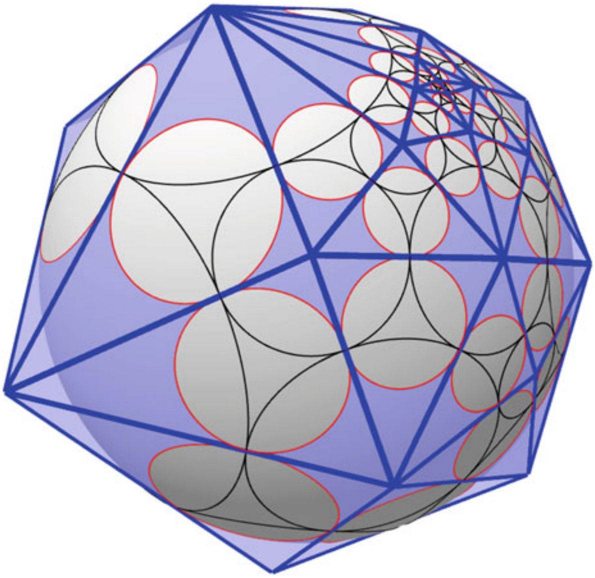


Fig. 5.10 A midscribed polyhedron. Each face meets the sphere $\mathbb{S}^2$ in a red circle and each vertex is the cone point of a black circle. Each edge e^* meets $\mathbb{S}^2$ in exactly one point, at the intersection of the two red circles determined by the faces incident to e^*, or at the intersection of the two black circles determined by the endpoints of e^*

is tangent to the sphere $\mathbb{S}^2$ at the point of intersection of the circles C_u and C_v, which are tangent since $\Phi(e) = 0$. Hence Q midscribes the sphere $\mathbb{S}^2$. See Fig. 5.10.

Uniqueness up to projective transformations that fix the unit sphere follows from the Möbius uniqueness of the circle packing $K(\mathcal{C})$ with edge angle data Φ and the fact that the Möbius group extends its action on $\mathbb{S}^2$ to a projective action of $\mathbb{RP}^3$ set-wise fixing $\mathbb{S}^2$. □

Schulte introduced in [68] the question of whether the sphere can be replaced by other convex bodies. Schramm [64] proved that when the convex polyhedron P is simplicial, then for any smooth convex body S, a combinatorially equivalent polyhedron Q exists that midscribes S. Of course this means that each edge of Q is tangent to the boundary ∂S. Shortly thereafter, Schramm improved his result by removing the requirement that P be simplicial. A convex body S is *strictly convex* if its boundary contains no non-degenerate line segment, and is *smooth* if each point of the boundary has a unique supporting plane. This latter condition is equivalent to the boundary being C^1-smooth. Schramm's definitive result on midscription is the main theorem of his *Inventiones* article [66] whimsically entitled *How to cage an egg*.

Convex Body Midscription (Schramm [66]) *Let P be a convex polyhedron and S a smooth strictly convex body in $\mathbb{E}^3$. Then there exists a convex polyhedron Q combinatorially equivalent to P that midscribes S.*

Discussion of Proof The proof is rather involved and so I am content to give the briefest of indication of its method. Schramm defines the *configuration space* $\mathbb{Z} = (\mathbb{E}^3)^{V(P)} \times \mathrm{G}(2,3)^{F(P)}$, where $V(P)$ and $F(P)$ are the respective sets of vertices and faces of P, and G(2, 3) is the manifold of oriented affine planes in $\mathbb{E}^3$. In this way P is identified with a single point of $\mathbb{Z}$, and the combinatorial type of P defines a submanifold $\mathbb{Z}_P$ of $\mathbb{Z}$ corresponding to various convex polyhedra in $\mathbb{E}^3$ that are combinatorially equivalent to P. Schramm then shows that there is a C^2 convex body S_0 with positively curved boundary that P midscribes. Let S_t, $0 \leq t \leq 1$, be a C^2-path of convex bodies with positively curved boundaries with $S_1 = S$. The idea now is to flow S_0 to S_1 along this path and drag combinatorial realizations of P along as midscribing polyhedra. The proof relies on a fine analysis of the configuration space $\mathbb{Z}$ and its submanifold $\mathbb{Z}_P$, and the method is to show that when S_t is midscribed by a realization of P, then so is $S_{t'}$ for all t' in an open interval about t. Then a delicate argument shows also that the set of parameter values for which S_t is midscribable by a realization of P is a closed set. Being open and closed, and nonempty since P midscribes S_0, this set of parameter values must be the whole of the unit interval, hence $S = S_1$ is midscribed by a realization of P. □

The remaining discussion on hyperbolic polyhedra has little to do, at least directly, with the Koebe–Andre'ev–Thurston Theorem. The arguments tend to be clever and technical, but ultimately involve the elementary geometry of hyperbolic space, often times realized as the unit ball in projective 3-space where the machinery of the Minkowski inner product and of de Sitter space is available. I include the discussion in order to complete for the reader the current state of affairs in the study of convex hyperbolic polyhedra, a study which I view as having been revitalized by Thurston's articulation of KAT I and pushed forward into the broader mathematical consciousness by the seminal work of Thurston's students, Oded Schramm and especially Igor Rivin.

5.7.2 *Compact and Convex Hyperbolic Polyhedra—Hodgson and Rivin*

In his doctoral thesis of 1986, Igor Rivin studied convex hyperbolic polyhedra. Therein he gave a characterization of compact, convex hyperbolic polyhedra that generalizes the Andre'ev results of [2],[21] and in articles in the early 1990s, extended his characterization to ideal polyhedra, generalizing Andre'ev's results in [3]. He

[21] See Roeder, Hubbard, and Dunbar's paper [63] for a readable proof of Andre'ev's classification of compact hyperbolic polyhedra with non-obtuse exterior dihedral angles.

used this latter generalization to answer definitively Steiner's question of 1832 asking for a characterization of those polyhedra that circumscribe a sphere. This of course is the content of the Circumscribable Type Characterization Theorem of the preceding section. In this section, I present an overview of Rivin's characterization of compact and convex hyperbolic polyhedra in terms of a generalized Gauss map. The overview embellishes Hodgson's outline presented in [45] (and repeated in [46]). In the section following, I outline Rivin's characterization of ideal polyhedra and make his observation that the Circumscribable Type Characterization Theorem is an immediate corollary of his characterization of ideal polyhedra.

To lay the groundwork, let's review the *Gauss map* $\mathtt{G}$ of a compact and convex Euclidean polyhedron P to the unit sphere $\mathbb{S}^2$. This is a set-valued map from the 2-complex forming the boundary of P that assigns to the point p of ∂P the set of outward pointing unit normals to support planes to P at p. Thus when p is a point of an open face f, $\mathtt{G}(p) = \mathtt{G}(f)$ is a single point determined by the outward unit normal to f. When p is in the open edge e incident to faces f and g, $\mathtt{G}(p) = \mathtt{G}(e)$ is the great circular arc connecting $\mathtt{G}(f)$ to $\mathtt{G}(g)$ of length equal to the exterior dihedral angle between f and g. Finally, for a vertex p of P, $\mathtt{G}(p)$ is the convex spherical polygon bounded by the arcs $\mathtt{G}(e)$ for edges e incident with p. When edges e and e' of the face f are incident at p, the interior angle of the polygon $\mathtt{G}(p)$ at the vertex $\mathtt{G}(f)$ is $\pi - \alpha$, where α is the interior angle of the face f at p. In this way the Gauss map realizes the Poincaré dual P^* of P as a geodesic cellular decomposition of the 2-sphere $\mathbb{S}^2$. Notice that the Gauss map does not encode all the information needed to reconstruct the polyhedron P. It encodes the interior angles of all the faces and the dihedral angles of all adjacent faces, but there is no encoding of side-lengths of the edges of P. For example, all rectangular boxes have the same image under the Gauss map, namely, a regular right-angled octahedral decomposition of the sphere $\mathbb{S}^2$.

Another way to describe the convex spherical polygon $\mathtt{G}(p)$ for a vertex p of P is as the polar dual $\mathtt{L}^*(p)$ of the infinitesimal link $\mathtt{L}(p)$ of p in P.[22] Note that $\mathtt{L}(p)$ is a convex spherical polygon with internal angles equal to the dihedral angles of the faces of P incident with p, and edge-lengths equal to the internal angles at the vertex p in the faces of P incident with p. Recall that an oriented great circle in $\mathbb{S}^2$ and its spherical center are polar duals of one another. The polar dual $\mathtt{L}^*(p)$ is obtained by replacing the edges of $\mathtt{L}(p)$ by the polar dual centers of their supporting great circles, and the vertices by appropriate arcs of the polar dual great circles. A nice exercise in spherical geometry verifies that $\mathtt{L}^*(p)$ is isometric to $\mathtt{G}(p)$. This gives an alternate construction of the Poincaré dual P^* as a geodesic, cellular decomposition of the 2-sphere—just isometrically glue the polar duals $\mathtt{L}^*(p)$ together as p ranges over the vertices of P along corresponding edges, $\mathtt{L}^*(p)$

[22] For a Euclidean polyhedron, $\mathtt{L}(p)$ is the intersection of P with a small sphere centered at p, one whose radius is smaller than the lengths of edges incident with p, rescaled to unit radius, and is oriented so that its interior is "to the left" as one traverses the polygon in its positive direction.

glued to $\mathtt{L}^*(q)$ whenever pq is an edge of P.[23] Obviously this gluing produces a 2-sphere, not only homeomorphic, but also isometric to the standard 2-sphere $\mathbb{S}^2$, and reproduces the cellular decomposition determined by the Gauss map.

It is this latter construction of the Poincaré dual P^* as a cellular decomposition of the 2-sphere that readily generalizes to convex and compact hyperbolic polyhedra. Indeed, let P now be a convex and compact hyperbolic polyhedron in $\mathbb{H}^3$ and for each vertex p, let $\mathtt{L}^*(p)$ be the polar dual of the infinitesimal link $\mathtt{L}(p)$ of p in P.[24] The link $\mathtt{L}(p)$, as in the Euclidean case, is an oriented convex spherical polygon in $\mathbb{S}^2$ with internal angles equal to the dihedral angles of the faces of P incident with p, and edge-lengths equal to the internal angles at the vertex p in the faces of P incident with p. The polar dual $\mathtt{L}^*(p)$ then encodes the exterior dihedral angles at the edges of P incident with p as the lengths of its edges, and the interior angles α of the faces incident with p as its interior angles in the form $\pi - \alpha$. This construction acts as a local Gauss map in a small neighborhood of the vertex p. Now exactly as before, isometrically glue the polar duals $\mathtt{L}^*(p)$ together as p ranges over the vertices of P along corresponding edges. The result is again a 2-sphere topologically, which is called the *Gaussian image* of P and denoted as $\mathtt{G}(P)$, with a spherical metric of constant unit curvature, except at the vertices. The vertices have cone type singularities with concentrated negative curvature. Indeed, at the vertex corresponding to the face $f = p_1 \cdots p_n$ of P, the angle sum is $\theta(f) = n\pi - \sum_{i=1}^n \alpha_i$, where α_i is the internal angle of f at the vertex p_i. In the hyperbolic plane, the compact and convex polygon f always has interior angle sum strictly less than $(n-2)\pi$ so that $\theta(f) > 2\pi$.

This brings us to Rivin's characterization of compact and convex hyperbolic polyhedra.

Compact Convex Hyperbolic Polyhedra Characterization (Rivin) *A metric space (M, g) homeomorphic to $\mathbb{S}^2$ can arise as the Gaussian image $\mathtt{G}(P)$ of a compact and convex polyhedron P in $\mathbb{H}^3$ if and only if these three conditions adhere.*

(i) *The metric g has constant curvature $+1$ except at a finite number of cone points.*
(ii) *The cone angle at each cone point is greater than 2π.*
(iii) *The lengths of the nontrivial closed geodesics of (M, g) are all strictly greater than 2π.*

Moreover, the metric g determines P uniquely up to hyperbolic congruence.

Recall that the Gauss map does not determine Euclidean polyhedra up to congruence since it contains no information about side lengths. In contrast, a

[23]The edge pq determines respective vertices u and v of $\mathtt{L}(p)$ and $\mathtt{L}(q)$ whose respective polar edges u^* and v^* have the same lengths, namely the exterior dihedral angle of P at edge pq.

[24]This is the link in the tangent space of $\mathbb{H}^3$ of the pre-image of the intersection of P with a small neighborhood of p under the exponential map.

hyperbolic polyhedron is determined up to a global hyperbolic isometry by its Gaussian image. The proof of this uniqueness uses Cauchy's toolbox that is reviewed in Addendum Sect. 5.7.5, wherein I recall the tools Cauchy used to prove his celebrated rigidity theorem of 1813. The necessity of items (i) and (ii) follows from the previous discussion and that of (iii) uses the fact that the total geodedic curvature of a non-trivial closed hyperbolic space curve is greater than 2π, a hyperbolic version of Fenchel's Theorem on Euclidean space curves. The proof of sufficiency is based on Aleksandrov's Invariance of Domain method used in his study of Euclidean polyhedra in [1].

Rivin also uses Cauchy's toolbox to prove this rather interesting theorem that illustrates again the enhanced rigidity of hyperbolic polyhedra vis-à-vis Euclidean ones.

Face Angle Rigidity (Rivin) *The face angles of a compact and convex polyhedron in $\mathbb{H}^3$ determine it up to congruence.*

The characterization of compact and convex hyperbolic polyhedra in terms of the Gaussian image surveyed here suffers from the same defect as Aleksandrov's characterization of compact and convex Euclidean polyhedra. Both characterizations posit a singular positively curved metric on a 2-sphere, but neither provides a way to decode from this metric space (M, g) the combinatorial type of the polyhedron P encoded in (M, g). The proof is not constructive, but depends on a topological analysis within the space of admissible metrics on the 2-sphere satisfying the three conditions of the characterization and yields, finally, the abstract fact of existence of an appropriate polyhedron, without describing its combinatorial type.

5.7.3 Convex Ideal Hyperbolic Polyhedra—Rivin

Rivin turns his attention to convex ideal polyhedra in $\mathbb{H}^3$ in [61] where he gives a full characterization in terms of exterior dihedral angles. The characterization begins with an analysis of the exterior dihedral angles of such a polyhedron reported in [47] with details in [59] that goes as follows. Label each edge e^* of the polyhedron P^* dual to the ideal convex polyhedron P by the exterior dihedral angle $\theta(e^*)$ of the corresponding edge e of P. Rivin's argument that these labels satisfy the following conditions is reproduced in the next two theorems.

(i) $0 < \theta(e^*) < \pi$ for all edges e of P.
(ii) If the edges $e_1^*, \ldots, e_n^*$ are the edges bounding a face of P^*, then $\theta(e_1^*) + \cdots + \theta(e_n^*) = 2\pi$.
(iii) If $e_1^*, \ldots, e_n^*$ forms a simple nontrivial circuit that does not bound a face of P^*, then $\theta(e_1^*) + \cdots + \theta(e_n^*) > 2\pi$.

Compare these conditions with the hypotheses of the Circumscribable Type Characterization on page 223. Now Condition (i) is a requirement of convexity and Condition (ii) is seen easily in the upper-half-space model by placing one of the ideal vertices v of P at infinity and observing that the link of v is a convex Euclidean polygon. Indeed, the faces incident with v lie on vertical Euclidean planes whose intersections with the xy-plane cut out a convex Euclidean polygon $L(v)$, and quite easily the sum $\theta(e_1^*) + \cdots + \theta(e_n^*)$ is precisely the sum of the turning angles of $L(v)$. Condition (iii) is a consequence of the following discrete, hyperbolic version of Fenchel's Theorem, in this case for closed polygonal curves in $\mathbb{H}^3$.

Discrete Total Curvature for Polygonal Hyperbolic Loops (Rivin [61]) *The total discrete geodesic curvature of a closed, polygonal, hyperbolic space curve is greater than 2π, unless the vertices are collinear, in which case the total curvature is 2π.*

Proof The total discrete geodesic curvature of the polygonal hyperbolic space curve γ with vertices $p_1, \ldots, p_k, p_{k+1} = p_1$ is $\sum_{i=1}^k \alpha_i$, where α_i is the turning angle of γ at p_i. The angle α_i is just the exterior angle at p_i of the triangle $\tau_i = p_{i-1}p_i p_{i+1}$. For $2 \leq i \leq k-1$, let T_i be the triangle $T_i = p_1 p_i p_{i+1}$ with internal angles a_i, b_i, and c_i at the respective vertices p_1, p_i, and p_{i+1}. Note that by considering the triangles τ_i, T_{i-1} and T_i with common vertex p_i, the spherical triangle inequality gives

$$c_{i-1} + b_i \geq \pi - \alpha_i \quad \text{for} \quad 3 \leq i \leq k-1,$$

and

$$b_2 = \pi - \alpha_2, \quad c_{k-1} = \pi - \alpha_k, \quad \text{and} \quad \sum_{i=2}^{k-1} a_i \geq \pi - \alpha_1.$$

Recalling that $\pi \geq a_i + b_i + c_i$ with equality only when p_1, p_i, and p_{i+1} are collinear, and then summing, one has

$$(k-2)\pi \geq \sum_{i=2}^{k-1}(a_i + b_i + c_i) \geq k\pi - \sum_{i=1}^{k} \alpha_i,$$

with equality only when $p_1, \ldots, p_k$ are collinear. □

Theorem 5.7.1 (Rivin [59]) *The edge label $\theta(e^*)$ of the polyhedron P^* dual to the ideal convex polyhedron P defined above satisfies Conditions* (i)–(iii).

Proof Conditions (i) and (ii) already are verified. For Condition (iii), the circuit e_1^*, $\ldots$, e_n^* that does not bound a face of P^* corresponds to a chain of contiguous faces $f_1, \ldots, f_n$ in P with $f_i \cap f_{i+1} = e_i$. $F = \cup_{i=1}^n f_i$ is a hyperbolic surface with boundary and cusps, and can be completed by extending geodesically across the boundary components to a complete immersed surface $\widetilde{F}$ in $\mathbb{H}^3$ without boundary. The surface $\widetilde{F}$ is an immersed hyperbolic cylinder with both ends of infinite-area. This observation uses the fact that the circuit $e_1^*, \ldots, e_n^*$ does not bound a face of P^*. Let γ be the unique closed geodesic path on the surface $\widetilde{F}$ that is freely homotopic to the meridian. The curve γ is immersed in $\mathbb{H}^3$ as a polygonal curve lying on $\widetilde{F}$ with turning angles at the edges e_i. But it is easy to see that the turning angle of γ at edge e_i is no more than the exterior dihedral angle of the faces f_i and f_{i+1} that meet along e_i. This implies that the sum, $\theta(e_1^*) + \cdots + \theta(e_n^*)$, which is the sum of these dihedral angles, is at least as large as the discrete geodesic curvature of γ, which in turn is greater than 2π by an application of the preceding theorem. □

Rivin was able to turn this around and prove a converse to the theorem, which gives the following characterization of convex, ideal hyperbolic polyhedra. The existence is proved in [61], uniqueness in [60], and necessity of the three conditions in [59].

Characterization of Convex Ideal Polyhedra (Rivin [61]) *Let P^* be an abstract polyhedron. Then for any label $\theta : E(P^*) \to (0, \pi)$ that satisfies Conditions* (i)–(iii)*, there is a convex, ideal hyperbolic polyhedron P in $\mathbb{H}^3$ whose Poincaré dual is P^*, and whose exterior dihedral angles at edges e are given by the values $\theta(e^*)$. Moreover, P is unique up to hyperbolic congruence. Conversely, every such polyhedron P satisfies Conditions* (i)–(iii) *as shown in Theorem 5.7.1.*

This characterization also proves the Circumscribable and Inscribable Type Characterizations, answering Steiner's question of 1832. This is because a convex, ideal hyperbolic polyhedron in the Beltrami–Klein projective model of $\mathbb{H}^3$ is represented by a convex Euclidean polyhedron inscribed in the 2-sphere $\mathbb{S}^2$.

Since Rivin's work of the 1990s, several topologists and geometers have taken up the mantel and continued to unearth these beautiful gems of discrete geometry. I'll close this survey with the mention of two examples in the next section, the first from the first decade of the new century, and the second of very recent origin.

5.7.4 New Millennium Excavations

Space constraints forbid too much further development of the topic, but I would be remiss if I didn't mention at least these two beautiful theorems, the first characterizing convex hyperideal hyperbolic polyhedra by Bao and Bonahon, and the second giving a complete answer to Steiner's original question when interpreted as broadly as possible, this time by Chen and Schlenker. I develop just enough of these topics to state the main results, and leave the interested reader the task of perusing the original articles for details of the proofs.

5.7.4.1 Hyperideal Polyhedra—Bao and Bonahon

A *hyperideal* polyhedron in $\mathbb{H}^3$ is a non-compact polyhedron that may be described most easily in the Beltrami–Klein projective model $\mathbb{H}^3 = \mathbb{B}^3 \subset \mathbb{RP}^3$ as the intersection with $\mathbb{B}^3$ of a projective polyhedron all of whose vertices lie outside of $\mathbb{B}^3$ while each edge meets $\mathbb{B}^3$. Bao and Bonahon [6] classify hyperideal polyhedra up to hyperbolic congruence in terms of their dihedral angles and combinatorial type in much the same vein as Rivin's classification of ideal hyperbolic polyhedra. Note that Bao and Bonahon do allow for the vertices to lie on the sphere $\mathbb{S}^2 = \partial\mathbb{B}^3$ and hence their characterization reduces to Rivin's for ideal polyhedra.

I will state the characterization in terms of conditions on the 1-skeletal graph of the dual polyhedron using Steinitz's famous characterization of those graphs that may serve as the dual graph of a convex polyhedron in $\mathbb{E}^3$ as precisely the planar, 3-connected graphs.

Characterization of Convex Hyperideal Polyhedra (Bao and Bonahon [6]) *Let $\mathcal{G}$ be a 3-connected graph embedded in $\mathbb{S}^2$ and $\theta : E(\mathcal{G}) \to (0, \pi)$. There is a hyperideal polyhedron P in $\mathbb{H}^3$ with dual graph isomorphic with $\mathcal{G}$ and exterior dihedral angles given by θ if and only if the following conditions are satisfied.*

(i) *If $e_1, \dots, e_n$ forms a simple nontrivial circuit of edges of $\mathcal{G}$, then $\theta(e_1) + \cdots + \theta(e_n) \geq 2\pi$, with equality possible only if $e_1, \dots, e_n$ bounds a component of $\mathbb{S}^2 - \mathcal{G}$.*

(ii) *If $\gamma = e_1, \dots, e_n$ forms a simple path of edges of $\mathcal{G}$ that connects two vertices of $\mathcal{G}$ that lie in the closure of a component C of $\mathbb{S}^2 - \mathcal{G}$, but γ does not lie in the boundary of C, then $\theta(e_1) + \cdots + \theta(e_n) > \pi$.*

Moreover if P' is the projective polyhedron with $P' \cap \mathbb{H}^3 = P$, a vertex v of P' is located on the sphere at infinity of $\mathbb{H}^3$ if and only if equality holds in Condition (i) *for the boundary of the corresponding component of $\mathbb{S}^2 - \mathcal{G}$.*

Finally, the hyperideal polyhedron P is unique up to hyperbolic congruence.

I should mention that Hodgson and Rivin's [46] characterization of compact and convex hyperbolic polyhedra can be applied to appropriate truncated polyhedra associated with those hyperideal polyhedra for which no vertex lies on the sphere at infinity to characterize them.

Define a *strictly hyperideal* polyhedron to be the intersection of $\mathbb{B}^3$ with a projective polyhedron P all of whose vertices lie outside the closed unit ball $\overline{\mathbb{B}}^3 = \mathbb{B}^3 \cup \mathbb{S}^2$, but all of whose faces meet $\mathbb{B}^3$. Note that this definition allows that an edge of P may lie entirely outside the closed ball $\overline{\mathbb{B}}^3$. These are yet to be characterized, but I mention that the article [23] verifies the rigidity of these that are bounded and convex, as long as no edges are tangent to the unit sphere. The proof again uses Cauchy's toolbox.

5.7.4.2 Weakly Inscribed Polyhedra—Chen and Schlenker

Recall Steiner's question of which polyhedra inscribe or circumscribe a sphere that Rivin answered. A more faithful translation of Steiner's question from the German is "Does every polyhedron have a combinatorially equivalent realization that is inscribed or circumscribed to a sphere, or to another quadratic surface? If not, which polyhedra have such realizations?" He includes the definition that "A polyhedron P is *inscribed to* a quadratic surface S if all the vertices of P lie on S," and further defines that P is *circumscribed to* S if all of its facets are tangent to S. As before I will concentrate on inscription since polarity relates circumscription to inscription. In the very recent preprint [30], Chen and Schlenker point out that the apparent grammar mistake—inscribed *to* instead of *in* S—makes a significant distinction.

Generally Steiner's question has been interpreted to ask about inscription of the polyhedron P to a quadratic surface S in Euclidean space $\mathbb{E}^3$, and in this setting P is contained in the bounded component of the complement of S, i.e., P is "inside" S, hence the change from inscribed "to" to "in". But Steiner's question makes sense in projective space as well, and in this setting a polyhedron may be inscribed to a surface without being inscribed in the surface. To be a bit more illustrative, consider the unit sphere $\mathbb{S}^2$ sitting in $\mathbb{E}^3 \subset \mathbb{RP}^3$. Now $\mathbb{S}^2$ usually is thought of as the boundary of the open unit ball $\mathbb{B}^3$ that serves as the projective model of hyperbolic space, and this is what Rivin exploited in his characterization of those polyhedra inscribable in $\mathbb{S}^2$. But $\mathbb{S}^2$ is also the boundary of the complement $\mathbb{RP}^3 - \overline{\mathbb{B}}^3$, which has a complete metric making it into a model of de Sitter space $d\mathbb{S}^3$. In this setting a projective polyhedron may have its vertices on the sphere $\mathbb{S}^2$ and yet not lie entirely in the ball $\mathbb{B}^3$ so that it is inscribed to $\mathbb{S}^2$, but not inscribed in $\mathbb{S}^2$ in the usual meaning. Following Chen and Schlenker, I will revise Steiner's terminology to emphasize the difference between *inscribed in* and *inscribed to but not in*.

Definition (Strong and Weak Inscription) In the real projective space $\mathbb{RP}^3$, a polyhedron P inscribed to a quadratic surface S is *strongly inscribed in* S if the interior of P is disjoint from S, and *weakly inscribed to* S otherwise.

Before presenting a characterization of those polyhedra weakly inscribed to a sphere in $\mathbb{RP}^3$, allow a word about polyhedra inscribed to other quadratic surfaces. This topic has been neglected until rather recently. There are only three quadratic surfaces in $\mathbb{RP}^3$ up to projective transformations, and these are the sphere, the one-sheeted hyperboloid, and the cylinder. Danciger, Maloni, and Schlenker in [31] characterized the combinatorial types of polyhedra that are strongly inscribable in a one-sheeted hyperboloid or in a cylinder, and of course Rivin takes care of those strongly inscribable in a sphere. Chen and Schlenker's work reported here characterizes those polyhedra weakly inscribable to a sphere, and the characterization of those weakly inscribable to the remaining two quadratic surfaces is the subject of current research by Chen and Schlenker.

Weak Inscription Characterization (Chen and Schlenker [30]) *A 3-connected planar graph Γ is the 1-skeleton of a polyhedron $P \subset \mathbb{RP}^3$ weakly inscribed to a sphere if and only if Γ admits a vertex-disjoint cycle cover by two cycles C_1 and C_2 with the following property. Color edge uv red if u and v both belong to C_1 or both belong to C_2, and color it blue otherwise. Then there is a weight function $w : E(\Gamma) \to \mathbb{R}$ such that*

(i) *$w > 0$ on red edges and $w < 0$ on blue ones;*
(ii) *w sums to 0 over the edges adjacent to a vertex v, unless v is the only vertex on C_1 or C_2 (trivial cycle), in which case w sums to -2π over the edges adjacent to v.*

I end this survey of progress in the characterization of polyhedra since Thurston's observation that every polyhedron type in $\mathbb{E}^3$ has a realization that midscribes a sphere with a description of the original rigidity theorem of Cauchy that is so instrumental in many of the proofs of the results surveyed here.

5.7.5 Addendum: Cauchy's Toolbox

In this bonus section I review Cauchy's celebrated rigidity theorem [29] of 1813 on the uniqueness of convex, bounded polyhedra in $\mathbb{E}^3$. The theorem concerns two convex polyhedra with equivalent combinatorics and with corresponding faces congruent. Cauchy's Rigidity Theorem states that the two polyhedra must be congruent globally, meaning that there is a Euclidean isometry of the whole of $\mathbb{E}^3$ mapping one to the other. Like many of the great theorems of mathematics, the proof is of more importance than the theorem itself. As stated earlier in the introduction to this section, the toolbox Cauchy developed has been instrumental in the past 200 year development of the theory of polyhedra, especially in its rigidity theory. The proof, though at places clever and even subtle, overall is rather straightforward with a simplicity that belies its importance.

Cauchy's proof has two components—the one geometric and the other combinatorial. The geometric component is the Discrete Four Vertex Lemma, which follows from an application of Cauchy's Arm Lemma. Denote a convex planar or spherical polygon P merely by listing its vertices in cyclic order, say as $P = p_1 \dots p_n$. The Euclidean or spherical length of the side $p_i p_{i+1}$ is denoted as $|p_i p_{i+1}|$ and the interior angle at p_i is denoted as $\angle p_i$.

Cauchy Arm Lemma *Let $P = p_1 \dots p_n$ and $P' = p'_1 \dots p'_n$ be two convex planar or spherical polygons such that, for $1 \le i < n$, $|p_i p_{i+1}| = |p'_i p'_{i+1}|$, and for $1 \le i < n-1$, $\angle p_{i+1} \le \angle p'_{i+1}$. Then $|p_n p_1| \le |p'_n p'_1|$ with equality if and only if $\angle p_{i+1} = \angle p'_{i+1}$ for all $1 \le i < n-1$.*

Cauchy's original proof of the lemma had a gap that subsequently was filled by Ernst Steinitz. A straightforward inductive proof, such as the one in [36], relies on the law of cosines and the triangle inequality.

Now let P and P' be convex planar or spherical polygons with the same number of sides whose corresponding sides have equal length. Label each vertex of P with a plus sign $+$ or a minus sign $-$ by comparing its angle with the corresponding angle in P': if the angle at p_i is larger than that at p'_i, label it with a $+$, if smaller, a $-$, and if equal, no label at all. Using the Cauchy Arm Lemma, the proof of the following lemma is straightforward.

Discrete Four Vertex Lemma *Let P and P' be as in the preceding paragraph and label the vertices of P as described. Then either P and P' are congruent, or a walk around P encounters at least four sign changes, from $-$ to $+$ or from $+$ to $-$.*

Proof First note that because a polygon is a cycle, the number of sign changes must be even. If no vertex is labeled, then the two polygons are congruent. Assume then that some of the vertices are labeled, but all with the same label. Then Cauchy's Arm Lemma implies that there exists a pair of corresponding edges in P and P' with different lengths, a contradiction.

Assume now that there are exactly two sign changes of the labels of P. Select two edges $p_i p_{i+1}$ and $p_j p_{j+1}$ (oriented counter-clockwise) of P such that all of the $+$ signs are along the subchain from p_{i+1} to p_j and all of the $-$ signs are along the subchain from p_{j+1} back to p_i. Subdivide both edges in two by adding a vertex at the respective midpoints X and Y of $p_i p_{i+1}$ and $p_j p_{j+1}$. Similarly, subdivide the corresponding edges $p'_i p'_{i+1}$ and $p'_j p'_{j+1}$ in P' at midpoints X' and Y'. Denote the subchain of P from X to Y by P_+ and the subchain from Y back to X by P_-. Similarly for P'_+ and P'_- in P'. Applying the arm lemma to P_+ and P'_+ implies that $|XY| > |X'Y'|$, and, similarly, an application to P_- and P'_- implies that $|XY| < |X'Y'|$, a contradiction. □

This brings us to the combinatorial component of Cauchy's proof. A nice proof of the following lemma appears in [36] and follows from an argument based on the Euler characteristic of a sphere.

Cauchy Combinatorial Lemma *Let P be an abstract spherical polyhedron. Then for any labeling of any non-empty subset of the edges of P with $+$ and $-$ signs, there exists a vertex v that is incident to an edge labeled with a $+$ or a $-$ sign for which one encounters at most two sign changes in labels on the edges adjacent to v as one walks around the vertex.*

Cauchy Rigidity Theorem *If two bounded, combinatorially equivalent, convex polyhedra in $\mathbb{E}^3$ have congruent corresponding faces, then they are congruent by a Euclidean isometry of $\mathbb{E}^3$.*

Proof Assume that bounded, convex polyhedra P and P' have the same combinatorics and congruent corresponding faces. For each edge of P, label its dihedral angle with a $+$ or a $-$ depending on whether it is larger or smaller than the corresponding dihedral angle in P'. If P and P' are not congruent, Cauchy's Combinatorial Lemma provides a vertex v that is incident to an edge labeled with a $+$ or a $-$ sign, and around which there are at most two sign changes. Intersect P with a small sphere centered at v (one that contains no other vertex of P on its interior) to obtain a convex spherical polygon, and intersect P' with a sphere centered at the corresponding vertex v' and of the same radius. By construction both spherical polygons have the same edge lengths, and the angles between edges are given by the dihedral angles between faces at v and v'. An application of the Four Vertex Lemma implies that there are at least four sign changes, contradicting that there are at most two. It follows that P and P' are congruent. □

Both the bounded and convex requirements are necessary. For example, a polyhedron $\hat{H}$ in the shape of a cubical house with a shallow pyramidal roof has a cousin $\check{H}$ obtained by inverting the roof. $\hat{H}$ is not congruent to $\check{H}$, though these are combinatorially equivalent with congruent corresponding faces.

5.8 In Closing, an Open Invitation

This has been a whirlwind tour through the four decade history of the influence of one theorem brought to prominence by the mathematician we celebrate in this volume. Any result that has spawned such a great body of significant work leaves in its wake a bounty of open questions, problems, conjectures, and possible applications that await the right insights for resolution and explanation. What of the Koebe Uniformization Conjecture, of the question of where EQ-type sits among EEL- and VEL-type, of circle packings that mimic rational functions with arbitrary branching, of the existence and rigidity of inversive distance circle packings, of characterizations of projective polyhedra up to Möbius equivalence generalizing Bao-Bonahon, or of combinatorial rather than metric characterizations of hyperbolic polyhedra of various stripes? I have not covered in this survey the myriad of applications that circle packing has spawned, particularly in the realm of computer graphics and imaging, where each month sees more and more new and original publications. And so I close this tribute to the influence of this one theorem of Bill Thurston with an invitation to any reader who has been captured by the beauty and elegance of the results outlined in this survey to explore further on his or her own the wider discipline of Discrete Conformal Geometry, in both its theoretical and practical bents, and perhaps to add to our understanding and appreciation of this beautiful landscape opened up by the imagination of Bill Thurston.

Acknowledgments I thank Prof. Athanase Papadopoulos for inviting me to write on a favorite theme of mine to honor Bill Thurston and his legacy. It has been a pleasure for me to review the impact derived from this one beautiful little gem of Thurston. I thank Ken Stephenson for permission to use the graphics of Figs. 5.1, 5.2, 5.5, and 5.7, and John Bowers for generating the graphics for Figs. 5.3, 5.4, 5.6, 5.8, 5.9, and 5.10.

References

1. A.D. Aleksandrov, *Convex Polyhedra* (Translation of 1950 Russian ed.). Springer Monographs in Mathematics (Springer, Berlin, 2005)
2. E.M. Andre'ev, On convex polyhedra in Lobachevski spaces. Mat. Sbornik **81**(123), 445–478 (1970)
3. E.M. Andre'ev, On convex polyhedra of finite volume in Lobachevski spaces. Mat. Sbornik **83**(125), 256–260 (1970)
4. O. Angel, O. Schramm, Uniform infinite planar triangulations, Commun. Math. Phys. **241**, 191–213 (2003)

5. J. Ashe, E. Crane, K. Stephenson, Circle packings with generalized branching (2016). arXiv160703404A
6. X. Bao, F. Bonahon, Hyperideal polyhedra in hyperbolic 3-space. Bull. Soc. Math. France **130**(3), 457–491 (2002)
7. A.F. Beardon, K. Stephenson, The uniformization theorem for circle packings. Indiana Univ. Math. J. **39**, 1383–1425 (1990)
8. A.F. Beardon, K. Stephenson, The Schwarz-Pick lemma for circle packings, Ill. J. Math. **35**, 577–606 (1991)
9. M. Berger, *Geometry Revealed: A Jacob's Ladder to Modern Higher Geometry* (Springer, Berlin, 2010)
10. A.I. Bobenko, B.A. Springborn, Variational principles for circle patterns and Koebe's theorem. Trans. Amer. Math. Soc **356**(2), 659–689 (2004)
11. A.I. Bobenko, T. Hoffmann, B.A. Springborn, Minimal surfaces from circle patterns: Geometry from combinatorics. Ann. Math. **164**(1), 231–264 (2006)
12. P.L. Bowers, The upper Perron method for labelled complexes with applications to circle packings. Proc. Camb. Phil. Soc. **114**, 321–345 (1993)
13. P.L. Bowers, Negatively curved graph and planar metrics with applications to type. Mich. Math. J. **45**, 31–53 (1998)
14. P.L. Bowers, Introduction to circle packing: the theory of discrete analytic functions [book review]. Bull. Amer. Math. Soc. **46**(3), 511–525 (2009)
15. J.C. Bowers, P.L. Bowers, Ma–schlenker c-octahedra in the 2-sphere. Discrete Comput. Geometry **60**, 9–26 (2017)
16. P.L. Bowers, M.K. Hurdal, Planar conformal mappings of piecewise flat surfaces, in *Visualization and Mathematics III*, chap. 1, (Springer, Berlin, 2003), pp. 3–34
17. P.L. Bowers, K. Stephenson, The set of circle packing points in the Teichmüller space of a surface of finite conformal type is dense. Math. Proc. Camb. Phil. Soc. **111**, 487–513 (1992)
18. P.L. Bowers, K. Stephenson, A branched Andreev–Thurston theorem for circle packings of the sphere. Proc. London Math. Soc. **73**(3), 185–215 (1996)
19. P.L. Bowers, K. Stephenson, A "regular" pentagonal tiling of the plane. Conform. Geom. Dyn. **1**, 58–86 (1997)
20. P.L. Bowers, K. Stephenson, *Uniformizing Dessins and Belyĭ Maps via Circle Packing*. Memoirs of the AMS, vol. 170, no. 805 (American Mathematical Society, Providence, 2004)
21. P.L. Bowers, K. Stephenson, Conformal tilings I: foundations, theory, and practice. Conform. Geom. Dyn. **21**(1), 1–63 (2017)
22. P.L. Bowers, K. Stephenson, Conformal tilings II: local isomorphism, hierarchy, and conformal type Conform. Geom. Dyn. **23**, 60 (2018)
23. J.C. Bowers, P.L. Bowers, K. Pratt, Rigidity of circle polyhedra in the 2-sphere and of hyperideal polyhedra in hyperbolic 3-space. Trans. Amer. Math. Soc. **371**, 4215–4249 (2018)
24. J.W. Cannon, The combinatorial structure of cocompact discrete hyperbolic groups. Geometriae Dedicata **16**(2), 123–148 (1984)
25. J.W. Cannon, The combinatorial Riemann mapping theorem. Acta Math. **173**, 155–234 (1994)
26. J.W. Cannon, W.J. Floyd, W.R. Parry, Finite subdivision rules. Conform. Geom. Dyn. **5**, 153–196 (2001)
27. C. Carathéodory, Untersuchungen über die konformen abbildungen von festen und veränderlichen gebieten. Math. Annal. **72**(1), 107–144 (1912)
28. C. Carathéodory, *Conformal Representation*. Cambridge Tracts in Mathematics and Mathematical Physics, vol. 28, Reprint of 1952 edn. (Cambridge University Press, Cambridge, 2008)
29. A. Cauchy, Sur les polygones et les polyèdres. J. Ecole Polytechnique XVIe Cahier **IX**, 87–98 (1813)
30. H. Chen, J.-M. Schlenker, Weakly inscribed polyhedra (2017). arXiv170910389C, to appear in Trans. Amer. Math. Soc., Series B
31. J. Danciger, S. Maloni, J.-M. Schlenker, Polyhedra inscribed in a quadric. Invent. Math. **221**(1), 237–300 (2020)

32. H.P. de Saint-Gervais, *Uniformization of Riemann Surfaces: Revisiting a Hundred-Year-Old Theorem*. Heritage of European Mathematics, vol. 11 (European Mathematical Society, Zürich, 2016)
33. Y.C. de Verdière, Une principe variationnel pour les empilements de cercles. Invent. Math. **104**, 655–669 (1991)
34. P.G. Doyle, J. Laurie Snell, *Random Walks and Electric Networks*, vol. 22, 1st edn. (Mathematical Association of America, Washington, 1984)
35. R.J. Duffin, The extremal length of a network. J. Math. Anal. Appl. **5**, 200–215 (1962)
36. D. Fuchs, S. Tabachnikov, *Mathematical Omnibus: Thirty Lectures on Classic Mathematics* (American Mathematical Society, Providence, 2007)
37. J.T. Gill, S. Rohde, On the Riemann surface type of random planar maps. Rev. Math. Iber. **29**, 1071–1090 (2013)
38. J. Gray, On the history of the Riemann mapping theorem. Rendiconti del Cir. Math. di Palermo **II**(34), 47/94 (1994)
39. M. Gromov, *Hyperbolic groups*, in Essays in Group Theory, ed. by S.M. Gersten (Springer, New York, 1987), pp. 75Ű-263
40. B. Grünbaum, *Convex Polytopes*. Graduate Texts in Mathematics, vol. 221, 2nd edn. (Springer, Belin, 2003)
41. R. Guo, Local rigidity of inversive distance circle packing. Trans Amer. Math. Soc. **363**(9), 4757–4776 (2011)
42. Z.-X. He, O. Schramm, Fixed points, Koebe uniformization and circle packings. Ann. Math. **137**, 369–406 (1993)
43. Z.-X. He, O. Schramm, Hyperbolic and parabolic packings. Discrete Comput. Geom. **14**, 123–149 (1995)
44. Z.-X. He, O. Schramm, On the convergence of circle packings to the Riemann map. Invent. Math. **125**(2), 285–305 (1996)
45. C.D. Hodgson, Deduction of Andeev's theorem from Rivin's characterization of convex hyperbolic polyhedra, in *Topology 90. Proceedings of the Research Semester in Low Dimensional Topology at O.S.U* (1993)
46. C.D. Hodgson, I. Rivin, A characterization of compact convex polyhedra in hyperbolic 3-space. Invent. Math. **111**, 77–111 (1993)
47. C.D. Hodgson, I. Rivin, W.D. Smith, A characterization of convex hyperbolic polyhedra and of convex polyhedra inscribed in the sphere. Bull. Amer. Math. Soc. **27**, 246–251 (1992)
48. P. Koebe, Uber die uniformisierung beliebiger analytischer Kurven, III. Nach. Ges. Wiss. Gott. 337–358 (1908)
49. P. Koebe, Abhandlungen zur theorie der konformen abbildung: VI. abbildung mehrfach zusammenhängender Bereiche auf Kreisbereiche, etc. Math. Z. **7**, 235–301 (1920)
50. P. Koebe, Kontaktprobleme der konformen abbildung. Ber. Sächs. Akad. Wiss. Leipzig, Math.-Phys. Kl. **88**, 141–164 (1936)
51. G.F. Lawler, V. Limic, *Random Walk: A Modern Introduction*. Cambridge Studies in Advanced Mathematics, no. 123 (Cambridge University Press, Cambridge, 2010)
52. F. Luo, Rigidity of polyhedral surfaces, III. Geom. Topol. **15**(4), 2299–2319 (2011)
53. J. Ma, J.-M. Schlenker, Non-rigidity of spherical inversive distance circle packings. Discrete Comput. Geom. **47**(3), 610–617 (2012)
54. A. Marden, B. Rodin, On Thurston's formulation and proof of Andreev's theorem, in *Proceedings of a Conference, held in Valparaíso Computational Methods and Function Theory, 1989*. Lecture Notes in Mathematics, vol. 1435 (Springer, Berlin, 1990), pp. 103–115
55. C. St. J.A. Nash–Williams, Random walk and electric currents in networks. Proc. Camb. Phil. Soc. **55**, 181–195 (1959)
56. U. Pinkall, K. Polthier, Computing discrete minimal surfaces and their conjugates. Exp. Math. **2**, 15–36 (2012)
57. G. Pólya, Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die irrfahrt im Strassennetz. Math. Annal. **84**(1–2), 149–160 (1921)

58. S. Rhode, Oded Schramm: From circle packing to SLE, in *Selected Works of Oded Schramm*. Selected Works in Probability and Statistics, vol. 1 (Springer, Berlin, 2011), pp. 3–45
59. I. Rivin, On geometry of convex ideal polyhedra in hyperbolic 3-space. Topology **32**(1), 87–92 (1993)
60. I. Rivin, Euclidean structures on simplicial surfaces and hyperbolic volume. Ann. Math. **139**(3), 553 (1994)
61. I. Rivin, A characterization of ideal polyhedra in hyperbolic 3-space. Ann. Math. **143**, 51–70 (1996)
62. B. Rodin, D. Sullivan, The convergence of circle packings to the Riemann mapping. J. Diff. Geom. **26**, 349–360 (1987)
63. R.K.W. Roeder, J.H. Hubbard, W.D. Dunbar, Andreev's theorem on hyperbolic polyhedra. Ann. de l'Inst. Fourier **57**(3), 825–882 (2007)
64. O. Schramm, Existence and uniqueness of packings with specified combinatorics. Isr. J. Math. **73**(3), 321–341 (1991)
65. O. Schramm, Rigidity of infinite (circle) packings. J. Amer. Math. Soc. **4**, 127–149 (1991)
66. O. Schramm, How to cage an egg. Invent. Math. **107**(1), 543–560 (1992)
67. O. Schramm, Transboundary extremal length. J. d'Analyse Math. **66**(1), 307–329 (1995)
68. E. Shulte, Analogues of Steinitz's theorem about noninscribable polytopes, in *Proceedings of the Intuitive geometry, Siófok 1985*. Colloquia Mathematica Societatis János Bolyai, vol. 48 (1985), pp. 503–516
69. J. Steiner, *Systematische Entwickelung der Abhängigkeit Geometrischer Gestalten von Einander* (Fincke, Berlin, 1832)
70. K. Stephenson, A probabilistic proof of Thurston's conjecture on circle packings. Rendiconti del Seminario Mate. e Fisico di Milano **LXVI**, 201–291 (1996)
71. K. Stephenson, *Introduction to Circle Packing: the Theory of Discrete Analytic Functions* (Cambridge University Press, New York, 2005). (ISBN 0-521-82356-0, QA640.7.S74)
72. D. Sullivan, On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions, in *Proceedings of the 1978 Stony Brook Conference on Riemann Surfaces and Related Topics* (Princeton University Press, Princeton, 1981), pp. 465–496
73. W.P. Thurston, *The Geometry and Topology of 3-Manifolds*. Lecture Notes (Princeton University, Princeton, 1980)
74. W. Woess, *Random Walks on Infinite Graphs and Groups*. Cambridge Tracts in Mathematics, no. 138 (Cambridge University Press, Cambridge, 2000)
75. W.E. Wood, Combinatorial type problems for triangulation graphs, Ph.D. Thesis, FSU, Tallahassee, Advisor Philip L. Bowers, 2006
76. W.E. Wood, Bounded outdegree and extremal length on discrete Riemann surfaces. Conform. Geom. Dyn. **14**, 194–201 (2010)

Chapter 6
On Thurston's Parameterization of $\mathbb{C}P^1$-Structures

Shinpei Baba

Abstract Thurston established a correspondence between $\mathbb{C}P^1$-structures (complex projective structures) and equivariant pleated surfaces in the hyperbolic-three space $\mathbb{H}^3$, in order to give a parameterization of the deformation space of $\mathbb{C}P^1$-structures. In this note, we summarize Thurston's parametrization of $\mathbb{C}P^1$-structures, based on [15] and [17], giving an outline and the key points of its construction.

In addition we give independent proofs for the following well-known theorems on $\mathbb{C}P^1$-structures by means of pleated surfaces given by the parameterization. (1) Goldman's Theorem on $\mathbb{C}P^1$-structures with quasi-Fuchsian holonomy. (2) The path lifting property of developing maps in the domain of discontinuities in $\mathbb{C}P^1$.

Keywords $\mathbb{C}P^1$-structures · Measured laminations · Pleated surfaces

AMS Classification 57M50

6.1 Introduction

Let P be the space of all (marked) $\mathbb{C}P^1$-structures on a closed oriented surface S of genus at least two (Sect. 6.2). Thurston gave the following parameterization of P, using pleated surfaces in the hyperbolic three-space $\mathbb{H}^3$.

Theorem A (Thurston, [15, 17])

$$\mathsf{P} \cong \mathsf{ML} \times \mathsf{T},$$

where ML is the space of measured laminations on S and T is the space of all (marked) hyperbolic structures on S.

S. Baba (✉)
Osaka University, Toyonaka, Japan
e-mail: baba@math.sci.osaka-u.ac.jp

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_6

In Sect. 6.4, we outline this correspondence, in part, giving more details, following the work of Kulkarni and Pinkall [17]. A hyperbolic structure on S is in particular a $\mathbb{C}P^1$ structure, and its holonomy is a discrete and faithful representation of $\pi_1(S)$ into $PSL(2, \mathbb{R})$, called a *Fuchsian representation*. One holonomy representation of a $\mathbb{C}P^1$-structure on S corresponds to countably many different $\mathbb{C}P^1$-structures on S. Indeed, there is an operation called 2π-grafting (or simply grafting) which transforms a $\mathbb{C}P^1$-structure to a new $\mathbb{C}P^1$-structure, preserving its holonomy representations. The following theorem of Goldman characterizes all $\mathbb{C}P^1$-structures with fixed Fuchsian holonomy.

Theorem B ([11]) *Every $\mathbb{C}P^1$-structure C on S with Fuchsian holonomy ρ is obtained by grafting the hyperbolic structure τ along a unique multiloop M.*

Goldman actually proved the theorem for more general quasi-Fuchsian groups, although the proof is immediately reduced to the case of Fuchsian representations by a quasiconformal map of $\mathbb{C}P^1$. Let C be a $\mathbb{C}P^1$-structure with Fuchsian holonomy $\pi_1(S) \to PSL(2, \mathbb{C})$. Then, by Theorem B, C corresponds to (τ, M), where τ is the hyperbolic structure $\mathbb{H}^2/\mathrm{Im}\rho$ and each loop of M has a 2π-multiple weight.

For a subgroup $\Gamma \subset PSL(2, \mathbb{C})$, the *limit set* of Γ is the set of accumulation points of a Γ-orbit in $\mathbb{C}P^1$, and the domain of discontinuity is the complement of the limit set in $\mathbb{C}P^1$. In Sect. 6.5, we give an alternative proof of Theorem B, directly using pleated surfaces given by the Thurston parameters.

The following Theorem is a technical part of the proof of Theorem B, which was originally missing.

Theorem C ([7], See also §14.4.1. in [12]) *Let (f, ρ) be a developing pair of a $\mathbb{C}P^1$-structure on S. Let Ω be the domain of discontinuity of* $\mathrm{Im}\rho$. *Then, for each connected component U of $f^{-1}(\Omega)$, the restriction of f to U is a covering map onto its image.*

Note that as developing maps are local homeomorphisms, Theorem C is equivalent to saying that f has the path lifting property in the domain of discontinuity of $\mathrm{Im}\rho$.

We also give an alternative proof of Theorem C in Sect. 6.6, using Thurston's parametrization.

Theorem B states that given two $\mathbb{C}P^1$-structures C_1 and C_2 with Fuchsian holonomy, C_1 can be transformed into C_2, via the hyperbolic structure, by a composition of an inverse-grafting and a grafting (where an inverse grafting is the opposite of grafting which removes a cylinder for 2π-grafting). The following question due to Gallo, Kapovich, and Marden remains open.

Conjecture 6.1.1 (§12.1 in [10]) Given two $\mathbb{C}P^1$-structures C_1, C_2 on S with fixed holonomy $\pi_1(S) \to PSL(2, \mathbb{C})$, there is a composition of grafts and inverses of grafts which transforms C_1 into C_2.

Although [10] stated this conjecture in the form of a question, we state it more positively since it has been solved affirmatively for generic holonomy representations, namely, for purely loxodromic representations [3, 4]. (For Schottky representations, see [2].) There is also a version of this question for branched $\mathbb{C}P^1$-structures (Problem 12:1:2 in [10]); see [5, 19] for some progress in the case of branched $\mathbb{C}P^1$-structures.

Recently, Gupta and Mj [13] gave a generalization of Theorem A to certain $\mathbb{C}P^1$-structures on a surface with punctures (namely, $\mathbb{C}P^1$-structures which corresponds to compact Riemann surfaces with meromorphic quadratic differentials whose poles are of order at least three); see also [1].

6.2 $\mathbb{C}P^1$-Structures on Surfaces

General references for $\mathbb{C}P^1$-structures can be found in, for example [8, 16].

A $\mathbb{C}P^1$-*structure* on S is a $(\mathbb{C}P^1, \mathrm{PSL}(2, \mathbb{C}))$-structure, i.e. a maximal atlas of charts embedding open subsets of S onto open subsets of $\mathbb{C}P^1$ such that their transition maps are in $\mathrm{PSL}(2, \mathbb{C})$. Let $\tilde{S}$ be the universal cover of S, which is topologically an open disk. Then, equivalently, a $\mathbb{C}P^1$-structure on S is defined as a pair (f, ρ) consisting of

- a local homeomorphism $f\colon \tilde{S} \to \mathbb{C}P^1$ (*developing map*) and
- a homomorphism $\rho\colon \pi_1(S) \to \mathrm{PSL}(2, \mathbb{C})$ (*holonomy representation*)

such that f is ρ-equivariant (i.e. $f\alpha = \rho(\alpha) f$ for all $\alpha \in \pi_1(S)$). This pair(f, ρ) is called the *developing pair* of C, and (f, ρ) is, by definition, equivalent to $(\gamma f, \gamma\rho\gamma^{-1})$ for all $\gamma \in \mathrm{PSL}(2, \mathbb{C})$. Due to the equivariance condition, we do not usually need to distinguish between an element of $\pi_1(S)$ and its free homotopy class. Let P be the deformation space of all $\mathbb{C}P^1$-structures on S; then P has a natural topology, given by the open-compact topology on the developing maps $f\colon \tilde{S} \to \mathbb{C}P^1$.

Notice that hyperbolic structures are, in particular, $\mathbb{C}P^1$-structures, as $\mathbb{H}^2$ is the upper half-plane in $\mathbb{C}$ and the orientation-preserving isometry group $\mathrm{Isom}\,\mathbb{H}^2$ is the subgroup $\mathrm{PSL}(2, \mathbb{R})$ of $\mathrm{PSL}(2, \mathbb{C})$.

6.3 Grafting

A grafting is a cut-and-paste operation of a $\mathbb{C}P^1$-structure inserting some structure along a loop, an arc or more generally a lamination, originally due to [14, 18, 20]. There are slightly different versions of grafting, but they all yield new $\mathbb{C}P^1$-structures without changing the topological types of the base surfaces.

A *round circle* in $\mathbb{C}P^1 = \mathbb{C} \cup \{\infty\}$ is a round circle in $\mathbb{C}$ or a straight line in $\mathbb{C}$ plus ∞. A *round disk* in $\mathbb{C}P^1$ is a disk bounded by a round circle. An arc α on a $\mathbb{C}P^1$-structure is *circular* if α is immersed to (or embedded in) a round circle on $\mathbb{C}P^1$ by the developing map. Similarly, a loop α on a $\mathbb{C}P^1$-structure C is *circular* if its lift $\tilde{\alpha}$ to the universal cover is immersed to a circular arc $\mathbb{C}\mathrm{P}^1$ by the developing map.

We first define a grafting along a circular arc on a $\mathbb{C}\mathrm{P}^1$-structure. For $\theta > 0$, consider the horizontal biinfinite strip $\mathbb{R} \times [0, \theta i]$ in $\mathbb{C}$ of height θ. Then let R_θ be the $\mathbb{C}\mathrm{P}^1$-structure on the strip whose developing map is the restriction of the exponential map $\exp\colon \mathbb{C} \to \mathbb{C} \setminus \{0\}$. This $\mathbb{C}\mathrm{P}^1$-structure is called the *crescent* of angle θ or simply θ*-crescent*.

Let ℓ be a (biinfinite) circular arc properly embedded in a $\mathbb{C}\mathrm{P}^1$-surface C. Then the *grafting* of C along ℓ by θ is the insertion of this strip R_θ along ℓ (θ*-grafting*), to be precise, as follows: Notice that $C \setminus \ell$ has two boundary components isomorphic to ℓ. Then we take a union of $C \setminus \ell$ and $\mathbb{R} \times [0, \theta i]$ by an isomorphism between $\partial(C \setminus \ell)$ and $\partial(\mathbb{R} \times [0, \theta i])$ so that there is "no shearing", i.e. for each $r \in \mathbb{R}$, the vertical arc $r \times [0, \theta i]$ connects the points of the different boundary components of $C \setminus \ell$ corresponding to the same point of ℓ.

Let ℓ be a circular loop on a projective surface C. We can similarly define a grafting along ℓ by grafting the universal cover $\tilde{C}$ of C in an equivariant manner: Letting $\phi\colon \tilde{C} \to C$ be the universal covering map, $\phi^{-1}(\ell)$ is a union of disjoint circular arcs property embedded in $\tilde{C}$ which is invariant under $\pi_1(S)$.

Then, we insert a θ-crescent along each arc of $\phi^{-1}(\ell)$ as above. By quotienting out the resulting structure by $\pi_1(S)$, we obtain a new $\mathbb{C}\mathrm{P}^1$-structure homeomorphic to C, since a cylinder is inserted to C along ℓ. Indeed, the stabilizer of an arc $\tilde{\ell}$ of $\phi^{-1}(\ell)$ is an infinite cyclic group generated by an element $\gamma \in \pi_1(S)$ whose free homotopy class is ℓ, and the cyclic group $\langle \gamma \rangle$ acts on R_θ so that the quotient is the inserted cylinder (*grafting cylinder of height* θ).

Note that R_θ is foliated by horizontal lines $\mathbb{R} \times \{y\}$, $y \in [0, \theta]$. Then it has a natural transverse measure given by the difference of the second coordinates. This measured foliation descends to a measured foliation on the grafting cylinder. In addition, there is a natural projection $R_\theta \to \mathbb{R}$ to the first coordinate (*collapsing map*). This projection descends to a collapsing map of a grafting cylinder to a circle.

Let $\mathrm{Gr}_{\ell,\theta}(C)$ denote the resulting $\mathbb{C}\mathrm{P}^1$-structure homeomorphic to C. Notice that the holonomy along the circular loop ℓ is hyperbolic, as it has exactly two fixed points on $\mathbb{C}\mathrm{P}^1$ which are the endpoints of the developments of ℓ.

In the case that θ is an integer multiple of 2π, the holonomy C is not changed by the θ-grafting, since the developing map does not change in $\phi^{-1}(C \setminus \ell)$. In particular, the 2π-grafting along a circular loop ℓ inserts a copy of $\mathbb{C}\mathrm{P}^1$ minus a circular arc along each lift of ℓ.

In fact, a 2π-grafting is still well-defined along a more general loop. A loop ℓ on $C = (f, \rho)$ is *admissible* if $\rho(\gamma)$ is hyperbolic and an (equivalently, every) lift $\tilde{\ell}$ of ℓ embeds into $\mathbb{C}\mathrm{P}^1$ by f. Given such a loop, we can insert a copy of $\mathbb{C}\mathrm{P}^1 \setminus (f(\tilde{\ell}) \cup \mathrm{Fix}(\rho(\gamma)))$ along $\tilde{\ell}$, where $\mathrm{Fix}(\rho(\gamma))$ denotes the fixed points of $\rho(\gamma)$. Note that the quotient of $\mathbb{C}\mathrm{P}^1 \setminus \mathrm{Fix}(\rho(\gamma))$ by the infinite cyclic group generated by $\rho(\gamma)$ is a projective structure T on a torus, and the development $f(\tilde{\ell})$ covers a simple loop on T isomorphic to ℓ. By abuse of notation, we also denote the loop on T by ℓ. Then the 2π-grafting of C along ℓ is given by identifying the boundary loops of $C \setminus \ell$

and $T \setminus \ell$ by the isomorphism. Denote by $\mathrm{Gr}_\ell(C)$ the 2π-grafting of C along an admissible loop ℓ.

A *multiloop* is a union of locally finite disjoint simple closed curves. Note that if there is a multiloop M on a projective surface consisting of admissible loops, then a grafting can be done along M simultaneously.

6.4 The Construction of Thurston's Parameters

In this section, we explain the correspondence stated in Theorem A in both directions, following [17].

6.4.1 *The Construction of* $\mathbb{C}\mathrm{P}^1$*-Structures from Measured Laminations on Hyperbolic Surfaces*

Let $(\tau, L) \in \mathsf{T} \times \mathsf{ML}$, where τ is a hyperbolic structure on S, and L a measured geodesic lamination on τ. Then (τ, L) corresponds to the $\mathbb{C}\mathrm{P}^1$-structure on S obtained by grafting τ along L as follows.

Suppose first that L consists of periodic leaves. Then, for each leaf ℓ of L, letting w be its weight, we insert a grafting cylinder of height w, and obtain a projective structure $C = (f, \rho)$ on S. Let $\tilde{L}$ be the pull back of L by the universal covering map. Then there is a ρ-equivariant pleated surface $\beta\colon \mathbb{H}^2 \to \mathbb{H}^3$, obtained by bending $\mathbb{H}^2$ along $\tilde{L}$ by the angles given by the weights.

Let $\kappa : C \to \tau$ be the collapsing map obtained by collapsing all grafting cylinders in C in Sect. 6.3. For each point p in $\tilde{C}$, there is an open neighborhood D, called a maximal disk, such that f embeds D onto a round disk in $\mathbb{C}\mathrm{P}^1$. Then, the boundary of $f(D)$ bounds a hyperbolic plane H_p in $\mathbb{H}^3$. Denote, by $\Psi_p\colon f(D) \to H_p$, the nearest projection. Then $\beta \circ \tilde{\kappa}(p) = \Psi_p \circ f(p)$ for all $p \in \tilde{C}$, where $\tilde{\kappa}\colon \tilde{C} \to \mathbb{H}^2$ is the lift of $\kappa\colon C \to \tau$.

Suppose next that L contains an irrational sublamination. Then, pick a sequence of measured laminations L_i consisting of closed leaves, such that L_i converges to L as $i \to \infty$. Then, for each i, as above there is a $\mathbb{C}\mathrm{P}^1$-structure $C_i = Gr_{L_i}(\tau)$ and a ρ_i-equivariant pleated surface $\beta_i\colon \mathbb{H}^2 \to \mathbb{H}^3$. As L_i converges to L, the surface β_i converges to a pleated surface $\beta\colon \mathbb{H}^2 \to \mathbb{H}^3$ uniformly on compact sets, and therefore C_i converges to a $\mathbb{C}\mathrm{P}^1$-structure on S. (See [6].)

6.4.2 The Construction of Measured Laminations on Hyperbolic Surfaces from $\mathbb{C}\mathrm{P}^1$-Structures

Let $C = (f, \rho)$ be a projective structure on S given by a developing pair. Let $\tilde{C}$ be the universal cover of C.

Identify $\mathbb{C}\mathrm{P}^1$ conformally with a unit sphere $\mathbb{S}^2$ in $\mathbb{R}^3$. Then, each round circle on $\mathbb{C}\mathrm{P}^1$ is the intersection of $\mathbb{S}^2$ with some (affine) hyperplane $\mathbb{R}^2$ in $\mathbb{R}^3$. A *(open) round disk* D in $\tilde{C}$ is an open subset of $\tilde{C}$ homeomorphic to an open disk, such that f embeds D onto an open round disk in $\mathbb{C}\mathrm{P}^1$ (we also say a maximal disk *of* $\tilde{C}$, emphasizing the ambient space for the maximality). A *maximal disk* D in $\tilde{C}$ is a round disk, such that there is no round disk in $\tilde{C}$ strictly containing D. Let D be a maximal disk in $\tilde{C}$. Then the closure of its image, $\overline{f(D)}$, is a closed round disk in $\mathbb{C}\mathrm{P}^1$.

We first see a basic example illustrating the pleated surface corresponding to a $\mathbb{C}\mathrm{P}^1$-structure. (See [9].) Let U be a region of $\mathbb{C}\mathrm{P}^1$ homeomorphic to an open disk such that $\mathbb{C}\mathrm{P}^1 \setminus U$ contains more than one point (i.e. $U \ncong \mathbb{C}\mathrm{P}^1, \mathbb{C}$). Regard $\mathbb{C}\mathrm{P}^1$ as the ideal boundary of hyperbolic three-space $\mathbb{H}^3$, and consider the convex full of $\mathbb{C}\mathrm{P}^1 \setminus U$ in $\mathbb{H}^3$. Then its boundary in $\mathbb{H}^3$ is a hyperbolic plane $\mathbb{H}^2$ bent along a measured lamination L_U [9]. This lamination corresponds to the lamination in the Thurston coordinates.

Let Ψ_U denote the orthogonal projection from U to $\partial\,\mathrm{Conv}(\mathbb{C}\mathrm{P}^1 \setminus U)$. Then, since $\partial\,\mathrm{Conv}(\mathbb{C}\mathrm{P}^1 \setminus U)$ is, in the intrinsic metric, a hyperbolic plane, Ψ yields a continuous map from U to $\mathbb{H}^2$. For each maximal disk D in U, let H_D be the hyperbolic plane in $\mathbb{H}^3$ bounded by its ideal boundary of D. Then H_D intersects $\partial\,\mathrm{Conv}(\mathbb{C}\mathrm{P}^1 \setminus U)$ in either a geodesic or the closure of a complementary region of L_U in $\mathbb{H}^2$. Thus, all maximal disks in U correspond to the *strata* of $(\mathbb{H}^2, L)$, where each stratum is either the closure of a complementary region of L in $\mathbb{H}^2$, a leaf of L with atomic measure, or a leaf of L not contained in the closure of some complementary region. In particular, two distinct complementary regions R_1, R_2 of $(\mathbb{H}^2, L)$ correspond to different maximal disks D_1, D_2, and if R_1 is close enough to R_2, then D_1 intersects D_2. Accordingly, the ideal boundary circles of D_1 and D_2 bound hyperbolic planes intersecting in a geodesic. Then the transverse measure of L_U is, infinitesimally, given by the angles between such hyperbolic planes.

Moreover there is a natural measured lamination $\mathcal{L}_U$ on U which maps to L_U by Ψ_U. If a leaf ℓ has a positive atomic measure $w > 0$, then $\Psi_U^{-1}(\ell)$ is a crescent region R_w of angle w, and R_w is foliated by circular arcs ℓ' which project to ℓ. Then Ψ_U is a homeomorphism in the complement of such foliated crescents, and Ψ_U isomorphically takes $\mathcal{L}_U$ to L_U in the complement (i.e. it preserves leaves and transverse measure). The transverse measure of $\mathcal{L}$ is given by infinitesimal angles between "very close" maximal disks.

As developing maps of $\mathbb{C}\mathrm{P}^1$-structures are, in general, not embedding, we need to find such projections somewhat more "locally" using maximal disks.

Let D be a maximal disk in the universal cover $\tilde{C}$, and let $\overline{D}$ be the closure of D in $\tilde{C}$. In other words, $\overline{D}$ is the connected component of $f^{-1}(\overline{f(D)})$ containing

D. Then $\overline{f(D)} \setminus f(\overline{D})$ is a subset of the boundary circle of the round disk $f(D)$, and the points in this subset are called the *ideal points* of D. (Given a point p of the boundary circle $f(D)$, pick a path $\alpha\colon [0, 1) \to f(D)$ limiting to p as the parameter goes to 1. Then p is an ideal point of D if and only if the lift of α to $\tilde{C}$ leaves every compact subset of $\tilde{C}$.)

Let $\partial_\infty D \subset \mathbb{C}P^1$ denote the set of all ideal points of D. As $f|D$ is an embedding onto a round disk, we regard $\partial_\infty D$ as a subset of the boundary circle of D abstractly (not as a subset of $\mathbb{C}P^1$). Then $\partial_\infty D$ is a closed subset of $\mathbb{S}^1$, since its complement is open. Identifying D with a hyperbolic disk conformally, we let $\mathrm{Core}(D) = \mathrm{Core}_{\tilde{C}}(D)$ be the convex hull of $\partial_\infty D$.

For each point p of $\tilde{C}$, there is a round disk containing p, and moreover, as C is not $\mathbb{C}P^1$ or $\mathbb{C}$, there is a maximal disk containing p. The *canonical neighborhood* U_p of C is the union of all maximal disks D_j $(j \in J)$ in $\tilde{C}$ which contain p.

In fact, $(C, \mathcal{L})$ completely determines the Thurston parameters (τ, L). Furthermore the Thurston parameters near $p \in \mathbb{C}P^1$ are determined by the Thurston parameters of its canonical neighborhood, in the way similar to the region U in $\mathbb{C}P^1$ homeomorphic to a disk as above. Namely Lemma 6.4.3 below implies that the Thurston lamination on $\tilde{C}$ near p is determined by the canonical neighborhood U_p of p, and the following Proposition states that U_p is a topological disk embedded in $\mathbb{C}P^1$.

Proposition 6.4.1 ([17], Proposition 4.1) *For every point p in $\tilde{C}$, $f\colon \tilde{S} \to \mathbb{C}P^1$ embeds its canonical neighborhood U_p into $\mathbb{C}P^1$. Moreover U_p is homeomorphic to an open disk.*

Proof Set $U_p = \cup D_j$, where D_j are maximal disks in $\tilde{C}$ containing p. Let x, y be distinct points in U_p; let D_x and D_y be maximal disks containing $\{p, x\}$ and $\{p, y\}$, respectively. By the definition of maximal disks, f embeds D_i and D_j onto round disks in $\mathbb{C}P^1$. Then $f(D_i) \cap f(D_j) = f(D_i \cap D_j)$ is either a crescent or a *round annulus*, i.e. a region in $\mathbb{C}P^1$ bounded by disjoint round circles. If it is a round annulus, then $f|D_i \cup D_j$ must be a homeomorphism onto $\mathbb{C}P^1$ and $D_i \cup D_j = \tilde{S}$, which cannot occur. Thus $f(D_i \cap D_j)$ is a crescent, and therefore f is injective on $D_x \cup D_y$. Hence $f(x) \neq f(y)$, and f embeds U_p into $\mathbb{C}P^1$.

The image $f(U_p)$ is not surjective (as S is not homomorphic to a sphere). Thus we can normalize $\mathbb{C}P^1 = \mathbb{C} \cup \{\infty\}$ so that $p = \infty$ and $0 \notin U_p$. Then $\mathbb{C}P^1 \setminus U_p$ is the intersection of the closed disks $\mathbb{C}P^1 \setminus D_j$ containing 0. Thus $\mathbb{C}P^1 \setminus U_p$ is a closed convex subset containing 0, and therefore U_p is topologically an open disk. □

Then in the setting of Proposition 6.4.1, we have

Corollary 6.4.2 *When $\mathbb{C}P^1 = \mathbb{C} \cup \{\infty\}$ is normalized so that $p = \{\infty\}$, the complement $\mathbb{C}P^1 \setminus U_p$ is a compact convex subset K of $\mathbb{C}$.*

The canonical neighborhood U_p can be regarded as a projective structure on an open disk (Proposition 6.4.1), and one can consider maximal disks in U_p, which are a priori unrelated maximal disks in $\tilde{C}$.

Lemma 6.4.3 *The maximal disks of U_p bijectively correspond to the maximal disks of $\tilde{C}$ whose closure contain p by the inclusion $U_p \subset \tilde{C}$.*

Moreover, if D is a maximal disk of U_p containing p, then the ideal points of D as a maximal disk of U_p coincide with its ideal points of D as a maximal disk of $\tilde{C}$.

Proof If D is a maximal disk in $\tilde{C}$ containing p, then, by the definition of U_p, clearly D is also a maximal disk in U_p. Similarly, if D is a maximal disk in $\tilde{C}$ whose boundary contains p, then there is a sequence of maximal disks D_i containing p with $D_i \to D$ as $i \to \infty$. Therefore every maximal disk D in $\tilde{C}$ whose closure contains p is a maximal disk in U_p.

We show the opposite inclusion. By Corollary 6.4.2, the complement $K = \mathbb{C}\mathrm{P}^1 \setminus U_p$ is a closed compact convex subset of C. If there is a (straight) line ℓ in $\mathbb{C}$ such that $\ell \cap K$ is a single point x, then, by the inclusions $\tilde{C} \supset U_p \subset \mathbb{C}\mathrm{P}^1$, x corresponds to an ideal point of a maximal ball of $\tilde{C}$ containing p. Next suppose that there is a line ℓ in $\mathbb{C}$ such that $\ell \cap K$ is a line segment. Then, letting P be the half-plane bounded by ℓ so that P and K have disjoint interiors, there is a sequence of maximal disks D_i of $\tilde{C}$ containing p such that D_i converges to P as $i \to \infty$. Thus the endpoints of the line segment correspond to ideal points of $\tilde{C}$.

Suppose that D is a maximal disk of U_p. Then $\overline{D}$ intersects K in ∂K. If ∂D intersects K in a line segment, then D is a half-plane in $\mathbb{C}$ with ∂D containing p. As the endpoints of the segment correspond to the ideal points of $\tilde{C}$, D is also a maximal disk in $\tilde{C}$.

If the closure of D does not intersect K in a line segment, then clearly D contains p. If a point on ∂K is not an interior point of any line segment of ∂K, then the point corresponds to an ideal point of $\tilde{C}$. Therefore no round disk in $\tilde{C}$ strictly contains D, and therefore D is also a maximal ball in $\tilde{C}$. Thus we have the opposite inclusion.

Finally, suppose that D is a maximal ball in U_p containing p. Then $\overline{D} \cap K$ contains no line segment, and therefore, $\overline{D} \cap K$ corresponds to the ideal points of D as a maximal ball in $\tilde{C}$. □

The following proposition yields a lamination on $\tilde{C}$ invariant under $\pi_1(S)$.

Proposition 6.4.4 ([17], Theorem 4.4) *The cores* $\mathrm{Core}(D)$ *of the maximal disks D in $\tilde{C}$ are all disjoint and their union is $\tilde{C}$.*

Proof We first show that the cores are disjoint. Let D_1 and D_2 be distinct maximal disks in $\tilde{C}$. If $D_1 \cap D_2 \neq \emptyset$, then $f(D_1)$ and $f(D_2)$ are round disks intersecting a crescent. Therefore $\mathrm{Core}(D_1)$ and $\mathrm{Core}(D_2)$ are disjoint. (Consider the circular arc in D_1 orthogonal to ∂D_1; then, indeed, this arc separates $\mathrm{Core}(D_1)$ and $\mathrm{Core}(D_2)$ in $D_1 \cup D_2$.) □

Claim 6.4.5 *Given a convex subset V of $\mathbb{C}$, there is a unique round disk D in $\mathbb{C}$ of minimal radius containing V.*

Proof Suppose, on the contrary, that there are two different round disks D_1, D_2 containing V which attain the minimal radius. Then, clearly, there is a round disk D_3 of strictly smaller radius which contains V (such that $D_3 \supset D_1 \cap D_2$ and $D_3 \subset D_1 \cup D_2$). This is a contradiction. □

Claim 6.4.6 *The convex hull of $\partial D \cap \overline{V}$ contains the center of c with respect to the complete hyperbolic metric on $D(\cong \mathbb{H}^2)$ given by the conformal identification.*

Proof Suppose not; then the closure of V is contained in the interior of a (Euclidean) half disk of D. Then one can easily find a round disk of smaller radius containing $\overline{V}$. □

Note that, the inversion of $\mathbb{C}\mathrm{P}^1 = \mathbb{C} \cup \{\infty\}$ about ∂D exchanges ∞ and the center of D, and it fixes ∂D pointwise. Then, by Claim 6.4.6, in the (conformal) hyperbolic metric on D, the convex hull of $\partial D \cap \overline{V}$ contains the center of D. Therefore, by the inversion, in the hyperbolic metric on $\mathbb{C}\mathrm{P}^1 \setminus D$, the point at ∞ is contained in the convex hull of $\partial D \cap \overline{V}$ in the interior of $\mathbb{C}\mathrm{P}^1 \setminus D$.

Using the above claims, we show that, for every $x \in \tilde{C}$, there is a maximal disk D in $\tilde{C}$ whose core contains x. Let $U_x = \cup_{j \in J} D_j$ be the canonical neighborhood of x, where D_j are the maximal disks in $\tilde{C}$ which contain x. Normalize $\mathbb{C}\mathrm{P}^1$ so that $f(x) = \infty$. Let $D_j^c = \mathbb{C}\mathrm{P}^1 \setminus f(D_j)$. Then $\mathbb{C}\mathrm{P}^1 \setminus f(U_x) = \cap_j D_j^c$. By Claim 6.4.6, let D be the maximal disk of U_x such that $x \in \mathrm{Core}_{U_x}(D)$. By Lemma 6.4.3, D is also a maximal disk of $\tilde{C}$ which contains x, and moreover the ideal points of D as a maximal disk of U_x coincide with those as a maximal ball of $\tilde{C}$. Then, $\mathrm{Core}_{\tilde{C}}(D)$ contains x. 6.4.4

By Proposition 6.4.4, $\tilde{C}$ is canonically decomposed into the cores of maximal disks in $\tilde{C}$, which yields a *stratification* of $\tilde{C}$. Note that this decomposition is invariant under $\pi_1(S)$, as the maximal balls and ideal points are preserved by the action. Moreover, for each maximal disk D in $\tilde{C}$, $\mathrm{Core}(D)$ is properly embedded in $\tilde{C}$. Then the one-dimensional cores and the boundary components of two-dimensional cores form a $\pi_1(S)$-invariant lamination $\tilde{\lambda}$ on $\tilde{C}$, which descends to a lamination λ on C.

Next we see that the angles between infinitesimally close maximal disks yield a natural transverse measure of this lamination. Given a point $x \in \tilde{C}$, let D_x be the maximal disk in $\tilde{C}$ whose core contains x. If $y \in \tilde{C}$ is sufficiently close to x, then D_y intersects D_x. Then let $\angle(D_x, D_y)$ denote the angle between the boundary circles of D_x and D_y. To be precise, this is the angle of the crescent $D_x \setminus D_y$ (or $D_y \setminus D_x$) at the vertices. Then $\angle(D_x, D_y) \to 0$ as $y \to x$.

Let x and y be distinct points of $\tilde{C}$ contained in different strata of $(\tilde{C}, \tilde{\lambda})$. Then pick a path $\alpha\colon [0, 1] \to \tilde{C}$ connecting x to y such that α is transverse to $\tilde{\lambda}$. Let $\Delta : 0 = t_0 < t_1 < \cdots < t_n = 1$ be a finite division of $[0, 1]$, and let $x_i = \alpha(t_i)$ for each $i = 0, \ldots, n$. Let $|\Delta| = \min_{i=0}^{n-1}(x_{i+1} - x_i)$, the smallest width of the subintervals. Then, let $\Theta(\Delta) = \Sigma_{i=1}^{n-1} \angle(D_{x_i}, D_{x_{i+1}})$ for a subdivision Δ of $[0, 1]$ with sufficiently small $|\Delta|$. Pick a sequence of subdivisions Δ_i such that $|\Delta_i| \to 0$ as $i \to \infty$. Then $\lim_{i\to\infty}(\Theta(\Delta_i))$ exists and it is independent of the choice of Δ_i as $i \to \infty$ [6, II.1]. We define the transverse measure of α to be $\lim_{i\to\infty}(\Theta(\Delta_i))$. Then $\tilde{\lambda}$ with this transverse measure yields a measured lamination $\tilde{\mathcal{L}}$ invariant under $\pi_1(S)$. Thus $\tilde{\mathcal{L}}$ descends to a measured lamination $\mathcal{L}$ on C.

By Lemma 6.4.3, for every $x \in \tilde{C}$, the measured lamination $\mathcal{L}$ near x is determined by the canonical neighborhood U_x of x. Let $\mathcal{L}_x$ be the measured lamination on U_x, which descends to the measured lamination on the boundary of $\mathrm{Conv}(\mathbb{C}\mathrm{P}^1 \setminus U_x)$. Then there is a neighborhood V of x in U_x such that $\mathcal{L}$ and $\mathcal{L}_x$ coincide in V by the inclusion $U_x \subset \tilde{C}$.

For each point $x \in \tilde{C}$, the boundary circle of the maximal disk D_x bounds a hyperbolic plane H_x in $\mathbb{H}^3$. Let $\Psi_x : f(D_x) \to H_x$ be the projection along geodesics in $\mathbb{H}^3$ orthogonal to H_x. Then H_x has a canonical normal direction pointing to D_x. By Lemma 6.4.3 there is a neighborhood V of x, such that $\Psi_y(y) = \Psi_x(y)$. Moreover, Ψ_x coincides with the projection onto the boundary pleated surface of $\mathrm{Conv}(\mathbb{C}\mathrm{P}^1 \setminus U_x)$. Therefore, as in the case of regions in $\mathbb{C}\mathrm{P}^1$, we have a pleated surface $\mathbb{H}^2 \to \mathbb{H}^3$ which is ρ-equivariant, as in the following paragraph.

We assume that crescents R in $\tilde{C}$ are always foliated by leaves of $\tilde{\mathcal{L}}$ sharing their endpoints at the vertices of R. We have a well-defined continuous map $\Psi\colon \tilde{C} \to \mathbb{H}^3$ defined by $\Psi(x) = \Psi_x(x)$. We shall take an appropriate quotient of $\tilde{C}$ to turn it into a hyperbolic plane. For each crescent R in $\tilde{C}$, Ψ takes each leaf in R to the geodesic in $\mathbb{H}^3$ connecting the vertices of R. Identify $x, y \in \tilde{C}$, if x, y are contained in a single crescent in $\tilde{C}$ and $\Psi_x(x) = \Psi_y(y)$; let $\tilde{\kappa}\colon \tilde{C} \to \tilde{C}/\sim$ be the quotient map by this identification, which collapses each foliated crescent region to a single leaf. Then by the equivalence relation, $\Psi\colon \tilde{C} \to \mathbb{H}^3$ induces a continuous map $\beta\colon (\tilde{C}/\sim) \to \mathbb{H}^3$ such that $\Psi_x(x) = \beta \circ \kappa$. Moreover, $\tilde{C}/\sim$ is $\mathbb{H}^2$ with respect to the path metric in $\mathbb{H}^3$ via Ψ, since, for every $x \in \tilde{C}$, Ψ coincides with the projection $U_x \to \partial\,\mathrm{Conv}(\mathbb{C}\mathrm{P}^1 \setminus U_x)$ in a neighborhood of x. Thus we have a ρ-equivariant pleated surface $\mathbb{H}^2 \to \mathbb{H}^3$.

The measured lamination $\tilde{\mathcal{L}}$ on $\tilde{C}$ descends to a measured lamination $\tilde{L}$ on $\mathbb{H}^2$ invariant under $\pi_1(S)$. By taking the quotient, we obtain a desired pair (τ, L) of a hyperbolic surface τ and a measured geodesic lamination L on τ.

Similarly, the collapsing map $\tilde{\kappa}\colon \tilde{C} \to \mathbb{H}^2$ descends to a *collapsing map* $\kappa\colon C \to \tau$. Then, for each periodic leaf ℓ of L, $\kappa^{-1}(\ell)$ is a grafting cylinder foliated by closed leaves of $\mathcal{L}$.

Finally we note that as $\beta\colon \mathbb{H}^2 \to \mathbb{H}^3$ is obtained by bending $\mathbb{H}^2$ in $\mathbb{H}^3$ along $\tilde{L}$, the pair (τ, L) corresponds to C by the correspondence in Sect. 6.4.1.

6.5 Goldman's Theorem on Projective Structures with Fuchsian Holonomy

Let C be a $\mathbb{C}\mathrm{P}^1$-structure on S with holonomy ρ, and let $(\tau, L) \in \mathsf{T} \times \mathsf{ML}$ be its Thurston parameters. Let $\psi\colon \mathbb{H}^2 \to \tau$ be the universal covering map, and $\tilde{L}$ be the measured lamination $\psi^{-1}(L)$ on $\mathbb{H}^2$. Let $\Gamma = \mathrm{Im}\rho$, and let Λ denote the limit set of $\mathrm{Im}\rho$.

Lemma 6.5.1 *Let $\beta\colon \mathbb{H}^2 \to \mathbb{H}^3$ be the associated pleated surface, where $\mathbb{H}^2$ is the universal cover of τ. Then, for every leaf $\tilde{\ell}$ of $\tilde{L}$, $\beta|\tilde{\ell}$ is a geodesic connecting different points of Λ.*

Proof If $\tilde{\ell}$ is a lift of a closed leaf of L, then the assertion clearly holds.

For every closed curve α on τ, let $\tilde{\alpha}$ be a lift of α to $\mathbb{H}^2$. Since the curve $\beta|\tilde{\alpha}$ is preserved by the hyperbolic element $\rho(\alpha)$, it is a quasi-geodesic in $\mathbb{H}^3$ whose endpoints are the fixed points of $\rho(\alpha)$. Note that the endpoints are contained in Λ.

Let ℓ be a non-periodic leaf of L, and let $\tilde{\ell}$ be a lift of ℓ to $\mathbb{H}^2$. There is a sequence of simple closed geodesics ℓ_i on τ such that ℓ_i converges to ℓ in the Hausdorff topology [6, I.4.2.14]. For each $i \in \mathbb{N}$, pick a lift $\tilde{\ell}_i$ of ℓ_i to $\mathbb{H}^2$ so that $\tilde{\ell}_i \to \ell$ uniformly on compact sets as $i \to \infty$. Then, $\beta|\tilde{\ell}_i$ converges to $\beta|\tilde{\ell}$ uniformly on compact sets. Moreover as $\angle_{\tau_i}(\tau_i, L_i) \to 0$, $\beta_i|\tilde{\ell}_i$ is asymptotically an isometric embedding: To be precise, for large enough i, it is a bilipschitz embedding, and its bilipschitz constant converges to 1 as $i \to \infty$ [3, Proposition 4.1].

As ℓ_i are closed loops, the endpoints of $\beta|\tilde{\ell}_i$ are in Λ. Then the endpoints of $\beta|\tilde{\ell}_i$ converge to the endpoints of $\beta|\tilde{\ell}$ in $\mathbb{C}\mathrm{P}^1$. Therefore, since Λ is a closed subset of $\partial\mathbb{H}^3$, the endpoints of $\beta|\ell$ are also contained in Λ. □

We immediately have the following.

Corollary 6.5.2 *For each stratum σ of $(\mathbb{H}^2, \tilde{L})$, let $D_\sigma \subset \tilde{C}$ be the maximal disk whose core corresponds to σ. Then its ideal points $\partial_\infty D_\sigma$ are contained in the limit set Λ.*

We reprove the following theorem by means of pleated surfaces.

Proposition 6.5.3 (See [21, Theorem 3.7.3.]) *Let C be a $\mathbb{C}\mathrm{P}^1$-structure with real holonomy $\rho\colon \pi_1(S) \to \mathrm{PSL}(2, \mathbb{R})$ and (L, τ) its Thurston parameters. Then each leaf of L is periodic, and its weight is a π-multiple. If ρ is, in addition, Fuchsian, then each leaf of L is periodic and its weight is a 2π-multiple.*

Proof We first show that L consists of periodic leaves. Suppose, on the contrary, that L contains an irrational minimal sublamination N. Then the transverse measure is continuous in a neighborhood of $|N|$ in τ (i.e. no leaf of N has an atomic measure).

Thus there are two-dimensional strata $\sigma, \sigma_1, \sigma_2, \ldots$ of $\mathbb{H}^2 \setminus \tilde{L}$, such that σ_i converges to an edge of σ as $i \to \infty$. Note that, as it is two-dimensional, each $\beta(\sigma_i)$ has at least three ideal points, which lie in a round circle in $\mathbb{C}\mathrm{P}^1$. Let $H, H_1, H_2, \ldots$ be the supporting oriented hyperbolic planes in $\mathbb{H}^3$ of $\sigma, \sigma_1, \ldots$. Let $\angle_{\mathbb{H}^3}(H, H_i) \in [0, \pi]$ be the angle between the hyperbolic planes H and H_i with respect to their orientations, if H and H_i intersect. Then, by continuity, $\angle_{\mathbb{H}^3}(H, H_i) \to 0$ as $i \to \infty$. Thus the ideal points of σ and σ_i cannot be contained in a single round circle if i is sufficiently large. By Corollary 6.5.2, this cannot happen as Λ is a single round circle.

We first show that the weight of each leaf of L is a multiple of π. Let σ_1 and σ_2 be components of $\mathbb{H}^2 \setminus \tilde{L}$ adjacent along a leaf of $\tilde{L}$. Let H_1 and H_2 be the support planes of σ_1 and σ_2, respectively. Then the angle between H_1 and H_2 is the weight of ℓ. As the ideal points of σ_1 and σ_2 must lie in the round circle Σ, the angle must be a multiple of π.

Suppose, in addition, that ρ is Fuchsian. Let $\beta_0 : \mathbb{H}^2 (= \tilde{\tau}) \to \mathbb{H}^3$ be the ρ-equivariant embedding onto the hyperbolic plane H_Λ bounded by Λ. For each $i = 1, 2$, as each boundary component m of σ_i covers a periodic leaf of L, $\beta = \beta_0$ on m. Therefore $H_1 = H_2 = \mathrm{Conv}(\Lambda)$, and $\beta_0 = \beta$ on σ_i for each $i = 1, 2$. As the orientation of H_1 coincides with that of H_2, the weight of m must be a multiple of 2π. □

Proof of Theorem B By Proposition 6.5.3, L is a union of closed geodesics ℓ with 2π-multiple weights. For each (closed) leaf ℓ of L, let $2\pi n_\ell$ denote the weight of ℓ, where n_ℓ is a positive integer. Let $\kappa : C \to \tau$ be the collapsing map. Then, $\kappa^{-1}(\ell)$ is a grafting cylinder of height $2\pi n_\ell$, the structure inserted by 2π-grafting n times. Therefore, C is obtained by grafting along a multiloop corresponding to L. □

6.6 The Path Lifting Property in the Domain of Discontinuity

Let $C = (f, \rho)$ be a $\mathbb{C}\mathrm{P}^1$-structure on S. Then, let Λ be the limit set of $\mathrm{Im}\rho$, and let $\Omega = \mathbb{C}\mathrm{P}^1 \setminus \Lambda$, the domain of discontinuity.

Proposition 6.6.1 *For every $x \in \Omega$, there is a neighborhood V_x in Ω such that, for every $y \in \tilde{S}$ with $f(y) \in V_x$, V_x is contained in the maximal disk whose core contains x.*

Proof The union $\mathbb{H}^3 \cup \partial\mathbb{H}^3$ is a unit ball in the Euclidean space and the visual distance is the restriction of the Euclidean metric.

Suppose, on the contrary, that there is no such neighborhood V_x. Then there is a sequence $x_1, x_2, \dots \in f^{-1}(x)$ such that, letting $H_1, H_2, \dots$ be their corresponding hyperbolic support planes, the visual distance from H_i to x goes to zero as $i \to \infty$. Let $y_i \in \mathbb{H}^3$ be the nearest point projection of $f(x_i)$ to H_i. Then, $y_i \to x$ in the visual metric. Let σ_i be the stratum of $(\mathbb{H}^2, \tilde{L}_i)$ which contains $\tilde{\kappa}(x_i)$. Then, as the orthogonal projection of $f(x_i)$ to H_i is y_i, the visual distance between x and $\beta_i(\sigma_i)$ goes to zero as $i \to \infty$. Therefore, there is an ideal point p_i of $\beta(\sigma_i)$ which converges to x as $i \to \infty$. As Ω is open, this is a contradiction by Corollary 6.5.2. □

As f embeds maximal disks of $\tilde{C}$ into $\mathbb{C}\mathrm{P}^1$, we immediately have the following.

Corollary 6.6.2 *For each point $x \in \Omega$, there is a neighborhood V_x of x such that, if $f(y) \in V_x$ for $y \in \tilde{S}$, then f embeds a neighborhood W_y of y in $\tilde{S}$ homomorphically onto V_x.*

Theorem C immediately follows from the corollary.

Acknowledgments I thank Gye-Seon Lee and the mathematics department of Universität Heidelberg for their hospitality where much of this chapter was written. The author is supported by JSPS Grant-in-Aid for Research Activity start-up (18H05833) and Grant-in-Aid for Scientific Research C (20K03610). He also thanks the anonymous referee and the editor for comments, which in particular made some parts of the arguments clearer.

References

1. D. Allegretti, T. Bridgeland, The monodromy of meromorphic projective structures (2018). Preprint. arXiv:1802.02505v1
2. S. Baba, Complex projective structures with Schottky holonomy. Geom. Funct. Anal. **22**(2), 267–310 (2012)
3. S. Baba, 2π–grafting and complex projective structures, I. Geom. Topol. **19**(6), 3233–3287 (2015)
4. S. Baba, 2π-grafting and complex projective structures with generic holonomy. Geom. Funct. Anal. **27**(5), 1017–1069 (2017)
5. G. Calsamiglia, B. Deroin, S. Francaviglia, The oriented graph of multi-graftings in the Fuchsian case. Publ. Mat. **58**(1), 31–46 (2014)
6. R.D. Canary, D.B.A. Epstein, P. Green, Notes on notes of Thurston, in *Analytical and Geometric Aspects of Hyperbolic Space (Coventry/Durham, 1984)*. London Mathematical Society Lecture Note Series, vol. 111 (Cambridge University Press, Cambridge, 1987), pp. 3–92
7. S. Choi, H. Lee, Geometric structures on manifolds and holonomy-invariant metrics. Forum Math. **9**(2), 247–256 (1997)
8. D. Dumas, Complex projective structures, in *Handbook of Teichmüller Theory. Vol. II*. IRMA Lectures in Mathematics and Theoretical Physics, vol. 13 (European Mathematical Society, Zürich, 2009), pp. 455–508
9. D.B.A. Epstein, A. Marden, Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces, in *Analytical and Geometric Aspects of Hyperbolic Space (Coventry/Durham, 1984)*. London Mathematical Society Lecture Note Series, vol. 111 (Cambridge University Press, Cambridge, 1987), pp. 113–253
10. D. Gallo, M. Kapovich, A. Marden, The monodromy groups of Schwarzian equations on closed Riemann surfaces. Ann. Math. (2) **151**(2), 625–704 (2000)
11. W.M. Goldman, Projective structures with Fuchsian holonomy. J. Differential Geom. **25**(3), 297–326 (1987)
12. W.M. Goldman, Geometric structures on manifolds (2018). Available at http://www.math.umd.edu/~wmg/gstom.pdf
13. S. Gupta, M. Mj, Meromorphic projective structures, grafting and the monodromy map (2019). Available at https://arxiv.org/abs/1904.03804
14. D.A. Hejhal, Monodromy groups and linearly polymorphic functions. Acta Math. **135**(1), 1–55 (1975)
15. Y. Kamishima, S.P. Tan, Deformation spaces on geometric structures, in *Aspects of Low-dimensional Manifolds*. Advanced Studies in Pure Mathematics, vol. 20 (Kinokuniya, Tokyo, 1992), pp. 263–299
16. M. Kapovich, *Hyperbolic Manifolds and Discrete Groups*. Progress in Mathematics, vol. 183 (Birkhäuser Boston Inc., Boston, 2001)
17. R.S. Kulkarni, U. Pinkall, A canonical metric for Möbius structures and its applications. Math. Z. **216**(1), 89–129 (1994)
18. B. Maskit, On a class of Kleinian groups. Ann. Acad. Sci. Fenn. Ser. A I No. **442**, 8 (1969)
19. L. Ruffoni, Bubbling complex projective structures with quasi-Fuchsian holonomy. J. Topol. Anal. (2019)

20. D. Sullivan, W. Thurston, Manifolds with canonical coordinate charts: some examples. Enseign. Math. (2) **29**(1–2), 15–25 (1983)
21. S.P. Tan, Representations of surface groups into $\mathrm{PSL}(2, \mathbb{R})$ and geometric structures. PhD thesis, 1988

Chapter 7
A Short Proof of an Assertion of Thurston Concerning Convex Hulls

Graham Smith

Abstract Let X be a closed subset of the ideal boundary $\partial_\infty\mathbb{H}^3$ of 3-dimensional hyperbolic space $\mathbb{H}^3$ and let K be its convex hull in $\mathbb{H}^3$. We provide a short proof of the fact that the topological boundary ∂K of K is intrinsically hyperbolic.

AMS Classification: 30F60

7.1 Introduction

In this chapter, we study the intrinsic geometries of the boundaries of convex hulls in space-forms. For m a positive integer and $\kappa \in \{-1, 0, 1\}$, let M_κ^m denote the m-dimensional space-form of constant sectional curvature equal to κ. We will say that a closed subset K of M_κ^m is *convex* whenever any length-minimising geodesic arc whose two extremities lie in K is also wholly contained in K. Observe that, in the positive-curvature case, with convexity defined in this manner, a convex subset of M_1^m is either contained in an open hemisphere or is equal to the whole of M_1^m. Finally, given a closed subset X of M_κ^m, we define its *convex hull*, denoted by $\mathrm{Conv}(X)$, to be the intersection of all closed, convex subsets of M_κ^m containing X.

Let K now be a convex subset with non-trivial interior of some space-form M_κ^m. Let ∂K denote its set-theoretic boundary. The intrinsic metric (distance function) of ∂K is defined by

$$d(x, y) := \mathop{\mathrm{Inf}}_{\gamma} \mathrm{Length}(\gamma),$$

where γ varies over all rectifiable curves $\gamma : [0, 1] \to \partial K$ with $\gamma(0) = x$ and $\gamma(1) = y$. Since ∂K is everywhere locally a Lipschitz graph (see, for example,

G. Smith (✉)
Instituto de Matemática, Universidade Federal do Rio de Janeiro, Cidade Universitária, Ilha de Fundão, Rio de Janeiro, RJ, Brazil
e-mail: g.a.c.smith.95@cantab.net

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_7

Theorem 4.12 of [7]), the topology generated over this subset by d coincides with the topology that it inherits from M^m_κ.

Suppose now that K is the convex hull of some closed subset X. In this case, the set $\partial K\setminus X$ is known to satisfy at every point x the *local geodesic property* (c.f. Section 4.5 of [7] and Chapter 8 of [8]), namely, that there exists an *open* geodesic segment $\gamma :]-\epsilon,\epsilon[\rightarrow \partial K\setminus X$ such that $\gamma(0) = x$. Furthermore, this property characterises convex hulls (see Theorem 4.18 of [7]).

Having established these preliminaries, we now consider the case where the ambient space is 3-dimensional, so that ∂K is 2-dimensional. If $\partial K\setminus X$ were smooth, then the local geodesic property would make this surface extrinsically flat, and thus instrinsically everywhere locally isometric to M^2_κ. In Chapter 8 of [8], Thurston argues heuristically to show that this property continues to hold even in the non-smooth case. He then explains, furthermore, that the canonical embedding of $\partial K\setminus X$ into M^3_κ is totally geodesic except over a singular set given by the union of disjoint geodesics. These observations play a key role in Thurston's approach to Teichmüller theory by providing a bridge between hyperbolic geometry, on the one hand, and the theory of measured geodesic laminations, on the other. In this chapter, we provide a new proof of Thurston's result which is both shorter and more direct than those currently available in the literature (see [3] and [6]).

We show

Theorem 7.1.1 *Let X be a closed subset of* M^3_κ. *If Conv*(X) *has non-trivial interior, then* $\Sigma := \partial Conv(X)\setminus X$ *is everywhere locally isometric to* M^2_κ. *Furthermore, the canonical embedding of* Σ *into* M^3_κ *is totally geodesic except over a closed set which is a union of disjoint geodesic segments.*

When $\kappa = -1$, M^3_κ is 3-dimensional hyperbolic space $\mathbb{H}^3$. Recall that the ideal boundary $\partial_\infty\mathbb{H}^3$ of $\mathbb{H}^3$ is defined to be the space of equivalence classes of oriented geodesic rays (see [1]). A more useful description of the ideal boundary for our current applications is given by the Kleinian parametrisation, which maps $\mathbb{H}^3$ onto the open unit ball $B_1(0)$ in $\mathbb{R}^3$ in such a manner as to send hyperbolic geodesics to straight lines. With this parametrisation, the ideal boundary of $\mathbb{H}^3$ identifies with the unit sphere $S_1(0)$ in $\mathbb{R}^3$, and the convex hull in $\mathbb{H}^3$ of any given subset X of $S_1(0)$ likewise identifies with its convex hull in $B_1(0)$. Theorem 7.1.1 now yields

Theorem 7.1.2 *Let X be a closed subset of* $\partial_\infty\mathbb{H}^3$. *Let Conv*$(X)$ *denote its convex hull in* $\mathbb{H}^3$. *If Conv*(X) *has non-trivial interior, then* $\Sigma := \partial Conv(X)\setminus X$ *is everywhere locally isometric to* $\mathbb{H}^2$. *Furthermore, the canonical embedding of* Σ *into* $\mathbb{H}^3$ *is totally geodesic except over a closed set which is a union of complete, non-intersecting geodesics.*

7.2 Convex Subsets Viewed Extrinsically

Consider first an arbitrary metric space $(Y,\ \delta)$. In what follows, for any subset X of Y and for any $r > 0$, we will denote by $B_r(X)$ the open neighbourhood of radius r about X and, in the case where $X = \{x\}$ consists of a single point, we will write $B_r(x)$ instead of $B_r(X)$. Let $\mathrm{CB}(Y)$ denote the set of closed, bounded subsets of Y. Recall (see Section 45 of [5]) that the *Hausdorff metric* is defined over this set by

$$d_{\mathrm{H}}(X_1,\ X_2) = \mathrm{Inf}\{r > 0 \mid X_1 \subseteq B_r(X_2)\ \&\ X_2 \subseteq B_r(X_1)\}.$$

Recall also that the metric space $(\mathrm{CB}(Y), d_{\mathrm{H}})$ is compact (resp. complete) if and only if Y is compact (resp. complete).

Suppose now that $Y = \mathbb{R}^m$ is m-dimensional euclidean space. Consider the set $\mathrm{CC}(Y)$ of compact, convex subsets of Y. Observe that $\mathrm{CC}(Y)$ is a closed subset of $\mathrm{CB}(Y)$ and that, furthermore, the operator Conv defines a projection from $\mathrm{CB}(Y)$ onto this subset. In this section, we study the topological properties of these objects. Our results will also extend to convex subsets of arbitrary space-forms via affine charts.

Recall now that a closed *half-space* in $\mathbb{R}^m$ is a subset of the form

$$H_{\alpha,\lambda} := \{y \mid \alpha(y) \leq \lambda\},$$

where $\alpha : \mathbb{R}^m \rightarrow \mathbb{R}$ is a linear form and $\lambda > 0$ is a real number. It is straightforward to show (see Theorem 5.2 of [7]) that the convex hull of X is also the intersection of all closed half-spaces containing X.

Lemma 7.2.1 *Let (X_n) be a sequence of compact subsets of $\mathbb{R}^m$. If this sequence converges in the Hausdorff sense to the compact subset X_∞, then the sequence (Conv(X_n)) of convex hulls also converges in the Hausdorff sense to Conv(X_∞). In other words, Conv maps $CB(\mathbb{R}^m)$ continuously onto $CC(\mathbb{R}^m)$.*

Proof For all $n \in \mathbb{N} \cup \{\infty\}$, denote $K_n := \mathrm{Conv}(X_n)$. First observe that there exists $R > 0$ such that, for all n, $X_n \subseteq B_R(0)$ so that, in particular, $K_n \subseteq B_R(0)$. Since the set of compact, convex subsets of the closed ball $\overline{B}_R(0)$ is compact in the Hausdorff topology, it suffices to show that K_∞ is the only concentration point of the sequence (K_n) in this topology. Suppose therefore that another such concentration point K'_∞ exists. In particular, K'_∞ is compact and convex and $X_\infty \subset K'_\infty$ so that, by definition of the convex hull, $K_\infty \subseteq K'_\infty$. Let x be a point of $K'_\infty \setminus K_\infty$. Let $H_{\alpha,\lambda}$ be a half-space which contains K_∞ but which does not contain x. Let $\epsilon > 0$ be such that $\alpha(x) = \lambda + 2\epsilon$. Since X_∞ is contained in $H_{\alpha,\lambda}$, for all sufficiently large n, X_n is contained in $H_{\alpha,\lambda+\epsilon}$. In particular, for all such n, K_n is also contained in $H_{\alpha,\lambda+\epsilon}$, so that $K_n \cap B_\epsilon(x) = \emptyset$. This is absurd, since some subsequence of (K_n) converges to K_∞ in the Hausdorff sense, and the result follows. □

Consider now a point $x \in \mathbb{R}^m$. For $\epsilon, r > 0$, we say that the subset X of $\mathbb{R}^m$ is *ϵ-dense* in $B_r(x)$ whenever every point of $B_r(x)$ lies at a distance of less than ϵ from X.

Lemma 7.2.2 *If X is a compact, convex subset of $\mathbb{R}^m$ which is r-dense in $B_r(x)$, then x lies in X.*

Proof Indeed, if H is a half-space that contains X, then H must also contain x, for otherwise there would be a point of $B_r(x)$ lying at a distance of greater than r from H, and therefore also from X, which is absurd. The result follows. □

For $x \in \mathbb{R}^m$ and $\lambda > 0$, let D_x^λ denote the affine transformation which dilates by a factor of λ about the point x, that is,

$$D_x^\lambda y = x + \lambda(y - x).$$

Lemma 7.2.3 *Let (K_n) be a sequence of compact, convex subsets of $\mathbb{R}^m$ converging in the Hausdorff sense to the compact, convex subset K_∞. If x is an interior point of K_∞ then, for all $\lambda > 1$, and for all sufficiently large n,*

$$D_x^{\frac{1}{\lambda}} K_\infty \subseteq K_n \subseteq D_x^\lambda K_\infty.$$

Proof Suppose that $B_{2r}(x)$ is contained in K_∞. We first show that $B_r(x)$ is also contained in K_n for sufficiently large n. Indeed, since (K_n) converges to K_∞ in the Hausdorff sense, for sufficiently large n, the set K_n is r-dense in $B_r(y)$ for all $y \in B_r(x)$. It follows by Lemma 7.2.2 that, for all such n, $B_r(x)$ is also contained in K_n, as asserted.

Observe now that, for any convex subset K, if $B_r(x) \subseteq K$, then for all $\lambda > 1$,

$$B_{(\lambda-1)r}(K) \subseteq D_x^\lambda K.$$

Thus, since (K_n) converges to K_∞ in the Hausdorff sense, for sufficiently large n,

$$K_n \subseteq B_{(\lambda-1)r}(K_\infty) \subseteq D_x^\lambda K_\infty,$$

and the second inclusion follows. Likewise, for sufficiently large n,

$$K_\infty \subseteq B_{(\lambda-1)r}(K_n) \subseteq D_x^\lambda K_n,$$

so that

$$D_x^{\frac{1}{\lambda}} K_\infty \subseteq K_n,$$

and the first inclusion follows. This completes the proof. □

7.3 Convex Subsets Viewed Intrinsically

First recall that, given two compact metric spaces (X_1, d_1) and (X_2, d_2), their *Gromov-Hausdorff distance* (see [4]) is defined by

$$d_{\mathrm{GH}}((X_1, d_1), (X_2, d_2)) := \operatorname*{Inf}_{\phi_1, \phi_2, (Y,\delta)} d_{\mathrm{H}}(\phi_1(X_1), \phi_2(Y_2)),$$

where the infimum is taken over all metric spaces (Y, δ) and functions $\phi_1 : X_1 \to Y$ and $\phi_2 : X_2 \to Y$ which are isometries onto their images. The following technical result will prove useful.

Lemma 7.3.1 *Let X_1 and X_2 be compact metric spaces with metrics d_1 and d_2 respectively. For $\epsilon \in]0, 1]$, suppose that there exist surjective maps $\Phi : X_1 \to X_2$ and $\Psi : X_2 \to X_1$ such that*

$$d_2(\Phi(x), \Phi(y)) \leq (1 + \epsilon) d_1(x, y),$$
$$d_1(\Psi(x), \Psi(y)) \leq (1 + \epsilon) d_2(x, y),$$

and

$$d_1(x,\ \Psi\Phi(x)) \leq \epsilon.$$

Then the Gromov-Hausdorff distance between X_1 and X_2 satisfies

$$d_{\mathrm{GH}}(X_1, X_2) \leq \epsilon(2 + \mathrm{Max}(Diam(X_1), Diam(X_2))).$$

Proof Indeed, consider first a compact metric space (X, d). Observe that the map $D : X \to L^\infty(X)$ given by $D(x)(y) := d(x, y)$ is an isometry onto its image. Furthermore, given another compact metric space (X', d') and a surjective map $\Phi : X \to X'$, the composition operator $\Phi^* : L^\infty(X') \to L^\infty(X)$ also defines an isometry onto its image and, in particular, restricts to an isometry from $D'(X')$ onto a subset of $L^\infty(X)$. For each i, define $D_i : X_i \to L^\infty(X_i)$ in this manner. From the above relations, we deduce that, for all $x, y \in X_1$,

$$\Phi^* D_2(\Phi(x))(y) - D_1(x)(y) \leq \epsilon \mathrm{Diam}(X_1),$$
$$(\Psi\Phi)^* D_1(\Psi\Phi(x))(y) - \Phi^* D_2(\Phi(x))(y) \leq \epsilon \mathrm{Diam}(X_2), \text{ and}$$
$$D_1(x)(y) - (\Psi\Phi)^* D_1(\Psi\Phi(x))(y) \leq 2\epsilon.$$

Together these relations yield

$$\|\Phi^* D_2(\Phi(x)) - D_1(x)\|_{L^\infty} \leq \epsilon(2 + \mathrm{Max}(\mathrm{Diam}(X_1), \mathrm{Diam}(X_2))),$$

from which the result follows. □

Now let Ω be an affine chart of M^m_κ in $\mathbb{R}^m$. That is, when $\kappa = -1$, and $M^m_\kappa = \mathbb{H}^m$ is hyperbolic space, Ω is the open unit ball in $\mathbb{R}^m$ which identifies with $\mathbb{H}^m$ via the Kleinian parametrisation; when $\kappa = 0$ and $M^m_\kappa = \mathbb{R}^m$ is Euclidean space, Ω is simply the whole of $\mathbb{R}^m$; and when $\kappa = 1$ and $M^m_\kappa = S^m$ is the unit sphere, Ω is also the whole of $\mathbb{R}^m$ which now identifies with an open hemisphere also via the Kleinian parametrisation. Let $\overline{g}$ denote the riemannian metric of this affine chart and let $\overline{d}$ denote the topological metric (distance function) that it defines. Throughout the rest of this section, for any subset X of Ω, and for all $r > 0$, $B_{r,\overline{d}}(X)$ will denote the open neighbourhood of radius r about X with respect to $\overline{d}$.

Consider now a compact, convex subset K of Ω, and let $\Pi : \Omega \to K$ be the closest point projection.

Lemma 7.3.2 *If $B_{r,\overline{d}}(K) \subseteq \Omega$, then for all $x, y \in B_{r,\overline{d}}(K)$,*

$$\overline{d}(\Pi(x), \Pi(y)) \leq \frac{1}{cos(r)}\overline{d}(x, y).$$

Remark In fact, when $\kappa \in \{-1, 0\}$, the closest point projection is a contraction (see [1]).

Proof It suffices to consider the case where x and y are elements of $B_{r,\overline{d}}(K)\setminus K$, as the remaining cases are similar and simpler. Consider the geodesic quadrilateral determined by the ordered sequence of points $(x, y, \Pi(y), \Pi(x))$. By convexity, the geodesic segment $\Pi(x)\Pi(y)$ is contained in K. In particular, since $\Pi(x)$ and $\Pi(y)$ are the closest points in K to x and y respectively, the angles $x\Pi(x)\Pi(y)$ and $y\Pi(y)\Pi(x)$, taken with respect to the metric $\overline{g}$, are both at least $\pi/2$, and the result now follows by standard comparison theory (see [2]). □

Lemma 7.3.3 *Let (K_n) be a sequence of compact, convex subsets of Ω with non-trivial interiors, let K_∞ be another compact, convex subset of Ω with non-trivial interior, and for all $n \in \mathbb{N} \cup \{\infty\}$, let d_n denote the intrinsic metric of ∂K_n with respect to $\overline{g}$. If (K_n) converges to K_∞ in the Hausdorff sense, then $(\partial K_n, d_n)$ converges to $(\partial K_\infty, d_\infty)$ in the Gromov-Hausdorff sense.*

Proof Let x be an interior point of K_∞. Choose $\lambda > 1$. By Lemma 7.2.3, for sufficiently large n,

$$D_x^{\frac{1}{\lambda}} K_\infty \subseteq K_n \subseteq D_x^\lambda K_\infty.$$

Now denote $\Phi := D_x^\lambda \Pi_\infty$ and $\Psi := \Pi_n D_x^\lambda$, where Π_∞ and Π_n are respectively the closest point projections onto $D_x^{1/\lambda} K_\infty$ and K_n with respect to the metric $\overline{g}$. Since Φ and Ψ are continuous with unit degree, they are surjective. Thus, by Lemma 7.3.2 and the smoothness of $\overline{g}$, for any given $\epsilon > 0$, Φ and Ψ satisfy the hypotheses of Lemma 7.3.1 provided that λ is chosen sufficiently close to 1. The result follows. □

Proof of Theorem 7.1.1 Let (X_n) be a sequence of finite subsets of Ω converging to X in the Hausdorff sense. For all n, $\mathrm{Conv}(X_n)$ is a convex polyhedron with vertices in X_n. In particular, for all n, the intrinsic metric of $\Sigma_n := \partial\mathrm{Conv}(X_n)\setminus X_n$ is locally isometric to M^2_κ. Since, by Lemma 7.2.1, the sequence $(\mathrm{Conv}(X_n))$ of convex hulls converges in the Hausdorff sense to the convex hull $\mathrm{Conv}(X_\infty)$, the first assertion now follows by Lemma 7.3.3. To prove the second assertion, consider a totally geodesic supporting plane P to $\mathrm{Conv}(X_\infty)$ at some point of $\Sigma := \mathrm{Conv}(X_\infty)\setminus X$. Since $\mathrm{Conv}(X_\infty)$ is a convex hull, the intersection of P with $\mathrm{Conv}(X)$ is either a geodesic segment with end-points in X, or a convex polygon with geodesic edges and vertices in X. From this the second assertion readily follows, and this completes the proof. □

Proof of Theorem 7.1.2 Fix a point $x \in \mathbb{H}^3$, and for all $r > 0$, let $\overline{B}_r(x)$ denote the closed ball of radius r about x in $\mathbb{H}^3$. For all r, $\mathrm{Conv}(X) \cap \overline{B}_r(x)$ is the convex hull of the compact set $\mathrm{Conv}(X) \cap \partial B_r(x)$, and the result follows by Theorem 7.1.1. □

References

1. W. Ballmann, M. Gromov, V. Schroeder, *Manifolds of Nonpositive Curvature* (Birkhäuser Verlag, Basel, 1985)
2. J. Cheeger, D.G. Ebin, *Comparison Theorems in Riemannian Geometry* (American Mathematical Society, Providence, 2008)
3. D.B.A. Epstein , A. Marden , Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces, in *Fundamentals of Hyperbolic Geometry: Selected Expositions* (Cambridge University Press, Cambridge, 2006), pp. 117–266
4. M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces* (Birkhäuser Verlag, Basel, 2007)
5. J.R. Munkres, *Topology* (Prentice-Hall, Upper Saddle River, 1975)
6. C. Rourke, Convex ruled surfaces, in *Analytical and Geometric Aspects of Hyperbolic Space* (Cambridge University Press, Cambridge, 1987), pp. 255–272
7. G.A.C. Smith, *Global Singularity Theory for the Gauss Curvature Equation*. Ensaios Mathematics, 2015
8. W.P. Thurston, *The Geometry and Topology of Three-Manifolds* (Mathematical Sciences Research Institute, Berkeley, 2002)

Chapter 8
The Double Limit Theorem and Its Legacy

Cyril Lecuire

Abstract This chapter surveys recent and less recent results on convergence of Kleinian representations, following Thurston's Double Limit and "AH(acylindrical) is compact" Theorems.

Keywords Kleinian groups · Deformation space · Ahlfors–Bers coordinates · Algebraic convergence

2020 Mathematics Subject Classification Primary 57M30; Secondary 30F60, 30F40

8.1 Introduction

Although Kleinian groups were discovered in the late nineteenth century (by Schottky, Klein and Poincaré), the story of the present chapter's topic really starts in the early sixties with the works of Ahlfors and Bers on quasi-conformal deformations of Fuchsian groups [2, 3, 8]. In particular, after further development by Maskit [49] and Kra [44], it led to the parametrization of the space of quasi-conformal deformations by the conformal structure at infinity. Combined with later works of Marden [48] and Sullivan [82] this provided coordinates for the interior of the deformation space $AH(\pi_1(M))$ usually called the Ahlfors–Bers coordinates. This also led to Bers' compactification of Teichmüller space, [7], who in particular introduced sequences of quasi-Fuchsian groups converging to non quasi-Fuchsian ones. Meanwhile Jørgensen developed methods to study sequences of Kleinian groups, showing that discreteness is a closed property and isolating two types of convergence, which he called algebraic and geometric. In contrast to this rich theory of deformations of quasi-Fuchsian groups, Schottky groups and other infinite

C. Lecuire (✉)
Institut de Mathématiques de Toulouse, Université Paul Sabatier, Toulouse, France
e-mail: lecuire@math.univ-toulouse.fr

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_8

covolume convex cocompact Kleinian groups, Mostow showed in the late sixties that cocompact Kleinian groups are rigid, [61]. Then in the late seventies, Thurston revolutionized the world of low-dimensional geometry, introducing original and exotic tools to prove beautiful and unexpected new results.

In an incomplete series of articles ([84] and sequel) Thurston planned to present the arguments involved in the proof of the Geometrization Theorem for Haken Manifolds. Convergence of Kleinian representations plays a central role in each of the three existing papers: the main result of [84] is that AH(acylindrical) is compact (Theorem 8.1.1 in the present chapter), the Double Limit Theorem (Theorem 8.1.3 below) is essential in [87] and [88] is devoted to the Broken Windows Theorem (Theorem 8.4.2) and related results.

Now that the historical context has been set up, let us get more technical. The deformation space $AH(\pi_1(M))$ of a hyperbolic 3-manifold M is the set of discrete and faithful representations $\rho : \pi_1(M) \to PSL_2(\mathbb{C})$ up to conjugacy, equipped with the quotient of the compact open topology. We will elaborate on the topology of $AH(\pi_1(M))$ in Sects. 8.2 and 8.7. For now let us consider two simple cases: when M is a product $I \times S$ over a closed surface and M is acylindrical. In both cases the conformal structures at infinity provides us with a homeomorphism $q : \text{int}(AH(\pi_1(M))) \to \mathcal{T}(\partial_{\chi<0}M)$ and we call $q(\rho)$ the *Ahlfors–Bers coordinates* of ρ. From this homeomorphism, we get that if a sequence $\{\rho_i\} \subset \text{int}(AH(\pi_1(M)))$ has bounded Ahlfors–Bers coordinates, then $\{\rho_i\}$ has a converging subsequence. The question we address in this chapter is: What is the behaviour of a sequence with diverging Ahlfors–Bers coordinates?

When M is acylindrical, a complete answer has been provided by Thurston ([84], and an alternate proof was given by Morgan–Shalen [60]) with the following result:

Theorem 8.1.1 (AH(acylindrical) **Is Compact)** *If M is any compact acylindrical 3-manifold with boundary, then $AH(\pi_1(M))$ is compact.*

Moreover, the fact that $AH(\pi_1(M))$ is compact characterizes acylindrical manifolds.

When $M = I \times S$, the Ahlfors–Bers coordinates are a pair of metrics $(\sigma^+, \sigma^-) \in \mathcal{T}(S) \times \mathcal{T}(S)$. A first condition for convergence comes from the works of Ahlfors and Bers (see Theorem 8.3.1):

Theorem 8.1.2 *Let $\{\rho_i\} \subset AH(\pi_1(S))$ be a sequence of representations with Ahlfors–Bers coordinates (σ_i^+, σ_i^-). If $\{\sigma_i^+\}$ converges in the Teichmüller space $\mathcal{T}(S)$ then $\{\rho_i\}$ has a converging subsequence.*

If we allow both coordinates σ_i^+ and σ_i^- to diverge, then $\{\rho_i\}$ may not have a converging subsequence. To undertake a finer analysis, we need a way to quantify the behavior of diverging sequences. Thurston used his compactification of Teichmüller space by projective measured laminations in the celebrated Double Limit Theorem:

Theorem 8.1.3 (Double Limit Theorem) *Let S be a closed surface and let μ^+, μ^- be two measured geodesic laminations that bind S. Then for any sequence*

$\{\sigma_i^+, \sigma_i^-\}$ *in* $\mathcal{T}(S) \times \mathcal{T}(S)$ *converging to (the projective classes of)* (μ^+, μ^-) *in* $\overline{\mathcal{T}}(S) \times \overline{\mathcal{T}}(S)$*, the sequence of quasi-Fuchsian representations with Ahlfors–Bers coordinates* (σ_i^+, σ_i^-) *has a converging subsequence.*

Otal gave an alternative proof of this result in [75]. Thurston's, Morgan–Shalen's and Otal's proofs of Theorems 8.1.1 and 8.1.3 have seen adaptations and improvements by different authors which led to various generalizations. In this chapter, we will survey those generalizations and outline the arguments that are involved in their proofs.

We conclude this introduction with a plan of the chapter. In the second section we introduce deformation spaces and Thurston's and Culler–Morgan–Shalen's compactifications. In Sect. 8.3, we explain Thurston's and Otal's proofs of the Double Limit Theorem. In the following section, we describe Thurston's and Morgan–Shalen's arguments leading to the proof of the compactness of AH(acylindrical) and its more general version, the Broken Window Only Theorem. Then we explain how to combine the Broken Window Only Theorem with the proof of the Double Limit Theorem to get a convergence Theorem for manifolds with incompressible boundary. In Sect. 8.5, we describe in details progress that led to a general statement for all compact hyperbolic 3-manifolds, answering a question of Thurston. In Sect. 8.6, we mention the obstacles encountered when trying to relax the conditions in the Double Limit Theorem until they are necessary and sufficient. Then we describe a change of setting, using the curve complex to define such necessary and sufficient conditions. Lastly in Sect. 8.7, we depict some of the applications of the theorems listed in this chapter, starting with Thurston's Hyperbolization Theorem.

8.2 Compactifications of Deformation Spaces

8.2.1 Definitions

8.2.1.1 Deformation Spaces

Let M be a compact n-manifold (we are only interested in the cases $n = 2$ and 3) and set $G = \pi_1(M)$. Let $\mathcal{D}(G) \subset Hom(G, \mathrm{Isom}^+(\mathbb{H}^d))$ denote the set of discrete and faithful representations. Given $\rho \in \mathcal{D}(G)$, the quotient $\mathbb{H}^d/\rho(G)$ is a complete hyperbolic n-manifold homotopy equivalent to M. We equip $Hom(G, \mathrm{Isom}^+(\mathbb{H}^d))$ (and hence $\mathcal{D}(G)$) with the compact open topology, so that $\rho_n \longrightarrow \rho$ if $\rho_n(g) \longrightarrow \rho(g)$ for any $g \in G$. This topology is also called the *algebraic topology*. Notice that when G is not Abelian, $\mathcal{D}(G)$ is a closed subset [30, 38]. The group $\mathrm{Isom}^+(\mathbb{H}^d)$ acts properly discontinuously by conjugacy on $\mathcal{D}(G)$ and the quotient $AH(G)$ is the *deformation space* of G. $AH(G)$ is also the space of marked hyperbolic structures (N, h) where N is a complete hyperbolic n-manifold and $h : M \to N$ is a homotopy equivalence, modulo the equivalence relation $(N, h) \sim (N', h')$ if there is an isometry $\psi : N \to N'$ such that h' is homotopic to $\psi \circ h$.

When $d = 2$ and $\partial M = \emptyset$, $AH(G) = \mathcal{T}(M) \cup \mathcal{T}(\overline{M})$ is the union of two copies of the Teichmüller space of M.

When $d = 3$, by Thurston's Hyperbolisation Theorem, $AH(G) \neq \emptyset$ if and only if M is irreducible and atoroidal. Let us focus on this case making our way towards the Ahlfors–Bers coordinates mentioned in the introduction. To simplify the notation and statements, we will use the same notation for a conjugacy class in $AH(G)$ and a representative of this conjugacy class and we will assume that M is orientable and that ∂M contains no tori.

8.2.1.2 Ahlfors–Bers Coordinates

Given $\rho \in AH(G)$, the group $\rho(G)$ acts by conformal transformations on $\hat{\mathbb{C}} = \partial_\infty \mathbb{H}^3$. Let Ω_ρ be the maximal invariant open subset on which this action is properly discontinuous. We say that ρ is *convex cocompact* if $(\mathbb{H}^3 \cup \Omega_\rho)/\rho(G)$ is compact (this is equivalent to more classical definitions, see [48]). By Marden [48] and Sullivan [82], ρ is in the interior of $AH(G)$ if and only if it is convex cocompact. To each component $\mathcal{C}$ of $\mathrm{int}(AH(\pi_1(M)))$ is associated a pair (N, h) where N is a compact 3-manifold and $h : M \to N$ is a homotopy equivalence (up to an equivalence relation, see [6], here we only need a representative). Then, for each $\rho \in \mathcal{C}$ there is a homeomorphism $f_\rho : N \to (Hp^3 \cup \Omega_\rho)/\rho(G)$ such that $(f_\rho \circ h)_* = \rho$. Since the only requirement on f_ρ is $(f_\rho \circ h)_* = \rho$, the isotopy class of f_ρ is uniquely defined up to the action of the group $\mathrm{Mod}_0(N)$ of isotopy classes of orientation-preserving homeomorphisms of N that are homotopic to the identity.

Associating to a each representation $\rho \in \mathcal{C}$ its conformal structure at infinity $\Omega_\rho/\rho(G)$, we get a map $q : \mathcal{C} \to \mathcal{T}(\partial N)/\mathrm{Mod}_0(N)$. As mentioned in the introduction, by results of Ahlfors–Bers [3], Bers [8], Maskit [49] and Kra [44], q is a homeomorphism. We call $q(\rho)$ the *Ahlfors–Bers coordinates* of ρ. When M has incompressible boundary, $Mod_0(M)$ and $Mod_0(N)$ are trivial and $\mathcal{C} \approx \mathcal{T}(\partial N)$ is an open ball.

Notice that when $M = S \times I$, M is acylindrical or M is a handlebody, $\mathrm{int}(AH(\pi_1(M)))$ has only one component (corresponding to (M, Id)). The interested reader may refer to [6] or [4] for an enumeration of the components of $AH(\pi_1(M))$ and $\mathrm{int}(AH(\pi_1(M)))$ in general.

To study sequences that do not converge in the interior of $AH(\pi_1(M))$, we want to describe how their Ahlfors–Bers coordinates diverge in Teichmüller space. This naturally leads us to introduce Thurston's compactification.

Before that, let us finish this section with a notation. In a compact connected n-manifold M, a closed curve γ defines through its free homotopy class a conjugacy class in the fundamental group that we will also denote by γ. Given $\rho \in AH(\pi_1(M))$, we denote by $\ell_\rho(\gamma)$ the length in $\mathbb{H}^n/\rho(\pi_1(M))$ of the geodesic γ^*_ρ in the free homotopy class corresponding to $\rho(\gamma)$.

8.2.2 Thurston's Compactification of Teichmüller Space

Thurston constructed a compactification of Teichmüller space by projective measured foliations or equivalently projective measured geodesic laminations. We will adopt the latter since it is better suited to applications and extensions to Kleinians group. Before proceeding, let us briefly mention that this compactification led to Thurston's celebrated classification of surface homeomorphisms ([87, Theorem 2.5], see also [34] or [35]).

A *geodesic lamination* L on a closed hyperbolic surface S is heuristically a Hausdorff limit of multi-curves, i.e. disjoint unions of simple closed geodesics. The actual definition, which follows, encompasses a slightly larger set but in practice we will only consider such limits. A *geodesic lamination* is a compact set that is a (non-empty) disjoint union of complete embedded geodesics. Note that this definition can be made independent of the choice of metric on S, see [75, Appendice] for example.

A *measured geodesic lamination* λ consists of a geodesic lamination $|\lambda|$ and a transverse measure on $|\lambda|$: any arc $k \cong [0, 1]$ embedded in S transverse to $|\lambda|$, such that $\partial k \subset S - |\lambda|$, is endowed with a transverse measure $d\lambda$ such that:

- the support of $d\lambda|_k$ is $|\lambda| \cap k$;
- if an arc k' can be homotoped to k by a homotopy preserving $|\lambda|$ then $\int_k d\lambda = \int_{k'} d\lambda$.

The simplest case of measured geodesic laminations is a weighted simple closed geodesic δc, i.e. a simple closed geodesic c equipped with a transverse Dirac measure with weight δ. Weighted multi-curves are dense in the space $\mathcal{ML}(S)$ of measured geodesic laminations equipped with the weak* topology. Thus measured geodesic laminations can simply be viewed as limits of weighted multi-curves.

Given a hyperbolic metric on S, and hence a faithful and discrete representation $\rho : \pi_1(S) \to PSL_2(\mathbb{R})$, the length of a weighted simple closed geodesic δc, is defined by homogeneity: $\ell_\rho(\delta c) = \delta \ell_\rho(c)$. Then the length of a weighted multi-curve is simply the sum of the length of its weighted leaves and the length of a measured geodesic lamination is defined by taking limits of lengths of weighted multi-curves. Alternatively, given a measured geodesic lamination μ, we may pick a family k of arcs transverse to its support $|\mu|$ so that the components of $|\mu| - k$ are arcs with bounded lengths. Then the length of μ is computed by integrating the lengths of these arcs over the transverse measure. It turns out that these two definitions are equivalent and it follows from this equivalence that the definition using limit of sequences of weighted multi-curves is independent of the choice of the sequence.

Given a simple closed geodesic c and $\lambda \in \mathcal{ML}(S)$, the intersection number $i(c, \lambda)$ is the total weight of the measure on c when c is transverse to λ and is 0 otherwise, i.e. when c lies in or is disjoint from the support of λ. This extends to weighted simple closed geodesics by homogeneity: $i(\delta c, \lambda) = \delta i(c, \lambda)$, to weighted multi-curves by additivity and then to measured geodesic laminations by continuity.

There is a natural action of $\mathbb{R}_+^*$ on $\mathcal{ML}(S)$ obtained by multiplying the measure and the space $\mathcal{PML}(S)$ of projective measured geodesic laminations is the quotient of $\mathcal{ML}(S) - \{0\}$ under this action.

Thurston uses the intersection number to define a compactification of Teichmüller space by projective measured geodesic laminations ([83, Theorem 2.2]):

Theorem 8.2.1 (Laminations Compactify Teichmüller Space) *The union $\overline{T(S)} = T(S) \cup \mathcal{PML}(S)$ has a natural topology homeomorphic to a closed ball.*

In this topology, a sequence $\{\rho_i\}$ of representations in $\mathcal{T}(S)$ converges to a lamination $[\mu] \in \mathcal{PML}(S)$ if and only if there is a sequence $\{\mu_i\} \longrightarrow \infty$ (i.e. there is an arc k with $\int_k d\mu_i \longrightarrow \infty$) of measured laminations converging projectively to μ such that for all $\mu' \in \mathcal{ML}(S)$ for which $i(\mu', \mu) \neq 0$,

$$\lim_{i \longrightarrow \infty} \frac{\ell_{\rho_i}(\mu')}{i(\mu_i, \mu')} = 1.$$

Furthermore, $\ell_{\rho_0}(\mu_i) \longrightarrow \infty$ but $\ell_{\rho_i}(\mu_i)$ remains bounded.

Moreover, there is a constant C such that

$$i(\mu', \mu_i) \leq \ell_{\rho_i}(\mu') \leq i(\mu', \mu_i) + C\ell_{\rho_0}(\mu'). \tag{8.1}$$

The first part of the statement defines the compactification of Teichmüller space by projective measured geodesic lamination. The general idea is that if a sequence eventually stays outside every compact set, the lengths of some closed geodesics go to infinity: the metric is stretched. Since the area is bounded, locally, the metric is stretched only in one direction, transversely to a measured geodesic lamination μ_i so that $\lim_{i \longrightarrow \infty} \frac{\ell_{\rho_i}(\mu')}{i(\mu_i, \mu')} = 1$. We may then extract a projectively converging subsequence from the sequence $\{\mu_i\}$.

Formula (8.1) gives a more precise and uniform approximation. This uniformity can be used to prove a convergence result for surfaces in the spirit of the Double Limit Theorem (compare with Theorem 8.3.2). We say that two measured geodesic laminations γ, λ *bind* S if $i(\gamma, \nu) + i(\lambda, \nu) > 0$ for any non-trivial $\nu \in \mathcal{ML}(S)$.

Theorem 8.2.2 *Let S be a closed surface and let $\mu^{\pm}$ be two measured geodesic laminations that bind S. Let $\{\mu_i^{\pm}\}$ be two sequences of weighted multi-curves converging $\mu^{\pm}$. Then any sequence $\{\rho_i\} \subset \mathcal{T}(S)$ such that $\{\ell_{\rho_i}(\mu_i^+)\}$ and $\{\ell_{\rho_i}(\mu_i^-)\}$ are bounded has a converging subsequence.*

Proof If $\{\rho_i\}$ does not have a converging subsequence then it has a subsequence converging to a projective measured geodesic lamination $[\nu]$. Theorem 8.2.1 provides a sequence $\{\nu_i\} \longrightarrow \infty$ converging projectively to ν, i.e. $\{\varepsilon_i \nu_i\}$ converges to ν for a sequence $\varepsilon_i \longrightarrow 0$, such that the inequalities (8.1) are satisfied. Since μ^+ and μ^- bind S, $i(\mu^+, \nu) + i(\mu^-, \nu) > 0$, say $i(\mu^+, \nu) > 0$. By continuity of the intersection number, $i(\mu_i^+, \nu_i) \longrightarrow \infty$. Now inequality (8.1) contradicts the assumption that $\ell_{\rho_i}(\mu^+)$ is bounded. □

8.2.3 Culler–Morgan–Shalen's Compactification

A different point of view on the compactification of deformation spaces, using methods from algebraic geometry, was introduced by Culler and Shalen in [32] and then further developed by Morgan and Shalen [57, 59] and [60]. In particular, in [57] and [55] (see also [58]), Morgan and Shalen use valuations to compactify deformation spaces for hyperbolic manifolds in any dimension. The added points are actions on Λ-trees from which one easily extracts an action on a real tree (more details about these below). By a result of Skora, [79], (small minimal) actions of surface groups on real trees are dual to measured geodesic laminations. Thus, in dimension 2, Thurston's and Culler–Morgan–Shalen's compactification of Teichmüller spaces are equivalent.

In [9, 77] and [29], Bestvina, Paulin and Chiswell give an alternative and more geometric approach (with some variations) to Culler–Morgan–Shalen's compactification by actions on real trees. Let us sketch the ideas behind that geometric approach.

Consider a sequence of faithful and discrete representation $\rho_i : G \to Isom(\mathbb{H}^d)$ of a non-Abelian finitely generated group G and set $K_i = \inf_{x\in\mathbb{H}^d}\{\max_{g\in S} d(x, gx)\}$ for a finite generating set $S \subset G$. Since $\rho_i(G)$ is discrete and non-Abelian, K_i is a minimum reached at some point x_i. Up to conjugating ρ_i, we may assume that $x_i = O$. The sequence $\{\rho_i\}$ stays in a compact subset of the deformation space if and only if K_i is bounded. When K_i goes to infinity, one rescales $\mathbb{H}^d$ by multiplying the distances by K_i^{-1} so that the action of $\rho_i(S)$ on $K_i^{-1}\mathbb{H}^d$ is bounded. In $\mathbb{H}^d$, geodesic triangles are δ-thin, in the sense that any edge lies in a δ-neighbourhood of the other two (with $\delta = \log 2$). When we rescale the metric, the triangles become $K_i^{-1}\delta$-thin with $K_i^{-1}\delta \longrightarrow 0$, so that they look more and more like tripods as i goes to ∞. One then just needs the appropriate formalism to find a subsequence such that the action of $\rho_i(G)$ on $K_i^{-1}\mathbb{H}^n$ tends in some way to an action on a geodesic metric space where every geodesic triangle is a tripod. Such a space is called a *real tree*, a generalisation of simplicial trees that allows more flexibility on the vertices (they can accumulate or form a continuum). The convergence "in some way" is made formal by using the pointed Gromov–Hausdorff topology and either sequences of expanding finite subsets of G (as in [9] and [77]) or ultra-filters ([29], see also [39, chapter 9]). Thus we have extracted a subsequence of ρ_i converging to an action of G on a real tree. Up to taking a substree, the action can be assumed to be *minimal*, i.e. there is no invariant subtree. Furthermore, one can deduce from Margulis' Lemma that the action is *small*, i.e. edge stabilizers are Abelian. Notice that if we choose a different generating set S, we may get a different sequence K_i and the limiting tree may differ by a homothety.

As mentioned above, in dimension 2, the compactification by actions on real trees is equivalent to Thurston's compactification by projective measured geodesic laminations. This identification goes through the *dual tree* $\mathcal{T}_\lambda$ to a measured geodesic laminations $\lambda \in \mathcal{ML}(S)$. To define $\mathcal{T}_\lambda$, we first replace the closed leaves by foliated neighbourhoods so that the transverse measure has no atoms. The preimage

$\tilde{\lambda} \subset \mathbb{H}^2$ of λ under the covering projection $\mathbb{H}^2 \to S$ defines a partition $\mathcal{P}$ of $\mathbb{H}^2$ into closed sets. An element of $\mathcal{P}$ is either the closure of a component of $\mathbb{H}^2 - |\lambda|$ or a leaf of $|\lambda|$ which is not in the closure of such a component. The transverse measure defines a distance on $\mathcal{P}$ turning it into a real tree $\mathcal{T}_\lambda$ and the action of $\pi_1(S)$ on $\mathbb{H}^2 = \tilde{S}$ induces an action on $\mathcal{T}_\lambda$. Notice that by the theorem of Skora [79], any small minimal action of $\pi_1(S)$ on a real tree is dual to a measured geodesic lamination.

If a sequence of representations $\rho_i : \pi_1(S) \to PSL(2,\mathbb{R})$ converges in Thurston's compactification to a (projective) measured geodesic lamination μ, then $\{\rho_i\}$ also converges in Culler–Morgan–Shalen's compactification to the action of $\pi_1(S)$ on $\mathcal{T}_\mu$. A simple way to see the unity of these two compactifications is to look at translation lengths and intersection numbers. Given an action of a group G on a real tree $\mathcal{T}$ and $g \in G$, define its translation distance by $\delta_{\mathcal{T}}(g) = \inf\{d(x, gx) | x \in \mathcal{T}\}$. By Culler and Morgan [31], a minimal action of G by isometries on a real tree is uniquely defined by the function $\delta : G \to \mathbb{R}^+$. If c is a simple closed curve on S and if we also denote by c the corresponding element of $\pi_1(S)$, then we have $\delta_{\mathcal{T}_\mu}(c) = i(\mu, c)$. Now in Thurston's compactification, we have $\varepsilon_i \longrightarrow 0$ such that $\varepsilon_i \ell_{\rho_i}(c) \longrightarrow i(c, \mu)$ while in Culler–Morgan–Shalen's, we have $\varepsilon_i \ell_{\rho_i}(c) \longrightarrow \delta_{\mathcal{T}}(c)$. Hence $\mathcal{T}$ is dual to μ.

8.3 The Double Limit Theorem

In this section, we will describe Thurston's and Otal's proofs of the Double Limit Theorem. Let us first recall its statement.

Theorem 1.3 *Let S be a closed surface and let μ^+, μ^- be two measured geodesic laminations that bind S. Then for any sequence $\{(\sigma_i^+, \sigma_i^-)\}$ in $\mathcal{T}(S) \times \mathcal{T}(S)$ converging to (μ^+, μ^-) in $\overline{\mathcal{T}}(S) \times \overline{\mathcal{T}}(S)$, the sequence of quasi-Fuchsian representations with Ahlfors–Bers coordinates (σ_i^+, σ_i^-) has a converging subsequence.*

The first step in both proofs consists in establishing a link between the lengths of curves with respect to the conformal structures at infinity and their lengths inside the quotient 3-manifold. Let $\rho : \pi_1(S) \to PSL_2(\mathbb{C})$ be a quasi-Fuchsian representation with Ahlfors–Bers coordinates (i.e. conformal structures at infinity) $(\sigma^+, \sigma^-) \in \mathcal{T}(S) \times \mathcal{T}(S)$. Given a closed curve $\gamma \subset S$, let $\ell_{\sigma^+}(\gamma)$, resp. $\ell_{\sigma^-}(\gamma)$, denote the length of the geodesic in the homotopy class of γ with respect to the metric σ^+, resp. σ^-. Let also $\ell_\rho(\gamma)$ denote the length of the geodesic of $\mathbb{H}^3/\rho(\pi_1(S))$ in the homotopy class defined by γ.

Lemma 8.3.1 *We have: $\ell_\rho(\gamma) \leq 2\inf\{\ell_{\sigma^+}(\gamma), \ell_{\sigma^-}(\gamma)\}$.*

This statement, which is a reformulation of [7, Theorem 3], follows also from the work of Ahlfors [2] (see [75, Lemma 5.1.1]).

If a sequence $\sigma_i^{\pm}$ converges to a lamination $\mu^{\pm}$, then by Theorem 8.2.1, there is a sequence of measured laminations $\mu_i^{\pm}$ converging projectively to $\mu^{\pm}$ such that $\ell_{\sigma_i^{\pm}}(\mu_i)$ remains bounded. Since weighted multi-curves are dense in $\mathcal{ML}(S)$, we can assume that $\mu_i^{\pm}$ is a multi-curve for any i. Combining this observation with Lemma 8.3.1, Theorem 8.1.3 reduces to the following generalization of Theorem 8.2.2:

Theorem 8.3.2 *Let S be a closed surface and let μ^+, μ^- be two measured geodesic laminations that bind S. Let $\{\mu_i^+\}, \{\mu_i^-\} \subset \mathcal{ML}(S) \times \mathcal{ML}(S)$ be two sequences of weighted multi-curves converging respectively to μ^+ and μ^-. Then any sequence $\{\rho_i\} \subset AH(\pi_1(S))$ such that $\{\ell_{\rho_i}(\mu_i^+)\}$ and $\{\ell_{\rho_i}(\mu_i^-)\}$ are bounded has a converging subsequence.*

8.3.1 Thurston's Arguments: Efficiency of Pleated Surfaces

Thurston's approach to prove the Double limit Theorem is to project the 3-manifold to an immersed surface especially constructed so that the induced distortion on the metric is controlled and thus reduce the problem to the 2-dimensional case. This is done through the "Efficiency of Pleated Surfaces" Theorem which allows one to estimate the lengths of geodesics in the 3-manifold based on the length of their representatives on some surfaces specifically immersed in it. These surfaces are *pleated surfaces*, namely the immersions have totally geodesic image except on a geodesic lamination called the *pleating locus*. Such a pleated surface is locally ruled and the induced metric is hyperbolic. For example, let us pick a finite maximal lamination $\lambda \in \mathcal{ML}(S)$, and a representation $\rho \in \mathrm{int}(AH(\pi_1(S)))$. A surface $f_\lambda : S \to N_\rho = \mathbb{H}^3/\rho(\pi_1(S))$ pleated along λ always exists, it maps the leaves of λ to geodesics and the complementary regions to geodesic triangles.

Theorem 8.3.3 (Efficiency of Pleated Surfaces, [87, Theorem 3.3]) *Let S be a closed surface. For any $\varepsilon > 0$, there is a constant $C < \infty$ such that the following holds:*

- *Let λ be any finite maximal lamination on S.*
- *Let ρ be any element of* $\mathrm{int}(AH(\pi_1(S)))$ *such that no closed leaf of λ has length less than ε in $N_\rho = \mathbb{H}^3/\rho(\pi_1(S))$, and let $f_\lambda : S \to N_\rho$ be a surface which is pleated along λ.*
- *Let $\mu \in ML(S)$ be a measured geodesic lamination.*

Then

$$\ell_\rho(\mu) \le \ell_{f_\lambda}(\mu) \le \ell_\rho(\mu) + Ca(\lambda, \mu).$$

We will describe the *alternation number* $a(\lambda, .)$ in the sketch of the Proof of Theorem 8.3.3. For the proof of the Double Limit Theorem we only need to know that $a(\lambda, .)$ is finite and continuous [84, Proposition 3.2].

Before describing the Proof of Theorem 8.3.3, let us explain how it is used to conclude the Proof of Theorem 8.3.2 (and hence of the Double Limit Theorem). First, Thurston produces a curve c (actually infinitely many such curves, see [87, Corollary 4.3]) which is not too short in any of the manifolds $N_i = \mathbb{H}^3/\rho_i(\pi_1(M))$ (up to extracting a subsequence), i.e. $\ell_{\rho_i}(c) \geq \varepsilon$ for all i and a constant ε that depends only on S. Adding spiraling leaves, c can easily be extended to a maximal lamination λ with no closed leaf except for c. Then there is a unique pleated surface $f_{\lambda,i} : S \to N_i$ which maps each component of $S - \lambda$ to a geodesic triangle. Applying Theorem 8.3.3, we get that both $\{\ell_{f_{\lambda,i}}(\mu_n^+)\}$ and $\{\ell_{f_{\lambda,i}}(\mu_n^-)\}$ are bounded (see also [87, Theorem 4.4]). By Theorem 8.2.2, the metric induced by $f_{\lambda,i}$ stays in a compact set. It follows that for any closed curve d on S, $\ell_{\rho_i}(d) \leq \ell_{f_{\lambda,i}}(d)$ is bounded and that the sequence $\{\rho_i\}$ has a converging subsequence.

Sketch of the Proof of Theorem 8.3.3 These inequalities need only be proved for simple closed curves. Then they holds for weighted multicurves and extend to measured lamination by continuity of the length function [14] and of the alternation number [84, Proposition 3.2]. The left hand inequality is obvious so we focus on the right hand one.

Let $d \subset S$ be a closed geodesic for the metric induced by f_λ. Approximate d on S by a piecewise geodesic curve p made up of segments in λ and small jumps between those segments. To ensure that the jumps are small, we pick successive segments in asymptotic leaves of λ, and to have a control on the number of segments we pick non-successive segments in non-asymptotic leaves. The number of segments is then the number $a(\lambda, d)$ of times the direction of asymptoticity of the leaves of λ changes as one goes around d.

Next, consider a simplicial annulus A joining $f_\lambda(p)$ to the geodesic $d^* \subset N_\rho$ in the homotopy class of $f_\lambda(d)$ and fix $\delta > 0$. From each point $x \in f_\lambda(p)$ draw in A an arc A_x orthogonal to $f_\lambda(p)$ which either has length δ or hit ∂A before reaching that length. By the Gauss–Bonnet formula, the contribution to the length of $f_\lambda(p)$ of the points x for which A_x has length δ, A_x hits d^* or x is close to a corner is at most $\ell_N(d) + O(a(\lambda, d))$. For the remaining points, A_x is a shortcut in N, and the Uniform Injectivity Theorem (Theorem 8.3.4 below) says that there is a shortcut in S joining the preimage of the endpoints of A_x. It is not difficult to ensure in the construction of p that there are not too many such shortcuts. Thus we get C depending only on S and ε such that $\ell_{f_\lambda}(d) \leq \ell_{f_\lambda}(p) \leq \ell_N(d) + Ca(\lambda, d)$.

To have a complete overview of the proof, it remains to examine the Uniform Injectivity Theorem. Given a differentiable manifold N, let $\mathbb{P}N$ denote the tangent line bundle.

Theorem 8.3.4 (Uniform Injectivity, [84, Theorem 5.7]) *Let $\varepsilon_0 > 0$ and let S be a closed surface. Given a representation $\rho \in AH(\pi_1(S))$, a pleated map $f : S \to N_\rho = \mathbb{H}^3/\rho(\pi_1(S))$ which induces ρ and a lamination $\lambda \subset S$ which is mapped geodesically by f, denote by $g : \lambda \to \mathbb{P}M_\rho$ the canonical lifting. There is*

$\delta_0 > 0$ *depending only on* ε_0 *and* S *such that for any two points* x *and* $y \in \lambda$ *whose injectivity radii are greater than* ε_0*, if* $d_f(x, y) \geq \varepsilon_0$ *then* $d_{N_\rho}(f(x), f(y)) \geq \delta_0$.

The uniformity comes from a limit argument. Thurston first shows that g is injective [84, Theorems 5.5 and 5.6] by contradiction. A non injective map g would map two leaves of λ to the same geodesic and hence their closures to the same set. This would produce two non-homotopic simple closed curves $c_1, c_2 \subset S$ with the same image under f. This would contradict the assumption that f induces ρ. From the injectivity he then goes to the uniform injectivity by establishing the compactness of pleated surfaces (in the appropriate topology). □

8.3.2 Otal's Proof: Real Trees and δ-Realization of Train Tracks

In his book on Thurston's Hyperbolization Theorem for manifolds which fiber over the circle, Otal introduces a different strategy to prove the Double Limit Theorem. It goes by contradiction, using the Culler–Morgan–Shalen compactification by actions on real trees (the geometric approach as described in Sect. 8.2.3). The idea is to approximate geodesic laminations in $\mathbb{H}^3/\rho_i(\pi_1(S))$ by piecewise geodesic arcs with the geodesic pieces belonging to a finite set of homotopy classes which do not depend on i. The convergence to an action on a real tree allows us to estimate the behavior of the length of those geodesic arcs and then the behavior of geodesic laminations. These alternative arguments require an additional hypothesis:

Theorem 8.3.5 *Let* S *be a closed surface and let* μ^+, μ^- *be two minimal measured geodesic laminations that bind* S*. Let* $\{\mu_i^+, \mu_i^-\}$ *be two sequences of weighted multi-curves converging in the Hausdorff topology to almost minimal laminations containing* μ^+ *and* μ^- *respectively. Then any sequence* $\{\rho_i\} \subset AH(\pi_1(S))$ *such that* $\{\ell_{\rho_i}(\mu_i^+)\}$ *and* $\{\ell_{\rho_i}(\mu_i^-)\}$ *are bounded has a converging subsequence.*

A geodesic lamination is *minimal* if any leaf is dense and *almost minimal* if it is made up of one minimal lamination μ and finitely many leaves accumulating on μ. Notice that if μ^+ and μ^- have simply connected complementary regions (for example when they are stable laminations of pseudo-Anosov mapping classes), we could equivalently request that $\{\mu_i^+\}$ and $\{\mu_i^-\}$ converge projectively to projective laminations supported by $|\mu^+|$ and $|\mu^-|$. In particular Theorem 8.3.5 is sufficient for the proof of the Hyperbolization Theorem.

As mentioned earlier, the proof goes by contradiction. We consider a sequence $\{\rho_i\} \subset AH(\pi_1(S))$ of quasi-Fuchsian representations tending to a small minimal action of $\pi_1(S)$ on a (projective) real tree $\mathcal{T}$. Namely, there is $\varepsilon_i \longrightarrow 0$ such that the action of $\rho_i(\pi_1(S))$ on $\varepsilon_i\mathbb{H}^3$ tends to the action of $\pi_1(S)$ on $\mathcal{T}$. By Skora's Theorem [79], this action is dual to a (projective) measured lamination ν. Since μ^+ and μ^- bind S, at least one crosses ν, say $i(\mu^+, \nu) > 0$, and denote by μ_h^+ the Hausdorff limit of $\{\mu_i^+\}$. The next step in the proof consists in constructing a train track carrying μ_h^+ (and hence μ_i^+ for i large enough), using a segment of ν that crosses μ^+ as its unique switch. Before that, let us take a short break to review some definitions.

A *(fattened) train track* on a compact surface S is a finite family of rectangles which intersect only at their vertical sides. A connected component of the union of the vertical sides is called a *switch* and such switches are required to be embedded arcs. This is a fattened version of the train tracks defined by Thurston in [86] (see also Penner and Harer [78]. The rectangles come with a vertical and a horizontal foliations. To carry the metaphor further, let us call *rail* a line made up of horizontal fibers and *tie* a leaf of the vertical foliation. A geodesic lamination is *carried* by a train track if (up to isotopy) it lies in the train track and is transverse to the ties.

Picking a segment $\kappa \subset |\nu|$ that crosses μ^+ and grouping the component of $|\mu^+|-\kappa$ by homotopy classes, Otal constructs a train track $\mathcal{R}$ carrying μ_h^+ with κ as its only switch. The fact that $\mathcal{T}$ is dual to ν naturally produces a $\pi_1(S)$-equivariant map $f_\nu : \mathbb{H}^2 \to \mathcal{T}$. By construction, this map f_ν is monotonous on the preimage of the rails of $\mathcal{R}$ and not constant on any rectangle. Otal uses this observation to turn f_ν into a *realization* of $\mathcal{R}$, i.e. a map f that is injective when restricted to a lift of a rail. Then f is also a realization of any geodesic lamination λ carried by $\mathcal{R}$, i.e. it is injective when restricted to a leaf of the preimage of λ.

Let $\tilde{\mathcal{R}} \subset \mathbb{H}^2$ be the preimage of $\mathcal{R}$ and let $\tilde{\kappa} \subset \tilde{\mathcal{R}}$ be a lift of κ (the switch of $\mathcal{R}$). Recall that the action $\rho_i(\pi_1(S))$ on $\varepsilon_i\mathbb{H}^3$ tends to the action of $\pi_1(S)$ on $\mathcal{T}$ and consider a sequence of points $p_i \in \mathbb{H}^3$ tending to $p = f(\tilde{\kappa})$. Consider the ρ_i-equivariant map $F_i : \tilde{\mathcal{R}} \to \mathbb{H}^3$ that maps $\tilde{\kappa}$ to p_i and each rectangle to a geodesic segment. For any rectangle $\tilde{R}$, $\varepsilon_i\ell(F_i(\tilde{R}))$ converges to the positive length of $f(\tilde{R})$. It follows that for any geodesic l carried by R, $F(\tilde{l})$ is made up of long geodesic segments. But we cannot guarantee that $F(\tilde{l})$ is a quasi-geodesic since we have no control on the angle between two successive geodesic segments.

In the last step of the proof, Otal changes the train track $\mathcal{R}$ by a subdivision operation, producing a new train track $\mathcal{R}'$ carrying μ_h^+ and a ρ_i-equivariant map $F_i' : \tilde{\mathcal{R}}' \to \mathbb{H}^3$ which maps rectangles to long segments such that the angles between two successive segments are close to π. Then for i large enough and for any closed curve c carried by R, the projection of $F'_i(\tilde{c})$ to $\mathbb{H}^3/\rho_i(\pi_1(S))$ is a quasi-geodesic and its length is close to the length of the geodesic $c_i^* \subset \mathbb{H}^3/\rho_i(\pi_1(S))$ in the same homotopy class. Thus the length of c_i^* is approximated by the sum of the lengths of the images of the rectangle of $\mathcal{R}$ it goes through and we get:

$$\varepsilon_i\ell_{\rho_i}(c_i^*) \geq K\ell_{s_0}(c).$$

where ℓ_{s_0} is the length for a fixed reference hyperbolic metric on S, a simple way to roughly evaluate the number of rectangles through which c goes, K is a constant that depends only on $\mathcal{R}$ and the inequality holds for i large enough and for any closed curve carried by $\mathcal{R}$.

In particular, we have $\ell_{\rho_i}(\mu_i^+) \longrightarrow \infty$ which is the desired contradiction.

Remark 8.3.6 The assumption that $\mu_h^{\pm}$ is almost minimal was used in two instances:

- to deduce that μ_h^+ is carried by a train track $\mathcal{R}$ realized in $\mathcal{T}$ from the assumption that λ intersects μ and
- to construct a train track $\mathcal{R}$ with only one switch carrying μ_h^+.

The fact that $\mathcal{R}$ has only one switch simplifies the construction but removing that constraint would only add more technicalities, whereas $\mathcal{R}$ being realized (or equivalently μ_h^+ being realized) is required to end up with a piecewise geodesic curve made up of long segments with incident angles close to π.

Thus we could relax the assumption on $\mu^{\pm}$ being almost minimal as long as we can guarantee that μ_h^+ or μ_h^- is realized in any dual tree.

We could also put aside Skora's Theorem and dual laminations and start from the assumption that μ^+ is realized in $\mathcal{T}$. Proceeding with the same arguments from that point on leads to:

Theorem 8.3.7 (Continuity Theorem) *Let M be a compact atoroidal 3-manifold and $\{\rho_i\} \subset AH(\pi_1(M))$ be a sequence tending to a small minimal action of $\pi_1(M)$ on a real-tree $\mathcal{T}$. Let $\varepsilon_i \longrightarrow 0$ be such that $\forall g \in \pi_1(M)$, $\varepsilon_i \delta_{\rho_i}(g) \longrightarrow \delta_{\mathcal{T}}(g)$ and let $\mu \subset \partial M$ be a geodesic lamination which is realized in $\mathcal{T}$. Then there exists a neighborhood $\mathcal{V}(\mu)$ of $|\mu|$, and constants K, i_0 such that for any simple closed curve $c \subset \mathcal{V}(\mu)$ and for any $i \geq i_0$,*

$$\varepsilon_i l_{\rho_i}(c^*) \geq K l_{s_0}(c).$$

8.4 Manifolds with Incompressible Boundary

Next, we will consider Kleinian representations of fundamental groups of 3-manifolds with incompressible boundary, starting with acylindrical manifolds. Let us recall that an *essential* disc, annulus or torus is an incompressible properly embedded disc, annulus or torus that is not boundary parallel, i.e. cannot be homotoped relative to its boundary in ∂M. A compact 3-manifold is *atoroidal* if it does not contain any essential torus and is *acylindrical* if it does not contain any essential disc, torus or annulus.

Before discussing the compactness of $AH(acylindrical)$, let us outline the importance of acylindrical manifolds in the topology of 3-manifolds. For this purpose, we introduce the theory of the characteristic submanifold (or JSJ decomposition). To give a general idea let us say that the *characteristic submanifold* Σ of a compact 3-manifold with incompressible boundary is the smallest submanifold that contains all the essential tori, Klein bottles, annuli and Möbius bands up to isotopy (a precise definition can be found in [36] and [37], see also [12, Theorem 3.8]). Its existence and uniqueness (up to isotopy) has been established independently by

Jaco–Shalen [36] and Johannson [37]. We are only interested in orientable atoroidal 3-manifolds, in which case the components of Σ are essential I-bundles, solid tori and thickened tori. The solid tori and thickened tori are only required to intersect ∂M along a collection of disjoint annuli and tori, which is why they are not viewed as essential I-bundles. The components of $M - \Sigma$ are acylindrical relative to ∂M, i.e. if W is the closure of a component of $M - \Sigma$ and $\partial_0 W = W \cap \partial M$ then any annulus $(A, \partial A) \subset (W, \partial_0 W)$ can be homotoped in ∂W relative to its boundary. A relative version of this theory produces a characteristic submanifold relative to an incompressible subsurface $\partial_0 M \subset \partial M$ of the boundary (see [59, §IV.4.]), it contains all the essential annuli $(A, \partial A) \subset (M, \partial_0 M)$. This will be especially interesting in the next section where we will consider more general 3-manifolds since ∂M is allowed to be compressible as long as $\partial_0 M$ is incompressible.

Let us draw a simple conclusion from this dense paragraph: a compact orientable atoroidal 3-manifold with incompressible boundary is made up of I-bundles, (relative) acylindrical submanifolds and submanifolds with abelian fundamental groups. Since we have already studied deformations of hyperbolic I-bundles in the previous sections, it now seems natural to follow up with acylindrical manifolds.

Theorem 1.1 (AH(acylindrical) **Is Compact)** *If M is any compact acylindrical 3-manifold with boundary, then $AH(\pi_1(M))$ is compact.*

This result is due to Thurston, [84, Theorem 1.2], and then was proved by Morgan–Shalen, [60, Theorem V.2.1] with very distinct ideas and tools. Their overall strategies also differ: Thurston first proves Theorem 8.1.1 in [84] and later introduces new arguments (in [88]) to extend of the proof to a more general setting whereas Morgan and Shalen directly prove a general statement in [60] and deduce Theorem 8.1.1 as a special case. Both strategies still lead to comparable generalizations, which essentially state the following: for a compact atoroidal 3-manifold M with incompressible boundary, a sequence in $AH(\pi_1(M))$ can only degenerate on the fundamental group of the characteristic submanifold.

8.4.1 Thurston's Proof and Generalizations: Degenerating Simplices and Broken Windows

Following the chronological order, let us first outline Thurston's Proof of Theorem 8.1.1. Consider a sequence of maps $f_i : M \to M_i = \mathbb{H}^3/\rho_i(\pi_1(M))$ mapping a fixed triangulation of M minus the vertices to ideal simplices so that the restriction to the boundary is a pleated surface. We separate the simplices of the triangulation of M into two families Δ_b and Δ_∞ depending on whether the geometry of f_i remains bounded or goes to infinity. Thurston deduces then from the Uniform Injectivity Theorem that a neighbourhood of the interface between these two families has boundary with small area and hence with Abelian fundamental group. It follows then from topological considerations that Δ_b carries the fundamental group. Thus the sequence $\{\rho_i\}$ is bounded. In a subsequent paper, [88], Thurston uses the same argument to establish a relative compactness Theorem:

Theorem 8.4.1 (Relative Boundedness, [88, Theorem 3.1]) *Let M be a 3-manifold, and γ a doubly incompressible multicurve on ∂M. Then for any constant $A > 0$, the subset of $AH(\pi_1(M))$ such that the total length of γ does not exceed A is compact.*

We say that a multi-curve on the boundary of a compact 3-manifold is *doubly incompressible* if it intersects the boundary of any essential disc or annulus (this is a special case of Thurston's original definition [88, p. 10] where $S = \partial M$ and f is the inclusion).

Since ∂M is not assumed to be incompressible, the Uniform Injectivity Theorem may not apply under the assumptions of Theorem 8.4.1 (we will give more insight on this issue in the next section). To overcome this difficulty Thurston extends the Uniform Injectivity Theorem at the price of loosing some of its uniformity: the constant δ depends also on a doubly incompressible multicurve γ that must be contained in the pleating locus and on a bound on the length of this multicurve. Once this is established, the Proof of Theorem 8.4.1 follows the same outline as the Proof of Theorem 8.1.1.

Thurston uses Theorem 8.4.1 for a final generalization of Theorem 8.1.1: the Broken Windows Only Theorem. He uses a slight variation on the characteristic submanifold made up only of I-bundles which he calls the *windows*: he does not take the solid tori and thickened tori and replace them with a collection of thickened annuli. In his usual picturesque style, Thurston derives the name from the idea that if the manifold was made of glass, the window would be the part through which one could see without distortion. He shows that for a sequence in $AH(\pi_1(M))$, degenerations can only happen on the fundamental group of the window, hence carrying the metaphor further: " only the window breaks".

Theorem 8.4.2 (Broken Windows Only, [88, Theorem 0.1]) *If $\Gamma \subset \pi_1(M)$ is any subgroup which is conjugate to the fundamental group of a component of $M - \text{window}(M)$, then the set of representations of Γ in $\text{Isom}(\mathbb{H}^3)$ induced from $AH(\pi_1(M))$ are bounded, up to conjugacy.*

The window is an I-bundle over a (usually disconnected) compact surface S called the *window base* (denoted wb above). Its boundary ∂S is the *window frame*.

The Broken Windows Only Theorem is deduced from Theorem 8.4.1 and a uniform bound on the length of the window frame:

Theorem 8.4.3 (Window Frame Bounded, [88, Theorem 1.3]) *For any manifold M with incompressible boundary, there is a constant C such that among all elements $\rho \in AH(\pi_1(M))$, the length in N_ρ of $\partial \text{wb}(M)$ is bounded.*

Thurston's Proof of Theorem 8.4.3 ([84], see also the appendix of [55]) uses the area growth rate of branched pleated surfaces. An alternate proof using the Uniform Injectivity Theorem appeared in [20, Appendix].

In [87], the Broken Windows Only Theorem (Theorem 8.4.2) has a second part, generalizing a previous result of Thurston on surface groups ([87, Theorem 6.2]), and setting up the existence of sequences of maximal convergence and submanifolds of maximal convergence. But, as observed by Ohshika, the Convergence on Subsurfaces Theorem, [87, Theorem 6.2], does not extend to manifolds with incompressible boundaries as described by Thurston (see the example in [71, §5.3]). On the other hand, Canary et al. [28, Theorem 5.5] observed that one may remark the representations and extract a subsequence so that it converges on most of M:

Theorem 8.4.4 ([21, Theorem 2.8]) *Let M be a compact 3-manifold with incompressible boundary and consider a sequence $\{\rho_i\} \subset AH(\pi_1(M))$ of representation uniformizing M. Then after passing to a subsequence, there is a collection $\mathcal{B}$ of essential annuli and a sequence of homeomorphisms $\phi_i : M \to M$ each supported on $window(M)$ such that*

(1) $\lim \ell_{\rho_n \circ (\phi_n)_*}(c) = 0$ *for any simple closed $c \subset \partial\mathcal{B}$ and*
(2) $\{\rho_n \circ (\phi_n)_*\}$ *converges on the fundamental group of each component of $M - \mathcal{B}$.*

The proof of the last statement combines the Broken Windows Only Theorem, the Efficiency of Pleated Surfaces and Mumford Compactness Theorem ([62], see also [28, Proposition 5.6]).

8.4.2 *Morgan and Shalen's Arguments: Trees and Codimension-1 Laminations*

Morgan and Shalen start in a very general setting by considering a compact irreducible 3-manifold M and an incompressible subsurface of its boundary $\partial_0 M \subset \partial M$. They associate to each small minimal action of $\pi_1(M)$ on a real tree $\mathcal{T}$ a measured codimension 1 lamination $\mathcal{L} \subset M$ and a morphism between its dual tree $\mathcal{T}_{\mathcal{L}}$ and $\mathcal{T}$. This morphism may not be injective: it may fold, i.e. map two adjacent segments onto one. This possible lack of injectivity cannot be overcome in general since there are small minimal actions of fundamental groups of compact atoroidal 3-manifolds on real trees which are not dual to any codimension 1 laminations (see [71]). A morphism is still enough to guarantee that the fundamental group of every component of $M - \mathcal{L}$ has a fixed point in $\mathcal{T}$. In a previous work [59], Morgan and Shalen have shown that such a lamination sits (up to some surgeries and isotopies) in the characteristic submanifold relative to $\partial M - \partial_0 M$. This leads to the following statement:

Theorem 8.4.5 ([60, Theorem IV.1.2]) *Let M be a compact irreducible 3-manifold, let $\partial_0 M \subset M$ be an incompressible subsurface and let $\Sigma \subset M$ be the characteristic submanifold relative to $\partial_0 M$. Let $\pi_1(M) \times \mathcal{T} \to \mathcal{T}$ be a small action on a real tree and suppose that for any component Z of $\partial M - \partial_0 M$, $\pi_1(Z)$ has a fixed point. Then for each component C of $M - \Sigma$, the group $\pi_1(C)$ has a fixed point in T.*

When M is acylindrical, the characteristic submanifold is empty and it follows from this statement that there is no small minimal action of $\pi_1(M)$ on a non-trivial real tree. Then the conclusion of Theorem 8.1.1 follows from Culler–Morgan–Shalen's compactification of the deformation space.

Let us add that a more general result about splitting of groups acting on real trees (from which Theorem 8.4.5 can be deduced) has been obtained by Rips, using combinatorial methods instead of topological arguments, see [10] and [39, §12].

8.4.3 Mixing the Arguments

Theorems 8.4.5 and 8.4.2 both tell us that to bound a sequence of representations $\rho_i \in AH(\pi_1(M))$ it suffices to bound its restriction to the fundamental group of the window. Using this observation, we will extend the Double Limit Theorem to manifolds with incompressible boundaries. Let us first set up a property of laminations on ∂M that will play the role of the binding property in the Double Limit Theorem. We say that a measured lamination $\lambda \in \mathcal{ML}(\partial M)$ on the boundary of a manifold with incompressible boundary is *acylindrical* if there is $\varepsilon > 0$ such that $i(\lambda, \partial A) \geq \varepsilon$ for any essential annulus $A \subset M$. As observed by Bonahon–Otal [13], when M is not an I-bundle, it is equivalent to require that $i(\lambda, \partial A) > 0$ for any essential annulus $A \subset M$.

Theorem 8.4.6 *Let M be a compact hyperbolizable 3-manifold with incompressible boundary, let $\mathcal{C}$ be a connected component of* $\text{int}(AH(\pi_1(M)))$ *containing a representation uniformizing M and let $\mu \in \mathcal{ML}(\partial M)$ be an acylindrical measured geodesic lamination. Then for any sequence $\{\sigma_i\}$ in $\mathcal{T}(\partial M)$ converging to μ in $\overline{\mathcal{T}}(\partial M)$, the sequence of convex cocompact representations in $\mathcal{C}$ with Ahlfors–Bers coordinates σ_i has a converging subsequence.*

Using Lemma 8.3.1, the hypothesis on $\{\sigma_i\}$ can be replaced with a bound on the length of a sequence of weighted multi-curves converging to μ. The resulting statement can then be established using Theorem 8.4.2 and the arguments explained in Sect. 8.3.1. If we add the assumption that the limit is almost minimal (see Remark 8.3.6), we can also build a proof on Theorem 8.4.5 and Otal's arguments (compare with [13, Lemme 14]). Let us mix the two approaches to provide an alternative and fairly short proof (see also [66, Theorem 3.7] and [69, Theorem 3.1] for different mixes of those arguments).

Proof As in the proof of the Double Limit Theorem, we use Theorem 8.2.1 and Lemma 8.3.1 to obtain a sequence of weighted multi-curves $\mu_i \in \mathcal{ML}(\partial M)$ such that $\mu_i \longrightarrow \mu$ and $\{\ell_{\rho_i}(\mu_i)\}$ is bounded (compare with the beginning of Sect. 8.3).

As we have seen in Sect. 8.2.3, if ρ_i has no converging subsequence then a subsequence converges to a small minimal action on a real tree $\mathcal{T}$, namely there is $\varepsilon_i \longrightarrow 0$ such that $\varepsilon_i \ell_{\rho_i}(c^*) \longrightarrow \delta_{\mathcal{T}}(c)$ for any closed curve $c \in M$. For each component S of ∂M with negative Euler characteristic, since M has incompressible boundary, the map $i_* : \pi_1(S) \to \pi_1(M)$ induced by the inclusion provides us with a small action of $\pi_1(S)$ on $\mathcal{T}$. We can apply Skora's Theorem [79] to the minimal invariant subtree $\mathcal{T}_S \subset \mathcal{T}$ to get a dual lamination ν_S.

Building a pleated surface $f_{\lambda,i} : S \to N_i = \mathbb{H}^3/\rho_i(\pi_1(M))$ with a pleating locus that never gets too short (as explained in the Proof of Theorem 8.3.2), we get from the Efficiency of Pleated Surfaces (Theorem 8.3.3) $\ell_{\rho_i}(d) \leq \ell_{f_{\lambda,i}}(d) \leq \ell_{\rho_i}(d) + Ca(\lambda_i, d)$ for any simple closed curve $d \subset S$. In particular $\varepsilon_i \ell_{f_{\lambda,i}}(d) \longrightarrow \delta_T(d) = \delta_{T_S}(d)$. It follows that the metric induced by $f_{\lambda,i}$ converges to ν_S in Thurston's compactification. In particular, there is a sequence $\{\nu_i\} \longrightarrow \infty$ of measured laminations converging projectively to ν such that $i(\gamma, \nu_i) \leq \ell_{f_{\lambda,i}}(\gamma) \leq i(\gamma, \nu_i) + C'\ell_{f_0}(\gamma)$ for any measured lamination γ on S. Combining these inequalities with the Efficiency of pleated surfaces, we get

$$i(\gamma, \nu_i) - Ca(\lambda_i, \gamma) \leq \ell_{\rho_i}(\gamma) \leq i(\gamma, \nu_i) + C'\ell_{f_0}(\gamma). \tag{8.2}$$

Set $\nu = \bigcup_{S \subset \partial M} \nu_S$ and denote by $S(\nu)$ its minimal supporting surface. Let $F \subset \partial M$ be an essential subsurface. It follows from the definition of ν that $\pi_1(F)$ has a fixed point in T if and only if F is disjoint from $S(\nu)$ (up to isotopy). It follows then from [60, Theorem IV.1.2] that there is a collection Σ_ν of essential I-bundles, solid tori and thickened tori such that $S(\nu) = \partial\Sigma \cap \partial M$. Assuming that M is not an I-bundle, consider an essential annulus $A \subset \partial\Sigma$. By assumption, $i(\partial A, \mu) > 0$. This is possible only if $i(\mu, \nu) > 0$. Then we get $i(\mu_i, \nu_i) \longrightarrow \infty$ and $\ell_{\rho_i}(\mu_i) \longrightarrow \infty$ by inequality (8.2). This contradiction concludes the proof. □

8.5 Manifolds with Compressible Boundary

In the previous section we saw that with some additional work, an analogue of the Double Limit Theorem could be established for manifolds with incompressible boundary. To prove a similar result in full generality, we need to consider manifolds with compressible boundary. As we will see in this section, some of the results that were crucial in each proof either are not known in this level of generality or fail to be true.

The first step in both Thurston's and Otal's proof was Lemma 8.3.1 and an essential hypothesis in its proof (see [7, Theorem 3] and [75, Theorem 5.1.1]) is that the domain of discontinuity is simply connected. Canary first overcame this issue in [26] by using new arguments and allowing the multiplicative constant to depend on the injectivity radius of the domain of discontinuity.

Theorem 8.5.1 ([26]) *Given $A > 0$, there exists R such that, if Γ is a nonelementary Kleinian group such that every geodesic in D_Γ has length (in the Poincaré metric on D_Γ) at least A and if c is any closed curve on $S = D_\Gamma/\Gamma$, then*

$$\ell_N(c^*) \leq R\ell_S(c)$$

where $N = \mathbb{H}^3/\Gamma$.

Notice that a geodesic in the domain of discontinuity is a meridian, i.e. it bounds an essential disk. One can prove that a sequence for which the length of a meridian goes to 0 necessarily diverges. Thus the dependence of the constant on A will not be an obstacle when proving convergence results. Furthermore, at the price of dropping the linearity, Sugawa obtained in [80, Proposition 6.1] a universal constant with the following inequality (with the notation of Theorem 8.5.1):

$$\ell_N(c^*) \leq 2\ell_S(c)e^{\ell_S(c)/2} \tag{8.3}$$

This definitively solves the issue of replacing Lemma 8.3.1, even though, as mentioned earlier, Theorem 8.5.1 was already enough.

A more critical obstacle when attempting to extend Thurston's arguments is that the Uniform Injectivity Theorem, which is essential in Thurston's proof of both the Double Limit Theorem and the compactness of AH(acylindrical), does not hold for compressible pleated surfaces. A sequence of compressible pleated surfaces for which the length of a meridian goes to 0 does not converge in any reasonable sense. A way to get around this obstacle was given previously with the Relative Boundedness Theorem (Theorem 8.4.1) where we required a bound on the length of a fixed doubly incompressible multicurve. Combining Theorem 8.4.1 and Sugawa's inequality (8.3), we get:

Theorem 8.5.2 *Let $\gamma \in \partial M$ be a doubly incompressible multicurve and consider a sequence $\{\rho_i\} \in AH(\pi_1(M))$. If $\ell_{\sigma_i}(\gamma)$ is bounded, then ρ_i has a converging subsequence.*

The idea of adding a bound on the length of a well-chosen multi-curve has been pushed further by Canary who enhances the arguments of the proof of the Double Limit Theorem to get:

Theorem 8.5.3 ([27]) *Let H be a handlebody and consider a sequence $\{\rho_i\} \subset \mathrm{int}(AH(\pi_1(H)))$ with Ahlfors–Bers coordinates converging to a Masur domain lamination. If $H = S \times I$ and $\ell_{\rho_i}(\partial S) \leq K$ for all i and some K independent of i, then $\{\rho_i\}$ has a convergent subsequence in $AH(\pi_1(H))$.*

This statement raises another, although less decisive, issue: deciding what condition will replace the assumption that the laminations are binding. In [50], Masur introduced an open subset of $\mathcal{PML}(\partial H)$ for a handlebody H which is now known as the *Masur domain*. Save for some exceptional cases, it consists in projective measured laminations which intersect every projective limits of meridians.

It was conjectured by Thurston (see [27]) that this domain was the appropriate setting to extend the Double Limit Theorem to handlebodies. Later, this definition was extended to compression bodies by Otal in [73] (see also [43]).

When trying to extend Otal's proof to manifolds with compressible boundary, one also encounters important difficulties. As already explained the issue of extending Lemma 8.3.1 and the assumption that laminations are binding are shared by both proofs. When M is not an I-bundle, we still use Culler–Morgan–Shalen's compactification to get a small minimal action of $\pi_1(M)$ on a real tree. If S is a component of ∂M, we get an action of $\pi_1(S)$ on the same real tree through the map $i_* : \pi_1(S) \to \pi_1(M)$ induced by the inclusion. But when S is compressible, this action is not small and hence may not be dual to a measured geodesic lamination. Showing that a measured lamination on the boundary is realised in the tree in order to use the Continuity Theorem becomes problematic. Again, one way to get around this obstacle is to assume a control on the length of some multi-curve whose complement is incompressible. With this idea, one can obtain statements that are close to Theorem 8.5.3 (with the limitations described in Remark 8.3.6), see [74]. Deducing from the work of Culler–Vogtmann [33] that any action of a rank-2 free group is dual to a measured lamination on a compact surface, Otal shows the following:

Theorem 8.5.4 ([74, Theorem 1.5]) *Let H be a genus-2 handlebody and $\{\rho_i\}$ a sequence in $\text{int}(AH(\pi_1(H)))$ with Ahlfors–Bers coordinates converging to a Masur domain lamination whose complementary regions are simply connected. Then $\{\rho_i\}$ has a converging subsequence.*

Before discussing further developments, we should mention the work of Ohshika on free products $\Gamma = \pi_1(S_1) * \pi_1(S_2)$ of two surface groups. In [67], he uses the Culler–Morgan–Shalen compactification and a careful study of actions Γ on real trees when both surface groups have fixed points to prove a convergence result for representations in $AH(\Gamma)$ whose exterior boundary tend to a Masur domain lamination.

The most important breakthrough regarding the convergence of sequences in $AH(\pi_1(M))$ when M has compressible boundary was achieved by Kleineidam–Souto in [43]. By pursuing the study of limits of meridians initiated in [73] and cleverly adapting some arguments from [79], they manage to prove the following:

Theorem 8.5.5 ([43, Corollary 3]) *Let H be a handlebody and $\pi_1(H) \times T \to T$ be a non-trivial small minimal action on a real tree T. Then at least one minimal component of every measured lamination in the Masur domain is realised in T.*

This allows them to use the Continuity Theorem to show that, for a handlebody M, a sequence in $AH(\pi_1(M))$ is precompact if we assume a bound on the length of a sequence of measured laminations converging to a Masur domain lamination. These results are extended to compression bodies in the same paper ([43] and then to compact atoroidal 3-manifolds in [46] (see also [45, Proposition 6.1 and Theorem 6.6]). This leads to some nice generalizations of the Double Limit Theorem such as [65, Theorem 8.1] and [70, Theorem 3.8] although the need to converge to an almost minimal lamination (see Remark 8.3.6) adds some technical hypothesis to the statements.

The final page in this story was written by Kim–Lecuire–Ohshika, [42], who lifted this last limitation with an area argument in a simplicial annulus as in the proof of Efficiency of Pleated Surface and the analysis of limits of boundaries of essential disks and annuli initiated in [73] and pursued in [43] and [46].

Theorem 8.5.6 ([42]) *Let M be a compact orientable irreducible atoroidal 3-manifold, let $\mathcal{C}$ be a connected component of* $\mathrm{int}(AH(\pi_1(M))$ *containing a representation uniformizing M. Let $\{\rho_i\} \subset \mathcal{C}$ be a sequence of convex cocompact representations with Ahlfors–Bers coordinates $\sigma_i \in \mathcal{T}(\partial M)$. If σ_i converges to a doubly incompressible measured lamination, then $\{\rho_i\}$ has a convergent subsequence.*

This statement uses a slight generalization of Masur domain introduced in [46]: a measured geodesic lamamination $\lambda \in \mathcal{ML}(\partial M)$ is *doubly incompressible* if there exists $\eta > 0$ such that $i(\lambda, \partial E) > \eta$ for any essential disc or annulus $E \subset M$.

Notice that if λ is not doubly incompressible, using Dehn twists along annuli, a diverging sequence $\{\rho_i\} \subset AH(\pi_1(M))$ can be constructed so that σ_i tend to λ (compare with [41]).

8.6 Necessary Conditions

The theorems mentioned in the previous sections provide necessary conditions for a sequence to have a converging subsequence. As we already mentioned, Theorems 8.1.3 and 8.5.6 are optimal in the sense that if a measured lamination λ does not satisfy their assumptions, then there is a diverging sequence whose Ahlfors–Bers coordinates tend to λ. On the other hand there are a lot of converging sequences which do not satisfy the condition of Theorem 8.5.6 (Theorem 8.1.3 is simply the special case where $M = S \times I$), i.e. their Ahlfors–Bers coordinates σ_i tend to a measured lamination that is not doubly incompressible. This condition is far from being necessary.

In the quasi-Fuchsian case, some necessary conditions have been established by Ohshika with [68, theorem 3.1] and [72, Theorems 3,5 and 12]. Let $\rho_i : \pi_1(S) \to PSL_2(\mathbb{C})$ be quasi-Fuchsian representations with Ahlfors–Bers coordinates $(\sigma_i^\pm)$ such that $\{\sigma_i^\pm\}$, converge to a (projective) measured lamination $\mu^\pm$. A very rough description of Ohshika's statements could be that if μ^+ and μ^- share something (a leaf or a boundary component of a supporting surface), then $\{\rho_i\}$ diverges.

There is a large gap between these necessary conditions and the sufficient conditions of Theorem 8.5.6. One origin of this gap is the coarseness of Thurston's compactification: it only records the part of the representation that degenerates the fastest. Let us illustrate this idea with an example. Let S be a closed surface and $c, d \subset S$ be two disjoint simple closed curves that are not isotopic and denote by ψ_c, ψ_d the right Dehn twist along c, resp. d. Fix $X \in \mathcal{T}(S)$ and consider for every $i > 0$ the quasi-Fuchsian group $\rho_i : \pi_1(S) \to PSL_2(\mathbb{C})$ with Ahlfors–Bers coordinates $(\psi_c^{i^i} X, \psi_d^{i^i} X)$. It is easy to deduce from Lemma 8.3.1 that $\{\rho_i\}$ has a converging subsequence. Let $\Psi_c : S \times I \to S \times I$ be the Dehn twist along the annulus $c \times I$. One can prove that $\theta_i = \rho_i \circ (\Psi_c^i)_*$ has no converging subsequence. On the other hand, the Ahlfors–Bers coordinates of θ_i, $(\psi_c^i \circ \psi_c^{i^i} X, \psi_c^i \circ \psi_d^{i^i} X)$, have the same limit (c, d) as the coordinates of ρ_i.

A way to have a finer compactification of Teichmüller space is given by the Culler–Morgan–Shalen compactification by actions on Λ-trees (see [57]) and their dual Λ-measured laminations (see [56]). A simpler alternative which ought to give similar results for this specific problem is to look at Hausdorff limits of short pants decompositions or short collections of binding curves. Necessary conditions have also been given by Ohshika with this idea, see [72, Theorem 4]. But there is still a gap which, as illustrated by Brock et al. [19, Example 1.4], cannot be filled within this framework.

Brock, Bromberg, Canary and Lecuire managed to close these gaps in [19] with a different approach based on Masur–Minsky's work on the curve complex (especially [51]).

Theorem 8.6.1 ([19]) *Let S be a compact, orientable surface and let $\{\rho_i\}$ be a sequence in $AH(\pi_1(S))$ with Ahlfors–Bers coordinates $\{\sigma_i^{\pm}\}$. Then $\{\rho_i\}$ has a convergent subsequence if and only if there exists a subsequence $\{\rho_j\}$ of $\{\rho_i\}$ such that $\{\sigma_j^{\pm}\}$ bounds projections.*

The statement is short because the authors have craftily hidden the technicalities in the definition of "bounds projections". To give the precise definition would require too many preliminaries but we will try to convey the spirit. The definition of "bounding projections" consists in two conditions that we will refer as (a) and (b), following [19]. Condition (a) essentially prevents the case where both $\{\sigma_i^+\}$ and $\{\sigma_i^-\}$ tend to filling projective measured laminations with the same support (see also [68]). Condition (b) sees the introduction of *combinatorial parabolics*. Those are simple curves on S for which the behavior of σ_i^+ (for upward pointing combinatorial parabolics) or σ_i^- (for downward pointing ones) indicates that they should be parabolics in the limit (if there was one). Condition (b) essentially says that a simple closed curve on S cannot be simultaneously an upward pointing and a downward pointing combinatorial parabolic. The possibility of wrapping of the algebraic limit, as described in [5], compels us to add some exceptions to this last condition (condition (b)(ii), see also [72, Theorem 6]).

The Proof of Theorem 8.6.1 as well as the proofs of the main Theorems in [72] make extensive use of the works of Masur–Minsky and Minsky [51] and [53]. Masur and Minsky associate a family of simple closed curves $\mathcal{H}_\nu^0$, to a pair of end invariants $\{\sigma^{\pm}\}$. They add some structure to $\mathcal{H}_\nu^0$ to form what they call a *hierarchy* $\mathcal{H}_\nu$. Minsky builds from this hierarchy a model M_ν, i.e. a piecewise Riemannian manifold homeomorphic to $S \times I$ whose metric depends only on the hierarchy $\mathcal{H}_\nu$. Then with the collaboration of Brock et al. [17] and Minsky [53] shows that for any hyperbolic manifold N_ρ with end invariants $\{\sigma^{\pm}\}$, there is a bilipschitz map $M_\nu \to N_\rho$. When ρ is convex cocompact, the end invariants are the Ahlfors–Bers coordinates. In general they are a mixture of conformal structure at infinity and ending laminations which describe the asymptotic behavior of the geometry of the ends of $\mathbb{H}^3/\rho(\pi_1(S))$. This work on the models led to the proof of Thurston's Ending Lamination Conjecture which asserts that a representation $\rho \in AH(\pi_1(S))$ is uniquely defined by its end invariants. One important result of [53] that makes the construction of the model work is the existence of a bound on the length in N_ρ of all the curves of $\mathcal{H}^0_\nu$. Furthermore, this bound depends only on S.

Let us go back to Theorem 8.6.1 and consider a sequence $\{\rho_i\}$ in $AH(\pi_1(S))$ with Ahlfors–Bers coordinates $\{\sigma_i^\pm\}$. When $\{\sigma_i^\pm\}$ bounds projections, Brock–Bromberg–Canary–Lecuire deduce from the structure of $\mathcal{H}^0_{\sigma_i}$ that it contains a family of binding curves independent of i. The convergence follows immediately. On the other hand, Brock–Bromberg–Canary–Minsky [18] and Ohshika [72] use Minsky's model to study the link between the behavior of the end invariants of a sequence and the end invariants of a limit. This leads to the divergence results in [72] and the necessity part in Theorem 8.6.1.

We conclude this section by noticing that all the results we have mentioned so far give conditions for a convergence up to extracting a subsequence. To have convergence of the actual sequence would mean to completely predict the end invariants of the limit. The fact that the geometric limit often differs from the algebraic limit makes such a prediction extremely difficult.

8.7 Some Applications

The original motivation for the Double Limit Theorem and the compactness of AH(acylindrical) was the Hyperbolization Theorem for Haken manifolds:

Theorem 8.7.1 (Hyperbolization Theorem) *Let M be a compact irreducible atoroidal Haken 3-manifold, then the interior of M has a complete hyperbolic structure.*

Even though Thurston never published a complete proof for reasons that he explained in [85], he shared his arguments on multiple occasions and wrote some of them in [84, 87] and [88]. The proof decomposes into two distinct cases, each one using a different convergence result.

In the case of manifolds that fiber over the circle, the Double Limit Theorem is used to construct an invariant metric on the cyclic cover. Thurston's arguments have been summarized in [81] and Otal wrote a complete proof in [75] (with different arguments to prove the Double Limit Theorem, as explained in Sect. 8.3.2).

In the other case, the compactness of AH(acylindrical) is used to establish the Bounded Image Theorem (see [40, Theorem 41]) which allows hyperbolic pieces to be glued together to form a larger hyperbolic manifold. Morgan summarized Thurston's arguments in [54] and Kapovich, [39], and Otal, [76], wrote complete proofs.

The Density Theorem is another example of a proof in which convergence results play an important role.

Theorem 8.7.2 (Bers–Thurston's Density) *Every finitely generated Kleinian group is an algebraic limit of geometrically finite groups.*

This statement resolves a generalization due to Sullivan and Thurston of a conjecture of Bers. Combined with works of Marden and Sullivan, Theorem 8.7.2 shows that the deformation space $AH(\pi_1(M))$ does not have any isolated point. Its proof has been written out by Namazi–Souto, [65], and Ohshika [70], it uses the Tameness Theorem [1] and [25], the Ending Lamination Theorem [17, 53], a convergence Theorem (for example Theorem 8.5.6, but a weaker statement is sufficient) and an additional argument to show that non-realizable laminations are ending laminations (see [65, Theorem 1.4] or [70, Proposition 6.5]). The fact that this proof uses the resolutions of two difficult conjectures is a good illustration of how unfathomable the topology of the deformation space $AH(\pi_1(M))$ can be. Notice that an alternate approach has been developed by Brock–Bromberg [15] when ∂M is incompressible.

Combining the Density Theorem with the Ahlfors–Bers coordinates, we get that $AH(\pi_1(M)$ is the closure of an union of topological balls (assuming that ∂M is incompressible to simplify the statements). Despite this apparent simplicity, various exotic phenomenons have been observed. First Anderson–Canary, [5] proved that two of those balls may have intersecting closures. Then McMullen, [52] and Bromberg–Holt, [24] showed that those balls may self-bump. Lastly, Bromberg, [23], and Magyd [47], concluded that $AH(\pi_1(M))$ may not be locally connected. On the other hand by studying the ways different sequences converge to a point, we can find points where none of these happen:

Theorem 8.7.3 ([22] and [16]) *Let M be a compact atoroidal 3-manifold with incompressible boundary. If ρ is a quasiconformally rigid point in $\partial AH(\pi_1(M))$ then ρ is uniquely approachable. In particular, $AH(\pi_1(M))$ is locally connected at ρ and there is no self–bumping at ρ.*

A representation ρ is *quasiconformally rigid* if $\Omega_\rho/\rho(\pi_1(M))$ is a union of three holed spheres.

The proofs of many more results could illustrate the usefulness of the convergence results presented in this chapter. To drive this point home, let us also mention the work of Bonahon–Otal [13] and Lecuire [46] on bending measured laminations and the work of Namazi [63], Namazi–Souto [64] and Brock–Minsky–Namazi–Souto [20] on models for compact and non-compact hyperbolic 3-manifolds.

We would like to conclude this chapter by mentioning an article of Biringer–Souto, [11], where the authors study sequence of unfaithful representations.

References

1. I. Agol, Tameness of hyperbolic 3-manifolds (2004). arXiv:math/0405568
2. L.V. Ahlfors, Finitely generated Kleinian groups. Amer. J. Math. **86**, 413–429 (1964)
3. L. Ahlfors, L. Bers, Riemann's mapping theorem for variable metrics. Ann. Math. **72**(2), 385–404 (1960)
4. J.W. Anderson, A brief survey of the deformation theory of Kleinian groups, in *The Epstein Birthday schrift*. Geometry & Topology Monographs, vol. 1 (Geometry & Topology, Coventry, 1998), pp. 23–49

5. J.W. Anderson, R.D. Canary, Algebraic limits of Kleinian groups which rearrange the pages of a book. Invent. Math. **126**(2), 205–214 (1996)
6. J.W. Anderson, R.D. Canary, D. McCullough, The topology of deformation spaces of Kleinian groups. Ann. Math. **152**(3), 693–741 (2000)
7. L. Bers, On boundaries of Teichmüller spaces and on Kleinian groups. I. Ann. Math. **91**(2), 570–600 (1970)
8. L. Bers, Spaces of Kleinian groups, in *Several Complex Variables, I (Process Conference, University of Maryland, College Park, MD, 1970)* (Springer, Berlin, 1970), pp. 9–34
9. M. Bestvina, Degenerations of the hyperbolic space. Duke Math. J. **56**(1), 143–161 (1988)
10. M. Bestvina, M. Feighn, Stable actions of groups on real trees. Invent. Math. **121**(2), 287–321 (1995)
11. I. Biringer, J. Souto, Algebraic and geometric convergence of discrete representations into $PSL_2\mathbb{C}$. Geom. Topol. **14**(4), 2431–2477 (2010)
12. F. Bonahon, Geometric structures on 3-manifolds, in *Handbook of Geometric Topology* (North-Holland, Amsterdam, 2002), pp. 93–164
13. F. Bonahon, J.-P. Otal, Laminations measurées de plissage des variétés hyperboliques de dimension 3. Ann. Math. **160**(3), 1013–1055 (2004)
14. J.F. Brock, Continuity of Thurston's length function. Geom. Funct. Anal. **10**(4), 741–797 (2000)
15. J.F. Brock, K.W. Bromberg, On the density of geometrically finite Kleinian groups. Acta Math. **192**(1), 33–93 (2004)
16. J.F. Brock, K.W. Bromberg, R.D. Canary, Y.N. Minsky, Local topology in deformation spaces of hyperbolic 3-manifolds. Geom. Topol. **15**(2), 1169–1224 (2011)
17. J.F. Brock, R.D. Canary, Y.N. Minsky, The classification of Kleinian surface groups, II: the ending lamination conjecture. Ann. Math. **176**(1), 1–149 (2012)
18. J.F. Brock, K.W. Bromberg, R.D. Canary, Y.N. Minsky, Convergence properties of end invariants. Geom. Topol. **17**(5), 2877–2922 (2013)
19. J. Brock, K. Bromberg, R. Canary, C. Lecuire, Convergence and divergence of Kleinian surface groups. J. Topol. **8**(3), 811–841 (2015)
20. J. Brock, Y. Minsky, H. Namazi, J. Souto, Bounded combinatorics and uniform models for hyperbolic 3-manifolds. J. Topol. **9**(2), 451–501 (2016)
21. J.F. Brock, K.W. Bromberg, R.D. Canary, Y.N. Minsky, Windows, cores and skinning maps (2016). arXiv:1601.05482
22. J.F. Brock, K. Bromberg, R.D. Canary, C. Lecuire, Y.N. Minsky, Local topology in deformation spaces of hyperbolic 3-manifolds II. Groups Geom. Dyn. **13**(3), 767–793 (2019)
23. K. Bromberg, The space of Kleinian punctured torus groups is not locally connected. Duke Math. J. **156**(3), 387–427 (2011)
24. K. Bromberg, J. Holt, Self-bumping of deformation spaces of hyperbolic 3-manifolds. J. Diff. Geom. **57**(1), 47–65 (2001)
25. D. Calegari, D. Gabai, Shrinkwrapping and the taming of hyperbolic 3-manifolds. J. Amer. Math. Soc. **19**(2), 385–446 (2006)
26. R.D. Canary, The Poincaré metric and a conformal version of a theorem of Thurston. Duke Math. J. **64**(2), 349–359 (1991)
27. R.D. Canary, Algebraic convergence of Schottky groups. Trans. Amer. Math. Soc. **337**(1), 235–258 (1993)
28. R.D. Canary, Y.N. Minsky, E.C. Taylor, Spectral theory, Hausdorff dimension and the topology of hyperbolic 3-manifolds. J. Geom. Anal. **9**(1), 17–40 (1999)
29. I.M. Chiswell, Nonstandard analysis and the Morgan-Shalen compactification. Quart. J. Math. Oxford Ser. **42**(167), 257–270 (1991)
30. V. Chuckrow, On Schottky groups with applications to kleinian groups. Ann. Math. **88**, 47–61 (1968)
31. M. Culler, J.W. Morgan, Group actions on **R**-trees. Proc. London Math. Soc. **55**(3), 571–604 (1987)

32. M. Culler, P.B. Shalen, Varieties of group representations and splittings of 3-manifolds. Ann. Math. **117**(1), 109–146 (1983)
33. M. Culler, K. Vogtmann, The boundary of outer space in rank two, in *Arboreal Group Theory (Berkeley, CA, 1988)*. Mathematical Sciences Research Institute Publications, vol. 19 (Springer, New York, 1991), pp. 189–230
34. A. Fathi, F. Laudenbach, V. Poénaru, *Travaux de Thurston sur les Surfaces*. Astérisque. Société Mathématique de France, Paris, 1979, vol. 66–67. Séminaire Orsay
35. A. Fathi, F. Laudenbach, V. Poénaru, *Thurston's Work on Surfaces* (Translated from the 1979 French original by Djun M. Kim and Dan Margalit). Mathematical Notes, vol. 48 (Princeton University Press, Princeton, 2012)
36. W.H. Jaco, P.B. Shalen, Seifert fibered spaces in 3-manifolds. Mem. Amer. Math. Soc. **21**(220), viii+192 (1979)
37. K. Johannson, *Homotopy Equivalences of* 3*-Manifolds with Boundaries*. Lecture Notes in Mathematics, vol. 761 (Springer, Berlin, 1979)
38. T. Jørgensen, On discrete groups of Möbius transformations. Amer. J. Math. **98**(3), 739–749 (1976)
39. M. Kapovich, *Hyperbolic Manifolds and Discrete Groups*. Progress in Mathematics, vol. 183 (Birkhäuser Boston, Boston, 2001)
40. R.P. Kent, IV, Skinning maps. Duke Math. J. **151**(2), 279–336 (2010)
41. I. Kim, Divergent sequences of function groups. Differ. Geom. Appl. **26**(6), 645–655 (2008)
42. I. Kim, C. Lecuire, K. Ohshika, Convergence of freely decomposable Kleinian groups. Invent. Math. **204**(1), 83–131 (2016)
43. G. Kleineidam, J. Souto, Algebraic convergence of function groups. Comment. Math. Helv. **77**(2), 244–269 (2002)
44. I. Kra, On spaces of Kleinian groups. Comment. Math. Helv. **47**, 53–69 (1972)
45. C. Lecuire, An extension of the Masur domain, in *Spaces of Kleinian Groups*. London Mathematical Society Lecture Note series, vol. 329 (Cambridge University Press, Cambridge, 2006), pp. 49–73
46. C. Lecuire, Plissage des variétés hyperboliques de dimension 3. Invent. Math. **164**(1), 85–141 (2006)
47. A.D. Magid, Deformation spaces of Kleinian surface groups are not locally connected. Geom. Topol. **16**(3), 1247–1320 (2012)
48. A. Marden, The geometry of finitely generated kleinian groups. Ann. Math. **99**(2), 383–462 (1974)
49. B. Maskit, Self-maps on Kleinian groups. Amer. J. Math. **93**, 840–856 (1971)
50. H. Masur, Measured foliations and handlebodies. Ergodic Theory Dyn. Syst. **6**(1), 99–116 (1986)
51. H.A. Masur, Y.N. Minsky, Geometry of the complex of curves. II. Hierarchical structure. Geom. Funct. Anal. **10**(4), 902–974 (2000)
52. C.T. McMullen, Complex earthquakes and Teichmüller theory. J. Amer. Math. Soc. **11**(2), 283–320 (1998)
53. Y. Minsky, The classification of Kleinian surface groups. I. Models and bounds. Ann. Math. **171**(1), 1–107 (2010)
54. J.W. Morgan, On Thurston's uniformization theorem for three-dimensional manifolds, in *The Smith Conjecture (New York, 1979)*. Pure and Applied Mathematics, vol. 112 (Academic, Orlando, 1984), pp. 37–125
55. J.W. Morgan, Group actions on trees and the compactification of the space of classes of SO(n, 1)-representations. Topology **25**(1), 1–33 (1986)
56. J.W. Morgan, J.-P. Otal, Relative growth rates of closed geodesics on a surface under varying hyperbolic structures. Comment. Math. Helv. **68**(2), 171–208 (1993)
57. J.W. Morgan, P.B. Shalen, Valuations, trees, and degenerations of hyperbolic structures. I. Ann. of Math. **120**(3), 401–476 (1984)

58. J.W. Morgan, P.B. Shalen, An introduction to compactifying spaces of hyperbolic structures by actions on trees, in *Geometry and Topology (College Park, Md., 1983/1984)*. Lecture Notes in Mathematics, vol. 1167 (Springer, Berlin, 1985), pp. 228–240
59. J.W. Morgan, P.B. Shalen, Degenerations of hyperbolic structures. II. Measured laminations in 3-manifolds. Ann. Math. **127**(2), 403–456 (1988)
60. J.W. Morgan, P.B. Shalen, Degenerations of hyperbolic structures. III. Actions of 3-manifold groups on trees and Thurston's compactness theorem. Ann. Math. **127**(3), 457–519 (1988)
61. G.D. Mostow, Quasi-conformal mappings in n-space and the rigidity of hyperbolic space forms. Inst. Hautes Études Sci. Publ. Math. **34**, 53–104 (1968)
62. D. Mumford, A remark on Mahler's compactness theorem. Proc. Amer. Math. Soc. **28**, 289–294 (1971)
63. H. Namazi, Heegaard splittings and hyperbolic geometry. ProQuest LLC, Ann Arbor, MI, 2005. Thesis (Ph.D.)–State University of New York at Stony Brook, 2005
64. H. Namazi, J. Souto, Heegaard splittings and pseudo-Anosov maps. Geom. Funct. Anal. **19**(4), 1195–1228 (2009)
65. H. Namazi, J. Souto. Non-realizability and ending laminations: proof of the density conjecture. Acta Math. **209**(2), 323–395 (2012)
66. K. Ohshika, On limits of quasi-conformal deformations of Kleinian groups. Math. Z. **201**(2), 167–176 (1989)
67. K. Ohshika, A convergence theorem for Kleinian groups which are free products. Math. Ann. **309**(1), 53–70 (1997)
68. K. Ohshika, Divergent sequences of Kleinian groups, in *The Epstein Birthday Schrift*. Geometry & Topology Monographs, vol. 1 (Geometry & Topology Publications, Coventry, 1998), pp. 419–450
69. K. Ohshika, Constructing geometrically infinite groups on boundaries of deformation spaces. J. Math. Soc. Japan **61**(4), 1261–1291 (2009)
70. K. Ohshika, Realising end invariants by limits of minimally parabolic, geometrically finite groups. Geom. Topol. **15**(2), 827–890 (2011)
71. K. Ohshika, Degeneration of marked hyperbolic structures in dimensions 2 and 3, in *Handbook of Group Actions. Vol. III*. Advanced Lectures in Mathematics (ALM), vol. 40 (International Press, Somerville, 2018), pp. 13–35
72. K. Ohshika, Divergence, exotic convergence and self-bumping in quasi-fuchsian spaces (2020). arXiv:1010.0070, to appear in Ann. Fac. Sci. Toulouse
73. J.-P. Otal, Courants géodésiques et produits libres. Thése d'Etat, Université Paris-Sud, 1988
74. J.-P. Otal, Sur la dégénérescence des groupes de Schottky. Duke Math. J. **74**(3), 777–792 (1994)
75. J.-P. Otal, Le théorème d'hyperbolisation pour les variétés fibrées de dimension 3. Astérisque **235**, x+159 (1996)
76. J.-P. Otal, Thurston's hyperbolization of Haken manifolds, in *Surveys in Differential Geometry, Vol. III (Cambridge, MA, 1996)* (International Press, Boston, 1998), pp. 77–194
77. F. Paulin, Topologie de Gromov équivariante, structures hyperboliques et arbres réels. Invent. Math. **94**(1), 53–80 (1988)
78. R.C. Penner, J.L. Harer, *Combinatorics of Train Tracks*. Annals of Mathematics Studies, vol. 125 (Princeton University Press, Princeton,1992)
79. R.K. Skora, Splittings of surfaces. J. Amer. Math. Soc. **9**(2), 605–616 (1996)
80. T. Sugawa, Uniform perfectness of the limit sets of Kleinian groups. Trans. Amer. Math. Soc. **353**(9), 3603–3615 (2001)
81. D. Sullivan, Travaux de Thurston sur les groupes quasi-fuchsiens et les variétés hyperboliques de dimension 3 fibrées sur S^1, in *Bourbaki Seminar, Vol. 1979/80*. Lecture Notes in Mathematics, vol. 842 (Springer, Berlin, 1981)
82. D. Sullivan, Quasiconformal homeomorphisms and dynamics. II. Structural stability implies hyperbolicity for Kleinian groups. Acta Math. **155**(3–4), 243–260 (1985)
83. W.P. Thurston, Three-dimensional manifolds, Kleinian groups and hyperbolic geometry. Bull. Amer. Math. Soc. **6**(3), 357–381 (1982).

84. W.P. Thurston, Hyperbolic structures on 3-manifolds I: Deformations of acylindrical manifolds. Ann. Math. **124**, 203–246 (1986)
85. W.P. Thurston, On proof and progress in mathematics. Bull. Amer. Math. Soc. **30**(2), 161–177 (1994)
86. W.P. Thurston, *The Topology and Geometry of 3-Manifolds*. Lecture Notes (Princeton University Press, Princeton, 1976–79)
87. W.P. Thurston, Hyperbolic structures on three-manifolds ii: surface groups and three-manifolds which fiber over the circle (1998). arXiv:math/9801045
88. W.P. Thurston, Hyperbolic structures on three-manifolds iii: deformations of 3-manifolds with incompressible boundary (1998). arXiv:math/9801058

Chapter 9
Geometry and Topology of Geometric Limits I

Ken'ichi Ohshika and Teruhiko Soma

Abstract In this chapter, we classify completely, up to isometry, hyperbolic 3-manifolds corresponding to geometric limits of Kleinian surface groups isomorphic to $\pi_1(S)$ for a finite-type hyperbolic surface S. In the first of the three main theorems which constitute the basic results of this chapter, we construct bi-Lipschitz model manifolds for such hyperbolic 3-manifolds, which have a structure called brick decomposition and are embedded topologically in $S \times (0, 1)$. In the second theorem, we show that conversely, any such model manifold admitting a brick decomposition with reasonable conditions is bi-Lipschitz homeomorphic to a hyperbolic manifold corresponding to some geometric limit of quasi-Fuchsian groups. In the third theorem, it is shown that we can define end invariants for hyperbolic 3-manifolds appearing as geometric limits of Kleinian surface groups, and that the homeomorphism type and the end invariants determine the isometric type of a manifold. This is analogous to the ending lamination theorem for the case of finitely generated Kleinian groups. These results constitute an attempt to give an answer to the 8th question among the famous 24 questions raised by Thurston.

Keywords Kleinian groups · Geometric limits · Hyperbolic 3-manifolds · Ending lamination theorem

2000 Mathematics Subject Classification Primary 57M50; Secondary 30F40

K. Ohshika (✉)
Department of Mathematics, Gakushuin University, Tokyo, Japan
e-mail: ohshika@math.gakushuin.ac.jp

T. Soma
Department of Mathematical Sciences, Tokyo Metropolitan University, Tokyo, Japan
e-mail: tsoma@tmu.ac.jp

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_9

9.1 Introduction

There are two notions of convergence in the theory of Kleinian group: algebraic convergence and geometric convergence. Algebraic convergence is a convergence with respect to the topology induced from the natural topology on the space of representations of a group into $\mathrm{PSL}_2\mathbb{C}$. On the other hand, geometric convergence corresponds to a convergence of the quotient hyperbolic 3-manifolds with respect to the pointed Gromov–Hausdorff topology. One of the main topics in the theory of Kleinian groups is studying the topological structure of deformation spaces. Deformation spaces have topologies induced from the algebraic convergence. Still, their singularities, for instance, those which are called self-bumping points, are caused by the difference between the algebraic and geometric convergences, as was shown by work of Anderson–Canary [3] and McMullen [33]. This suggests that studying geometric limits is important for understanding the deformation spaces.

For an algebraically convergent sequence of Kleinian groups, its geometric limit, which always exists up to passing to a subsequence, contains the algebraic limit, but may be larger than it in general. The difference between algebraic limit and geometric one was first observed by Jørgensen and Marden. In [22], they gave an example of algebraically convergent sequence of infinite cyclic groups in $\mathrm{PSL}_2\mathbb{C}$ which converges geometrically to a rank-2 parabolic group. This is a typical phenomenon for geometric limits, and is a cause of the existence of geometric limits larger than algebraic ones in more complicated situations such as in the example of Kerckhoff–Thurston [24] which we now explain.

Kerckhoff and Thurston considered a sequence in the Bers slice of a quasi-Fuchsian space of a surface S, parametrised as $(m_0, \tau^n n_0) \in \mathcal{T}(S) \times \mathcal{T}(\bar{S})$ for a Dehn twist τ along an essential simple closed curve c on S, where m_0 and n_0 are arbitrary points in the Teichmüller spaces $\mathcal{T}(S)$ and $\mathcal{T}(\bar{S})$. They took a sequence of quasi-Fuchsian groups representing $(m_0, \tau^n n_0)$ so that it converges algebraically, which can always be done by Bers's compactness theorem, and showed that such a sequence converges geometrically to a group G such that $\mathbb{H}^3/G$ is homeomorphic to $S \times (0, 1) \setminus c \times \{\frac{1}{2}\}$. Here the end $c \times \{\frac{1}{2}\}$ in $S \times (0, 1)$ corresponds to a $\mathbb{Z} \times \mathbb{Z}$-cusp of $\mathbb{H}^3/G$ where a phenomenon as in the case of Jørgensen–Marden occurs. By iterating this kind of procedure, it is also possible to construct an example of a geometric limit G' of quasi-Fuchsian groups such that $\mathbb{H}^3/G'$ has infinitely countably many $\mathbb{Z} \times \mathbb{Z}$-cusps as was shown by Bonahon–Otal [7], (see also Ohshika [38]). In particular, this shows that the geometric limit of quasi-Fuchsian groups isomorphic to $\pi_1(S)$ with a finite type surface S can be infinitely generated.

Another important example of geometric limits of quasi-Fuchsian groups was given by Brock [11]. He considered a homeomorphism $\phi : S \to S$ which is pseudo-Anosov on some essential subsurface H of S and is the identity outside, and a sequence parametrised as $(m_0, \phi^n n_0)$ in the Bers slice as in the case of Kerckhoff–Thurston. He showed that the geometric limit of such a sequence is a Kleinian group G'' such that $\mathbb{H}^3/G''$ is homeomorphic to $S \times (0, 1) \setminus H \times \{\frac{1}{2}\}$, where $H \times \{\frac{1}{2}\}$ corresponds to a pair of geometrically infinite tame ends.

A natural problem arising from these examples is to determine what kind of Kleinian groups can appear as geometric limits of quasi-Fuchsian groups, or more generally as a geometric limit of a sequence in the deformation space of a Kleinian group. The purpose of this series of papers is to answer this question. In the present chapter, we shall consider only geometric limits of Kleinian groups isomorphic to surface groups preserving the parabolicity, which are sometimes called Kleinian surface groups. In Theorem A, which is the first of the three main theorems of this chapter, we shall give (bi-Lipschitz) model manifolds for geometric limits of Kleinian surface groups and present conditions which the model manifolds should satisfy. In Theorem C we shall show that these conditions are in fact sufficient, i.e. that any model manifold satisfying the conditions in Theorem A is homeomorphic to some geometric limit of quasi-Fuchsian groups. Combining these two theorems, we characterise completely Kleinian groups which can appear as geometric limits of Kleinian surface groups.

Another natural problem is to classify completely hyperbolic manifolds corresponding to geometric limits up to isometries, which is the subject of Theorem D. The classification problem of finitely generated Kleinian groups, which was called the ending lamination conjecture and is now the ending lamination theorem, was solved by Minsky, collaborating with Brock, Canary and Masur [12, 29, 30, 35]. (An alternative approach can be found in Bowditch [10]. See also an excellent expository paper by Lecuire [26].) Since geometric limits of isomorphic non-elementary finitely generated Kleinian groups can be infinitely generated in general, as explained above, the ending lamination theorem is not sufficient for our situation. Using our model manifolds constructed in Theorem A, we shall prove that the homeomorphism type and (generalised) end invariants completely determine the isometry type of geometric limits. Indeed this is what Theorem D claims for geometric limits of Kleinian surface groups.

In [48], Thurston listed 24 questions in the field of hyperbolic 3-manifolds and Kleinian groups which were open at that time. The question 8 reads "Analyse limits of quasi-fuchsian groups with accidental parabolics". Otal in [45], which is a very informative and well-written review of Thurston's paper, interpreted this problem as one of analysing geometric limits of algebraically convergent quasi-Fuchsian groups. The results of the present chapter give a complete answer to Thurston's question 8 interpreted in this way.

There are applications of the results of this chapter, which appeared in [36, 40, 42]. In [42], Theorem A is used to analyse which points on the boundary of the quasi-Fuchsian space can be bumping points. The second paper [40] studies a quotient space of the Bers boundary of Teichmüller space, called the reduced Bers boundary, on which the mapping class group action on the Teichmüller space extends continuously. Mahan Mj and Ohshika in [36] give necessary and sufficient conditions for the Cannon–Thurston maps of quasi-Fuchsian groups to converge pointwise to that of their limit group applying our classification of geometric limits. We refer the reader also to [41] for the overall picture of geometric limits.

An article sequel to this chapter [43], which will appear as a chapter in a later volume of this book, will deal with geometric limits of Schottky groups.

9.2 Main Results

In this section, we shall state the main results of this chapter. We shall also give definitions of terms which are necessary for stating the main results, and a short outline of their proofs.

For a hyperbolic 3-manifold N, we denote by N_0 the complement of the open ε-cusp neighbourhoods in N for $\varepsilon > 0$ less than the three-dimensional Margulis constant, and call it the non-cuspidal part of N. Its homeomorphism type does not depend on the choice of the constant ε. We call the ends of N_0 the relative ends of N. By the resolution of Marden's tameness conjecture by Agol [2] and Calegari and Gabai [13], the relative ends of any hyperbolic 3-manifold with finitely generated fundamental group are topologically tame, i.e. each of them has a neighbourhood homeomorphic to $F \times (0, \infty)$, where $F = F \times \{0\}$ corresponds to the frontier component of a relative compact core of N_0 facing the relative end. It follows then from the results of Bonahon [6] and Canary [14], that such a relative end is either geometrically finite or simply degenerate: the latter means that there is a sequence of closed geodesics tending to the end which are projected in $F \times \mathbb{R}$ to simple closed curves on F whose projective classes converge in the projectivised Masur domain. However, in general, a hyperbolic 3-manifold N with infinitely generated fundamental group may have infinitely many relative ends which are neither geometrically finite nor simply degenerate. We call such (relative) ends *wild*. To our knowledge, suitable invariants of wild ends which play the role of end invariants for tame ends have not been known up to now. Still, we shall show that for hyperbolic 3-manifolds corresponding to geometric limits of surface Kleinian groups, wild ends are controlled in some way and are determined only by the homeomorphism types, as we shall see in Theorem C.

Now, we are going to state our main results. The first theorem, Theorem A, says that every geometric limit of Kleinian surface groups isomorphic to $\pi_1(S)$, for a surface S of finite type, has a bi-Lipschitz model which admits a decomposition into standard blocks, and can be embedded topologically into $S \times (0, 1)$. This gives also necessary conditions which hyperbolic 3-manifolds corresponding to geometric limits of Kleinian surface groups must satisfy. Before stating the theorem, we shall explain terms which will be used in the statement. A detailed account of these notions can be found in Sect. 9.4.1. A *brick* B is a 3-manifold homeomorphic to $F \times J$ for a compact connected essential subsurface F of S with $\chi(F) < 0$ and an interval J which is either closed or half-open. A *brick manifold* is a union of countably many bricks $F_n \times J_n$ which are glued to each other along essential subsurfaces on their fronts $F_n \times \partial J_n$.

In a brick manifold, we attach to the end of each half-open brick either a conformal structure at infinity or an ending lamination (i.e. a filling geodesic lamination). We call the brick geometrically finite in the former case and simply degenerate in the latter. Each half-open end of a brick constitutes an end of M, and the end is called geometrically finite or simply degenerate accordingly. The ending lamination or the conformal structure attached there is called the end invariant. The union of ideal boundaries on which conformal structures are given is called the boundary at infinity of M, and is denoted by $\partial_\infty M$. A brick manifold endowed with these end invariants is called *a labelled brick manifold*. An end of a brick manifold which does not correspond to an end of a half-open brick is called *wild*. A wild end corresponds to a sequence of bricks which accumulates in $S \times (0, 1)$.

We say that a labelled brick manifold admits a block decomposition when the manifold is decomposed into blocks in the sense of Minsky and solid tori in such a way that each block has horizontal and vertical directions coinciding with those of bricks. We also require the block decomposition for a half-open brick to have solid tori whose core curves are vertically projected to simple closed curves converging to the end invariant of the brick. The blocks have standard metrics and we can choose the gluing maps to be isometries. By identifying a solid torus with a Margulis tube which is determined by information coming from the block decomposition, we can put a metric on the labelled brick manifold. We call such a metric *a model metric*. (See Sects. 9.4.4 and 9.4.5 for details.)

Theorem A *Let S be an orientable connected hyperbolic surface of finite type. Let $\{G_n\}$ be a sequence of Kleinian surface groups isomorphic to $\pi_1(S)$ preserving the parabolicity, and converging geometrically to a non-elementary Kleinian group G. Then there are a labelled brick manifold M which admits a block decomposition, and a K-bi-Lipschitz homeomorphism from M with the model metric to the non-cuspidal part N_0 of the hyperbolic 3-manifold $N = \mathbb{H}^3/G$, with the constant K depending only on $\chi(S)$, and which satisfies the following properties.*

(i) *Each component of ∂M is either a torus or an open annulus.*
(ii) *There is no properly embedded incompressible annulus in M whose boundary components lie on distinct boundary components of M.*
(iii) *If there is an embedded, incompressible half-open annulus $S^1 \times [0, \infty)$ in M such that $S^1 \times \{t\}$ tends to a wild end e of M as $t \to \infty$, i.e. such that for any neighbourhood U of e there is t_0 such that $S \times (t_0, \infty)$ is contained in U, then its core curve is homotopic into an open annulus component of ∂M tending to e.*
(iv) *The manifold M is (not necessarily properly) embedded in $S \times (0, 1)$ in such a way that each brick has a form $F \times J$ with an interval J and an essential subsurface F of S with respect to the product structure of $S \times (0, 1)$ and the ends of geometrically finite bricks lie $S \times \{0, 1\}$. (We shall say that geometrically finite ends are peripheral, to refer to the last condition.)*

We call the labelled brick manifold M in this theorem a *model manifold* for the geometric limit N. It should be noted that a result similar to this one was announced in the introduction of the first preprint version of Brock–Canary–Minsky [12]. By (iv), we see that the geometric limit manifold N_0 has at most $-2\chi(S)$ geometrically finite ends.

The following corollary is easily deduced from Theorem A.

Corollary B *Let G be a non-elementary geometric limit of Kleinian surface groups isomorphic to $\pi_1(S)$ preserving the parabolicity for S as in Theorem A. Then $N = \mathbb{H}^3/G$ has at most countably many relative ends.*

The next theorem guarantees the existence of a geometric limit which is bi-Lipschitz equivalent to a brick manifold with the properties in Theorem A provided that there are no two simply degenerate ends with homotopic ending laminations.

Theorem C *Suppose that M is a labelled brick manifold satisfying conditions (i)–(iv) in Theorem A such that the ending laminations of two simply degenerate ends of M are not homotopic to each other in M. (This condition is necessary only when M is homeomorphic to $F \times (0, 1)$ for a compact essential subsurface F of S since ending laminations are filling and by condition (ii) no two ends are homotopic except for this case.) Then M has a block decomposition, and if we put on M the model metric associated with the decomposition, then there exists a non-elementary geometric limit G of quasi-Fuchsian groups which are isomorphic to $\pi_1(S)$ such that $N = \mathbb{H}^3/G$ admits a K-bi-Lipschitz homeomorphism $f : M \to N_0$ which can be extended continuously to a conformal map $\partial_\infty M \to \partial_\infty N$ between the boundaries at infinity for a constant $K \geq 1$ depending only on $\chi(S)$.*

We shall often use the term "uniform bi-Lipschitz map" to mean that its bi-Lipschitz constant depends only on $\chi(S)$, and hence is independent of the end invariants.

By applying Theorem C, we can construct various examples of geometric limits G of quasi-Fuchsian groups isomorphic to $\pi_1(S)$; for instance, one such that N_0 has infinitely many simply degenerate ends and infinitely many wild ends simultaneously.

The last theorem is a classification theorem which is analogous to the ending lamination theorem for the finitely generated case.

Theorem D *Suppose that G_1 and G_2 are non-elementary geometric limits of Kleinian surface groups isomorphic to $\pi_1(S)$ preserving the parabolicity. If $f : \mathbb{H}^3/G_1 \to \mathbb{H}^3/G_2$ is a homeomorphism preserving their end invariants, then f is properly homotopic to an isometry.*

Remark 9.2.1 In the beginning of the present work, we tried to use a more classical topological approach involving only hyperbolic geometry to study topological properties of geometric limits of quasi-Fuchsian groups. Subsequently we found that, by invoking the bi-Lipschitz model theorem by Brock–Canary–Minsky, it is possible to simplify the proofs of some results and moreover to obtain a deeper result on geometric properties of geometric limits. Therefore, we have changed our original plan and adopted the method relying upon the work of [12, 35]. On the other hand, we have noticed that our original approach on geometric limits gives rise to a rather short proof of the bi-Lipschitz model theorem. See Soma [46].

Now we outline the proofs of the main theorems. To prove Theorem A, we shall first apply Minsky's bi-Lipschitz model theorem to each $\mathbb{H}^3/G_n$ for the given sequence of Kleinian groups $\{G_n\}$ in the statement, and get a model manifold M_n which can be decomposed into blocks with a bi-Lipschitz homeomorphism g_n from M_n to $(\mathbb{H}^3/G_n)_0$. We define M and a bi-Lipschitz homeomorphism from M to N_0 to be the geometric limits of M_n and g_n. We shall verify that these satisfy the required conditions (i)–(iv) one by one, among which the most difficult is (iv). Since M is the geometric limit of $\{M_n\}$, each union of finite bricks can be proved to be embedded in $S \times (0, 1)$ preserving the product structures, but this does not imply immediately that the entire M can also be embedded. We shall need to rearrange the embeddings of sub-bricks by twisting them in such a way that the twisting stabilises on each brick, as will be shown in Lemma 9.4.1.

Next we turn to Theorem C. We shall first consider an ascending exhausting sequence of sub-brick-manifolds W_n consisting of finite bricks within the given labelled brick manifold M. These W_n may have very complicated homeomorphism types, but we shall construct from the W_n brick manifolds Z_n corresponding to geometrically finite Kleinian surface groups by applying Thurston's uniformisation theorem for compact irreducible atoroidal 3-manifolds with boundary, whose geometric limit is also M. We shall approximate in the geometric topology these Kleinian groups corresponding to Z_n by quasi-Fuchsian groups, which are the groups we wanted.

Finally, we outline the proof of Theorem D. We are given two geometric limits G_1 and G_2 such that $N_1 = \mathbb{H}^3/G_1$ and $N_2 = \mathbb{H}^3/G_2$ share the same topological type and end invariants. By Theorem A, we can construct a labelled model manifold M of $(N_1)_0$. By our assumption, there is a homeomorphism from M to $(N_2)_0$ preserving the end invariants. In Theorem 9.5.1, which is a generalisation of the bi-Lipschitz model theorem by Brock–Canary–Minsky [12], we shall prove that such a homeomorphism can be homotoped to a uniform bi-Lipschitz homeomorphism. This shows that G_1 and G_2 are quasi-conformally conjugate by a quasi-conformal homeomorphism which is conformal on the domain of discontinuity. The second statement of Corollary B makes it possible to apply McMullen's generalisation of Sullivan's rigidity theorem and we shall be able to show that G_1 and G_2 are conformally conjugate.

9.3 Preliminaries

We refer the reader to Thurston [47], Benedetti and Petronio [4], Matsuzaki and Taniguchi [31], and Marden [27, 28] for the general theory of hyperbolic manifolds and Kleinian groups, and to Hempel [21] for 3-manifold topology.

Throughout this chapter, all manifolds are assumed to be oriented, and all homeomorphisms between manifolds are assumed to be orientation-preserving. When we talk about a surface S, we always assume that it is a connected surface of finite type possibly with punctures and $\chi(S) < 0$. Sometimes, we fix a hyperbolic structure of finite area on it for convenience. We denote by $\Sigma_{0,3}$, $\Sigma_{0,4}$, $\Sigma_{1,1}$ compact surfaces homeomorphic respectively to a three-holed sphere, a four-holed sphere and a one-holed torus.

9.3.1 The Curve Graph and Tight Geodesics

In this subsection we shall review the basic terminology and results on curve graphs and tight geodesics. Most of these results are due to Masur–Minsky and can be found in [29, 30].

A subsurface Σ of S is called *essential* if no component of the frontier of Σ is null-homotopic in S. We also regard S itself as an essential subsurface of S. When Σ is an open annulus we further assume that the frontier of Σ is not homotopic to a puncture of S. We consider both closed essential subsurfaces and open ones. When we consider two essential subsurfaces, we assume that they do not have inessential intersection. If two essential subsurfaces are isotopic, they are assumed to coincide.

Let Σ be a connected surface of finite type, possibly with punctures. In this chapter, when we talk about curve graphs, we only consider the situation where Σ is an open essential subsurface of some fixed surface S, including the case when $\Sigma = S$. The complexity of Σ is defined by $\xi(\Sigma) = 3g + p$, where g is the genus of Σ and p its number of punctures. (For our purpose, this is more convenient than the Euler characteristic $\chi(S)$.) A surface Σ with $\xi(\Sigma) = 3$ (resp. $\xi(\Sigma) = 4$) is homeomorphic to the interior of $\Sigma_{0,3}$ (resp. the interior of either $\Sigma_{0,4}$ or $\Sigma_{1,1}$).

When $\xi(\Sigma) > 4$, we define the *curve graph* $\mathcal{C}(\Sigma)$ of Σ to be a simplicial graph whose vertices are homotopy classes of non-contractible simple closed curves on Σ which are not homotopic to punctures such that two vertices are connected by an edge if and only if they have disjoint representatives. We call a vertex of $\mathcal{C}(\Sigma)$ or its representative a *curve* on Σ. For our convenience, we fix a complete hyperbolic structure on Σ of finite area and take a uniquely determined closed geodesic as a representative for any curve on Σ. The notion of curve graph was first introduced by Harvey [20] and extended and modified in [29, 30, 34]. In the case when $\xi(\Sigma) = 4$, the curve graph $\mathcal{C}(\Sigma)$ is defined so that the vertices are curves on Σ and two curves v, w are joined by an edge if and only if they have minimum geometric intersection, i.e. $i(v, w) = 1$ when Σ is $\mathrm{Int}\Sigma_{1,1}$ and $i(v, w) = 2$ when Σ is $\mathrm{Int}\Sigma_{0,4}$. When Σ is an open annulus embedded in S, we consider the covering $\tilde{\Sigma}$ of S (with a fixed hyperbolic structure) associated to $\pi_1(\Sigma)$ and compactify $\tilde{\Sigma}$ to $\bar{\Sigma}$ by attaching to it its ideal boundary. The vertices of the curve graph $\mathcal{C}(\Sigma)$ are homotopy classes of essential arcs on $\bar{\Sigma}$ fixing the endpoints. Two vertices are connected by an edge if and only if they can be homotoped fixing the endpoints to arcs whose interiors are disjoint.

We put the path metric $d = d_{\mathcal{C}(\Sigma)}$ on $\mathcal{C}(\Sigma)$ by setting the length of each edge to be 1. In the case when $\xi(\Sigma) > 4$, a finite subset v of vertices in $\mathcal{C}(\Sigma)$ is said to constitute a *simplex* if any two curves of v are represented by disjoint and non-parallel simple closed curves on Σ. This naming comes from the fact that they actually span a simplex in the curve complex of Σ. We only use this term and do not need to consider the curve complex itself. The graph $\mathcal{C}(\Sigma)$ is not locally finite but was proved to be Gromov hyperbolic as a metric space by Masur and Minsky [29]. (See also Bowditch [8] for an alternative approach.)

Let $\mathcal{ML}(\Sigma)$ be the space of compact measured laminations on Σ and $\mathcal{UML}(\Sigma)$ the quotient space of $\mathcal{ML}(\Sigma)$ obtained by forgetting the measures, and let $\mathcal{EL}(\Sigma)$ be the subspace of $\mathcal{UML}(\Sigma)$ consisting of filling laminations, which we call the ending lamination space of Σ. Here a lamination μ in $\mathcal{UML}(\Sigma)$ is said to be *filling* if, for any $\mu' \in \mathcal{UML}(\Sigma)$, either $\mu' = \mu$ or μ' intersects μ non-trivially and transversely. (The term "arational lamination" is used in some literature, with the same meaning.) Refer to [5, 17] for the definition and basic facts about measured lamination space.

Gromov showed that there is a natural boundary at infinity for a Gromov hyperbolic space. According to Klarreich [25] (see also Hamenstädt [19]), there exists a homeomorphism k from the Gromov boundary $\partial\mathcal{C}(\Sigma)$ of $\mathcal{C}(\Sigma)$ to $\mathcal{EL}(\Sigma)$ such that a sequence $\{v_i\}$ of vertices of $\mathcal{C}(\Sigma)$ converges to $\beta \in \partial\mathcal{C}(\Sigma)$ if and only if $\{v_i\}$ regarded as a sequence in $\mathcal{UML}(\Sigma)$ converges to $k(\beta)$ in $\mathcal{UML}(\Sigma)$.

Definition 9.3.1 A sequence $g = \{v_i\}_{i\in J}$ of simplices in $\mathcal{C}(\Sigma)$ is called a *tight sequence* if it satisfies one of the following conditions depending on whether $\xi(\Sigma)$ is greater than 4 or not, where J is a finite or an infinite interval of $\mathbb{Z}$.

(i) When $\xi(\Sigma) > 4$, for any vertices w_i of v_i and w_j of v_j with $i \neq j$, we have $d(w_i, w_j) = |i - j|$. Moreover, if $\{i-1, i, i+1\} \subseteq J$, then v_i is represented by the union of all components of $\partial\Sigma_{i-1}^{i+1}$ that are non-peripheral in Σ, where Σ_{i-1}^{i+1} is a subsurface smallest up to isotopy (with respect to the inclusion) in Σ with essential boundary containing the geodesic representatives of all the vertices of v_{i-1} and v_{i+1}.
(ii) When $\xi(\Sigma) = 4$, each v_i is a vertex in $\mathcal{C}(\Sigma)$ and $d(v_i, v_j) = |i - j|$.

The sequence g is said to connect $v_{\inf J}$ with $v_{\sup J}$, where we define $v_{\inf J}$ to be $\lim_{i\to-\infty} v_i$ when $\inf J = -\infty$ and $v_{\sup J}$ to be $\lim_{i\to\infty} v_i$ when $\sup J = \infty$. The surface Σ is called the support of g and is denoted by $D(g)$. The length of g is defined to be $\#J - 1$, where $\#J - 1$ is defined to be ∞ when $\#J = \infty$.

We regard a single vertex as a tight sequence of length 0. It follows from the definition that for any tight sequence $\{v_i\}$, if a vertex w of $\mathcal{C}(\Sigma)$ meets v_i transversely, then w meets at least one of v_{i-1} and v_{i+1} transversely.

For an open essential subsurface F of Σ and a tight geodesic g in $\mathcal{C}(\Sigma)$, we denote by $\phi_g(F)$ the union of simplices on g which are disjoint from F. Here being disjoint means that they can be made disjoint by an isotopy. For a curve c on F, we use the symbol $\phi_g(c)$ to denote $\phi_g(A(c))$, where $A(c)$ is an annular neighbourhood of c. The following property of tight geodesics is essentially used in this chapter.

Lemma 9.3.2 (Lemma 4.10 in [30]) *Let Y be an essential subsurface of Σ and g a tight geodesic in $\mathcal{C}(\Sigma)$. Then $\phi_g(Y)$ consists of 0, 1, 2, or 3 contiguous simplices of g.*

The following theorem is Lemma 5.14 in [35] (see also Theorem 1.2 in [9]), which was crucial in the proof of the ending lamination conjecture.

Theorem 9.3.3 *Let u, w be distinct vertices or laminations in $\mathcal{C}(\Sigma) \cup \mathcal{EL}(\Sigma)$. Then there exists a tight sequence connecting u with w.*

A *marking* on Σ is a simplex in $\mathcal{C}(\Sigma)$ with some of its vertices (possibly none) having transversals. Here a transversal of a curve c is defined to be a vertex of the curve graph of an annular neighbourhood of c. For a marking $\boldsymbol{I}$, we denote by $B(\boldsymbol{I})$ its vertices with the transversals forgotten, and call it the base curves. Suppose that each of $\boldsymbol{I}, \boldsymbol{T}$ is either a marking on Σ or a lamination in $\mathcal{EL}(\Sigma)$. Then a tight sequence $g = \{v_i\}_{i \in I}$ on Σ is said to be a *tight geodesic* with *initial marking* $I(g) = \boldsymbol{I}$ and *terminal marking* $T(g) = \boldsymbol{T}$ if it satisfies the following conditions.

– If $i_0 = \inf J > -\infty$, then v_{i_0} is a vertex of $\mathcal{C}(\Sigma)$ contained in $B(\boldsymbol{I})$. If $\inf J = -\infty$, then $\boldsymbol{I} = \lim_{i \to -\infty} v_i \in \mathcal{EL}(\Sigma)$.
– If $j_0 = \sup J < \infty$, then v_{j_0} is a vertex of $\mathcal{C}(\Sigma)$ contained in $B(\boldsymbol{T})$. If $\sup J = \infty$, then $\boldsymbol{T} = \lim_{j \to \infty} v_j \in \mathcal{EL}(\Sigma)$.

For a simplex v_j of $g = \{v_j\}$ supported on Σ, a component of $\Sigma \setminus v_j$ and an annulus with core curve in v_j is called a *component domain* of v_j, and also a component domain of g. We also define the predecessor $\mathrm{pred}(v_j)$ of v_j to be v_{j-1} if $j \neq 1$, and $I(g)$ if $j = 1$. Similarly we define the successor $\mathrm{succ}(v_j)$. For a component domain Y of v_j, we denote $\mathrm{pred}(v_j)|Y$ by $I(Y, g)$ and $\mathrm{succ}(v_j)|Y$ by $T(Y, g)$. Here in the case when Y is an annulus $\mathrm{pred}(v_j)|Y$ denotes a vertex in $\mathcal{C}(Y)$ which $\mathrm{pred}(v_j)$ determines when $j \neq 1$ and the transversal of the vertex v_j determines if $j = 1$. The same definition applies for $\mathrm{succ}(v_j)|Y$. If $T(Y, g) \neq \emptyset$, then we write $Y \overset{d}{\searrow} g$ and say that Y is forward subordinate to g at v_j. Similarly we write $g \overset{d}{\swarrow} Y$ and say that Y is backward subordinate to g at v_j if $I(Y, g) \neq \emptyset$. If a tight geodesic k is supported on Y, the domain Y is forward subordinate to g at v_j, and $T(k) = T(Y, g)$, we say that k is forward subordinate to g at v_j and denote this by $k \overset{d}{\searrow} g$. Similarly, we define $g \overset{d}{\swarrow} k$.

Definition 9.3.4 A *hierarchy* H of geodesics on S is a family of tight geodesics on essential open subsurfaces of S with the following properties.

(1) There is a unique geodesic g_H in H with $D(g_H) = S$, which we call the *main geodesic*.
(2) Let Y be a component domain of both a simplex v of $g \in H$ and w of $g' \in H$ such that $g \overset{d}{\swarrow} Y \overset{d}{\searrow} g'$. (The geodesics g and g' may be the same.) Then there exists a unique geodesic h in H such that $D(h) = Y$ and $g \overset{d}{\swarrow} h \overset{d}{\searrow} g'$.
(3) For any geodesic g in H other than g_H, there exist geodesics $h, k \in H$ such that $h \overset{d}{\swarrow} g \overset{d}{\searrow} k$.

For a hierarchy H, we define $|H|$ to be the sum of the lengths of the geodesics constituting H.

A hierarchy H is said to be *complete* if for each component domain X of $\xi(X) \neq 3$, there is a geodesic in H supported on X. A geodesic g in a hierarchy in H whose domain $D(g)$ satisfies $\xi(D(g)) = 4$ is called a 4-*geodesic*. A sub-hierarchy of a complete hierarchy H consisting of all the geodesics in H supported on domains with $\xi \geq 4$ is called the 4-sub-hierarchy.

Definition 9.3.5 Let H be a hierarchy of geodesics on S. A *slice* of H is a set of pairs $\sigma = \{(g, v)\}$ of a geodesic $g \in H$ and a simplex v on g which has the following properties.

(1) If (g, v_1) and (g, v_2) are contained in σ, then $v_1 = v_2$.
(2) There is a pair $(g_\sigma, v_\sigma) \in \sigma$ called the bottom pair, and except for the bottom pair every pair $(h, w) \in \sigma$ is supported in a component domain of some other $(k, u) \in \sigma$.

We also call g_σ the bottom geodesic and v_σ the bottom simplex of σ.

A slice σ is said to be *saturated* if for any $(g, v) \in \sigma$ and its component domain D for which there is a geodesic h in H supported on D, there is some simplex w of h such that $(h, w) \in \sigma$. We say that σ is *non-annular saturated* if the above holds provided that D is not an annulus. For a slice σ, **base**(σ) denotes the union of all vertices contained in simplices which appear in σ, which forms a simplex of $\mathcal{C}(D(g_\sigma))$.

9.3.2 Hyperbolic 3-manifolds and Geometric Limits

A *Kleinian group* Γ is a discrete subgroup of $\mathrm{PSL}_2\mathbb{C}$. When Γ contains an abelian subgroup of finite index, it is called *elementary*. In this chapter, we always assume that Kleinian groups are torsion-free, or equivalently that they contain no elliptic elements. Under this assumption, a Kleinian group is elementary if and only if it is isomorphic to a free abelian group of rank at most two. For a Kleinian group Γ, the quotient space $N = \mathbb{H}^3/\Gamma$ is called the *hyperbolic* 3-*manifold corresponding to* Γ.

The *limit set* Λ_Γ of Γ is the set of accumulation points of the orbit space Γx_0 in the closed 3-ball $\mathbb{H}^3 \cup \hat{\mathbb{C}}$ for a fixed point $x_0 \in \mathbb{H}^3$. It should be noted that Λ_Γ is contained in $\hat{\mathbb{C}}$ since Γ acts on $\mathbb{H}^3$ properly discontinuously. The complement of Λ_Γ in $\hat{\mathbb{C}}$ is called the *region of discontinuity* of Γ, and is denoted by Ω_Γ. We can regard N as the interior of the manifold $(\mathbb{H}^3 \cup \Omega_\Gamma)/\Gamma$, which is called the *Kleinian manifold* corresponding to Γ. The boundary at infinity Ω_Γ/Γ is also denoted by $\partial_\infty N$. The Nielsen convex hull H_Γ is the smallest closed convex set in $\mathbb{H}^3$ containing all geodesics with endpoints on Λ_Γ. The Nielsen convex core is also Γ-invariant. Its quotient $C_\Gamma = H_\Gamma/\Gamma$ is called the *convex core* of N. The Kleinian group Γ is said to be *geometrically finite* if the volume of the δ-neighbourhood of C_Γ in N is finite for some $\delta > 0$.

For a positive number ε, the *ε-thin part* $N_{(0,\varepsilon]}$ of N is the set consisting of all points $x \in N$ such that there exists a non-contractible loop l of length $\leq \varepsilon$ based at x. The complement of its interior $N_{[\varepsilon,\infty)} = N \setminus \mathrm{Int} N_{(0,\varepsilon]}$ is called the *ε-thick part* of N. A *Margulis tube* is an embedded, equidistant, tubular neighbourhood of a simple closed geodesic in N. A $\mathbb{Z}$ or a $\mathbb{Z} \times \mathbb{Z}$*-cusp neighbourhood* P is a subset of N such that each component of $p^{-1}(P)$ is a horoball whose stabiliser in Γ is isomorphic to either $\mathbb{Z}$ or $\mathbb{Z} \times \mathbb{Z}$, where $p : \mathbb{H}^3 \to N$ is the universal covering. By Margulis' lemma [47, Corollary 5.10.2], there exists a constant $\epsilon_0 > 0$ independent of Γ, called the *Margulis constant*, such that, for any $0 < \varepsilon < \epsilon_0$, each component of $N_{(0,\varepsilon]}$ is either a Margulis tube or a $\mathbb{Z}$ or a $\mathbb{Z} \times \mathbb{Z}$-cusp neighbourhood. Let $N_0 = N_0^\varepsilon$ be the union of $N_{[\varepsilon,\infty)}$ and all the Margulis tube components of $N_{(0,\varepsilon]}$, which we call the *non-cuspidal part* of N. For any $\varepsilon_1 < \varepsilon_2 < \epsilon_0$, there exists a K-bi-Lipschitz deformation retraction $N_0^{\varepsilon_2} \to N_0^{\varepsilon_1}$ for some constant $K \geq 1$ depending only on ε_1 and ε_2. It should also be noted that that N_0 is a deformation retract of N. Ends of N_0 are called *relative ends* of N. Each component of the boundary ∂N_0 is either a Euclidean torus or a Euclidean open annulus. Since every cusp neighbourhood of N is covered by a horoball in $\mathbb{H}^3$ based at a single point of $\widehat{\mathbf{C}}$, the boundary at infinity $\partial_\infty N_0$ of N_0 is equal to $\partial_\infty N$.

A sequence $\{(X_n, x_n)\}$ of complete metric spaces with base points *converges geometrically* (in the sense of Gromov) to a complete metric space (Y, y) if there exist (K_n, L_n)-quasi-isometric, L_n-dense map $g_n : B_{R_n}(X_n, x_N) \to B_{K_n R_n}(Y, y)$ with $K_n \searrow 1, L_n \searrow 0$ and $R_n \to \infty$, where $B_R(X, x)$ denotes the R-metric ball in X centred at x. A sequence of Kleinian groups $\{G_n\}$ is said to converge *geometrically* to a Kleinian group G if (i) each $\gamma \in G$ is the limit of a sequence $\{\gamma_n\}$ with $\gamma_n \in \Gamma_n$ and (ii) the limit of any convergent sequence $\{\gamma_{n_i}\}$ with $\gamma_{n_i} \in \Gamma_{n_i}$ is an element of G. It is well known that $\{\mathbb{H}^3/G_n\}$ converges geometrically to $\mathbb{H}^3/G$ with basepoints chosen to be the projections of a common point x_0 in $\mathbb{H}^3$ if and only if $\{G_n\}$ converges to G geometrically. Refer to [22], [4, Chapter E] for more details on properties of geometric limits.

Suppose that Σ is an open essential subsurface of S, possibly S itself. The Teichmüller space of Σ is denoted by $\mathcal{T}(\Sigma)$, for which we assume that every frontier or puncture corresponds to a parabolic cusp. For a point $\sigma \in \mathcal{T}(\Sigma)$, the surface Σ with a hyperbolic metric representing σ is denoted by $\Sigma(\sigma)$. A proper map f from $\Sigma(\sigma)$ to a hyperbolic 3-manifold N with $\sigma \in \mathcal{T}(\Sigma)$ is called a *pleated surface* realising a geodesic lamination λ in $\Sigma(\sigma)$ if f satisfies the following conditions.

(i) f maps each parabolic cusp of $\Sigma(\sigma)$ to a parabolic cusp in N.
(ii) The path-metric induced from N by f coincides with σ, that is, for any rectifiable path α in $\Sigma(\sigma)$, its image $f(\alpha)$ is also a rectifiable path in N with $\mathrm{length}_{\Sigma(\sigma)}(\alpha) = \mathrm{length}_N(f(\alpha))$.
(iii) $f(l)$ is a geodesic in N for each leaf l of λ.
(iv) For each component Δ of $\Sigma \setminus \lambda$, the restriction $f|\Delta$ is a totally geodesic immersion into N.

A relative end e of hyperbolic 3-manifold N is said to be *topologically tame* if there is a properly embedded compact surface F in N_0 which separates a submanifold containing e homeomorphic to $F \times [0, \infty)$. All topologically tame ends of hyperbolic 3-manifolds considered in this chapter are assumed to be *incompressible*, i.e. the inclusion $F \subset N$ is π_1-injective. A topologically tame relative end e is called *geometrically finite* if e has a neighbourhood which intersects no closed geodesics. (Here we need to assume e to be topologically tame since we are considering also the case when $\pi_1(N)$ is infinitely generated.) For a geometrically finite end, the conformal structure $\nu(e)$ on the component of $\partial_\infty N$ corresponding to e is defined to be the end invariant of e. If Γ itself is geometrically finite, then every relative end of N is geometrically finite.

As was shown by Bonahon [6], if e is topologically tame and incompressible but not geometrically finite, then there exists a sequence of closed geodesics tending to e in a neighbourhood $E \cong F \times [0, \infty)$ of e which are homotopic in E to essential simple closed curves c_n on F. Moreover, it is shown in [47] that $\{c_n\}$ converges in $\mathcal{UML}(\mathrm{Int}F)$ to a lamination $\nu(e)$ contained in $\mathcal{EL}(\mathrm{Int}F)$ which is determined uniquely, independently of the choice of closed geodesics tending to e. This $\nu(e)$ is called the *ending lamination* of e. In this situation, we say that the relative end e is *simply degenerate* and define the end invariant of e to be the ending lamination $\nu(e)$. An end which is not topologically tame is called *wild*. (Recall that we are *not* assuming the fundamental group of N to be finitely generated.) No reasonable invariant for a wild end is known up to now. This forces us to define the *end invariants* of N to be only those of topologically tame relative ends of N.

9.3.3 Least-Area Surfaces

In some arguments in this chapter, it will be necessary to homotope an immersion to an embedding or to make intersecting surfaces disjoint by a homotopy. For that, we shall make use of the following result proved by Freedman–Hass–Scott [18].

Theorem 9.3.6 ([18]) *Let M be an orientable irreducible Riemannian 3-manifold possibly with boundary having inward (or zero) mean curvature vectors, and $f : F \to M$ a proper continuous map from a compact (possibly disconnected) surface F which is properly homotopic to an incompressible embedding. Then any least-area surface f_0 homotopic to f is either an embedding or can be homotoped to an embedding within an arbitrarily small regular neighbourhood of its image. In particular, considering a Riemannian metric on M which is multiplied by a very large scalar outside a regular neighbourhood of the image of f, it follows that f can be homotoped to an embedding which is contained in an arbitrarily small regular neighbourhood of the image of f.*

9.4 Brick Manifolds

9.4.1 Embeddings of Brick Manifolds with Infinite Bricks

We shall first introduce some notation to denote the union of sets in a family. The notation will be convenient in the following discussion on brick manifolds. Let $\mathcal{Y} = \{Y_\alpha\}_{\alpha \in A}$ be a family of subsets of some set X. We denote by $\bigvee \mathcal{Y}$ the subset $\bigcup_{\alpha \in A} Y_\alpha$ of X. It should be noted that even when we consider a sequence of families $\{\mathcal{Y}_n\}$ of subsets of X, the union $\bigvee \mathcal{Y}_n$ is taken for each n.

Now we shall give a precise definition of brick manifolds, upon which we have touched lightly before stating the main results in Sect. 9.2. As we explained there, model manifolds of geometric limits which we shall use to prove our main results have structures of brick manifolds.

Throughout this subsection, S denotes some fixed surface of finite type with $\xi(S) \geq 4$. A brick is a 3-manifold homeomorphic to $F \times J$ for a compact essential subsurface F of S with $\xi(F) \geq 3$ and J is either $[0, 1]$ or $[0, 1)$ or $(0, 1]$. In the latter two cases, the brick is said to be *half open*. We define $\xi(B)$ to be $\xi(F)$. For a brick B, we set $\partial_- B = F \times \{0\}$ and $\partial_+ B = F \times \{1\}$ and call them the *upper front* and the *lower front* respectively, even when B is half open. When B is half open, a front which is not contained in B is called the *ideal front* of B. On the other hand, $\partial F \times J$ is called the vertical boundary of B, and is denoted by $\partial_{\mathrm{v}} B$. A brick $B = F \times J$ has two foliations: the horizontal (codimension-1) foliation whose leaves consist of $F \times \{t\}$ and vertical (codimension-2) foliation whose leaves consist of $\{x\} \times J$. A map from a brick to $S \times I$ (where I is an interval in $\mathbb{R}$) is said to be *leaf-preserving* when leaves of the horizontal and the vertical foliations are mapped to leaves of the corresponding foliation of the range. Here, for $S \times I$, the horizontal foliation consists of $S \times \{t\}$ whereas the vertical foliation consists of $\{x\} \times I$.

Before defining brick complexes and brick manifolds in general, we shall first define finite brick complexes and finite brick manifolds. A finite brick complex is a family of finitely many bricks $\mathcal{K} = \{B_1, \ldots, B_m\}$ realised as subsets of a 3-manifold with pairwise disjoint interiors satisfying the following two conditions:

(1) $\bigcup_{i=1}^m B_i$ is connected.
(2) For any two bricks B_i, B_j in $\mathcal{K}$ with $F_{ij} = B_i \cap B_j \neq \emptyset$, there exists a leaf-preserving embedding $\eta : B_i \cup B_j \to S \times [-1, 1]$ with $\eta(B_i) \subset S \times [-1, 0]$, $\eta(B_j) \subset S \times [0, 1]$ such that $\eta(F_{ij})$ is an essential subsurface of $S \times \{0\}$ with $\xi(\eta(F_{ij})) \geq 3$.

The union $\bigvee \mathcal{K}$ is called a *finite brick manifold* with brick decomposition $\mathcal{K}$. We call F_{ij} in the second condition the *joint* of B_i and B_j. A joint F_{ij} is said to be *inessential* if $\partial_- B_i = F_{ij} = \partial_+ B_j$.

Now we define brick complexes and brick manifolds. Let $\{\mathcal{K}_n\}_{n=1}^\infty$ be an ascending sequence of finite brick complexes. Then the union $\mathcal{K} = \bigcup_{n=1}^\infty \mathcal{K}_n$ is called a brick complex, and $\bigvee \mathcal{K}$ is said to be a brick manifold with brick decomposition $\mathcal{K}$. In the situation where a leaf-preserving embedding $\eta : M \to$

$S \times (0, 1)$ of a brick manifold is given, a half-open brick B in $\mathcal{K}$ is said to be *peripheral* with respect to η if the ideal front of $\eta(B)$ is contained in $S \times \{0\} \cup S \times \{1\}$.

The following lemma is a key step in the proof of Theorem A, to whose proof the rest of this subsection is devoted. In the setting of Theorem A, the model manifold M for N is a brick manifold which is a geometric limit of model manifolds for $(\mathbb{H}^3/G_n)_0$. It follows that M contains an ascending exhausting sequence of finite brick manifolds which admit leaf-preserving embeddings into $S \times (0, 1)$. The following lemma then implies that there is a leaf-preserving embedding of M itself into $S \times (0, 1)$.

Lemma 9.4.1 *Let $\{M_n\}$ be a sequence of finite brick manifolds with brick complexes $\mathcal{K}_n$ such that $\mathcal{K}_n \subsetneq \mathcal{K}_{n+1}$. If there exists a leaf-preserving embedding $\eta_n : M_n \to S \times (0, 1)$ for each $n \in \mathbb{N}$, then the brick manifold $M = \bigcup_{n=1}^{\infty} M_n$ has the following properties.*

(i) *There exists a leaf-preserving embedding $\eta_\infty : M \to S \times (0, 1)$.*
(ii) *The ends of M are countable.*
(iii) *If $B \in \mathcal{K}_m$ is peripheral with respect to η_n for all $n \geq m$, then B is also peripheral with respect to η_∞.*

We use the symbols $\mathrm{pr_h} : S \times [0, 1] \to [0, 1]$ to denote the projection to the second factor, and $\mathrm{pr_v} : S \times [0, 1] \to S$ to denote that to the first factor. For any brick $B_i \in \mathcal{K}_n$, we set $\mathrm{pr_h} \circ \eta_n(\partial_- B_i) = \alpha_{i,n}$ and $\mathrm{pr_h} \circ \eta_n(\partial_+ B_i) = \beta_{i,n}$. (Here we regard η_n as extended to ideal fronts continuously.) A half-open brick B_i is peripheral with respect to η_n if and only if either $\alpha_{i,n} = 0$ or $\beta_{i,n} = 1$. For integers n, m with $1 \leq n \leq m$, let $T_{n,m}$ be the subset of $[0, 1]$ consisting of the $\alpha_{i,m}, \beta_{i,m}$ for $B_i \in \mathcal{K}_n$, and set $T_n = T_{n,n}$. Consider the correspondence $\tau_{n,m} : T_n \to T_m$ which transfers $\alpha_{i,n}, \beta_{i,n}$ respectively to $\alpha_{i,m}, \beta_{i,m}$. Note that $\tau_{n,m}$ may not be a map in general. In fact, it may occur that $\alpha_{i,n} = \alpha_{j,n}$ (resp. $\alpha_{i,n} = \beta_{j,n}$) but $\alpha_{i,m} \neq \alpha_{j,m}$ (resp. $\alpha_{i,m} \neq \beta_{j,m}$) etc.

To prove Lemma 9.4.1, we shall make use of the following two kinds of rearrangement for $\{\mathcal{K}_n\}$ and $\{\eta_n\}$. In Rearrangement I, by taking a subsequence and modifying the embeddings η_n, we shall make $\alpha_{i,n}$ and $\beta_{i,n}$ independent of n.

Rearrangement I Fix $n \in \mathbb{N}$. Then by passing to a subsequence, we can make $\tau_{m,m'}|T_{n,m}$ be a map for $m' > m \geq n$. Moreover, since there are only finitely many bricks in $\mathcal{K}_n$, there are only finitely many ways to give them an order. Therefore, we can take a subsequence $\{\mathcal{K}_{n_k}\}$ of $\{\mathcal{K}_m\}_{m \geq n}$ so that the restriction $\tau_{n_k,n_l}|T_{n,n_k} : T_{n,n_k} \to T_{n,n_l}$ is an order-preserving bijection whenever $n_k \leq n_l$. For any $k \geq n$, we define a new embedding η_k to be the old $\eta_{n_k}|M_k$. Repeating the same argument, we can assume that $\tau_{m_1,m_2}|T_{n,m_1} : T_{n,m_1} \to T_{n,m_2}$ is an order-preserving bijection for any triple $n \leq m_1 \leq m_2$. Since η_n and η_m embed $\{\partial_- B_i, \partial_+ B_i \mid B_i \in \mathcal{K}_n\}$ in the same order, we can deform the new η_n by ambient isotopies of $S \times I$ in such a way that we have $\alpha_{i,n} = \alpha_{i,m}$ and $\beta_{i,n} = \beta_{i,m}$ for any $n \leq m$ and any i with $B_i \in \mathcal{K}_n$. In particular, T_n can be made a subset of T_m.

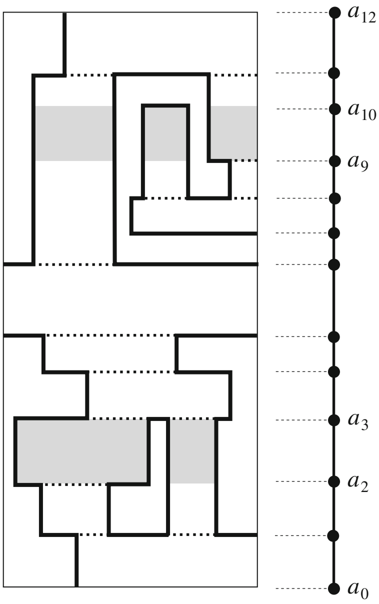


Fig. 9.1 The union of the shaded regions in the lower (resp. higher) level is $\eta_n(R_n^3)$ (resp. $\eta_n(R_n^{10})$)

Rearrangement II Set $T_n = \{a_0, a_1, \ldots, a_t\}$, where elements are subscripted in increasing order, and $R_n^j = \eta_n^{-1}(S \times [a_{j-1}, a_j])$. See Fig. 9.1. Passing again to a subsequence of $\{\eta_n\}$ if necessary, we may assume that, for any $j = 1, \ldots, t$, all $\eta_m|R_n^j$ $(m \geq n)$ define the same embedding up to isotopies and changes of the markings of $S \times [a_{j-1}, a_j]$, i.e. there exists an orientation-preserving homeomorphism $\gamma_{m,n} : S \times [a_{j-1}, a_j] \to S \times [a_{j-1}, a_j]$ with $\gamma_{m,n} \circ (\eta_m|R_n^j) = \eta_n|R_n^j$. For, if we fix a topological type of a compact essential subsurface F of S, there are only finitely many embeddings, up to isotopies and changes of markings, of F into S as an essential subsurface.

We note that this $\gamma_{m,n} \circ (\eta_m|R_n^j)$ may not extend to the entire M_m. In fact even for a brick B in $\mathcal{K}_m \setminus \mathcal{K}_n$ with both $\partial_+ B$ and $\partial_- B$ contained in M_n, it may be possible that $\gamma_{m,n} \circ \eta_m(\partial_- B)$ and $\gamma_{m,n} \circ \eta_m(\partial_+ B)$ are not isotopic.

To construct embeddings of the M_n which stabilise on each brick after finitely many steps, we need to modify the above embeddings η_n by composing 'twists' which will be defined below. Before the definition, we shall observe the local structure of the embeddings $\eta_n(M_n)$ at horizontal levels near the accumulation points of $\cup_m T_m$.

For each $c \in I$ and $n \in \mathbb{N}$, we call $\Sigma_c^{(n)} := (S \times \{c\}) \setminus \mathrm{Int}(\eta_n(M_n))$ the *slit* for $\eta_n(M_n)$ at c. By Rearrangements I and II, for every fixed c, if we take a sufficiently large n_0, the topological type of $\Sigma_c^{(n)}$ does not vary with $n \geq n_0$. The slit $\Sigma_c^{(n)}$ is said to be *stable* if all the $\Sigma_c^{(m)}$ $(m \geq n)$ are homeomorphic. For $c \in I$, we define $\chi_{\mathrm{stab}}(\Sigma_c)$ to be $\chi(\Sigma_c^{(n)})$ for stable $\Sigma_c^{(n)}$. Since the embedding of every brick

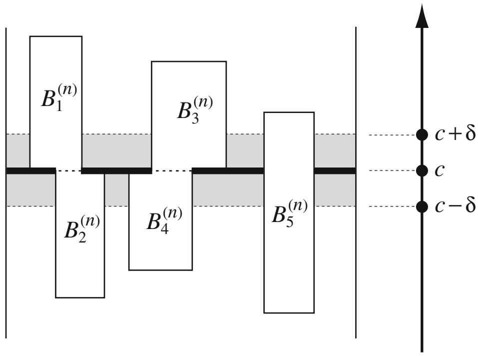


Fig. 9.2 The union of the bold horizontal segments represents $\Sigma_c^{(n)}$. The union of $\Sigma_c^{(n)}$ and the shaded regions is the δ-region $Q_\delta(\Sigma_c^{(n)})$

intersects $S \times \{c\}$ at an essential subsurface with negative Euler characteristic, we see that $\chi(\Sigma_c^{(n)})$ is monotone increasing and once the equality $\chi(\Sigma_c^{(n)}) = \chi_{\mathrm{stab}}(\Sigma_c)$ holds, $\Sigma_c^{(n)}$ is stable.

Let T'_∞ be the set of accumulation points of $T_\infty := \bigcup_{n \geq 1} T_n$. For $c \in T'_\infty$, consider a sufficiently large n such that $\Sigma_c^{(n)}$ is stable. Suppose that $B_1^{(n)}, \ldots, B_k^{(n)}$ are the bricks in $\mathcal{K}_n$ with $\eta_n(B_i^{(n)}) \cap S \times \{c\} \neq \emptyset$ $(i = 1, \ldots, k)$. Take a sufficiently small $\delta > 0$ so that $S \times ([c-\delta, c) \cup (c, c+\delta])$ meets none of the images under η_n of the fronts of $B_i^{(n)}$ $(i = 1, \ldots, k)$. Then we call the set

$$Q_\delta(\Sigma_c^{(n)}) := \big(S \times ([c-\delta, c) \cup (c, c+\delta]) \setminus \eta_n(B_1^{(n)}) \cup \cdots \cup \eta_n(B_k^{(n)})\big) \cup \Sigma_c^{(n)}$$

the *δ-region* of the slit $\Sigma_c^{(n)}$ for $\eta_n(M_n)$. See Fig. 9.2. When c is 0 or 1, we need to modify the definition a little: we define $Q_\delta(\Sigma_c^{(n)})$ to be $S \times (0, \delta]$ when $c = 0$ and $S \times [1-\delta, 1)$ when $c = 1$.

For $m \geq n$, if $\Sigma_d^{(m)}$ $(d \in I)$ is contained in $Q_\delta(\Sigma_c^{(m)}) \setminus \Sigma_c^{(m)}$ then $\chi(\Sigma_d^{(m)}) \geq \chi_{\mathrm{stab}}(\Sigma_c)$. If the equality holds, then $\Sigma_d^{(m)}$ is parallel to $\Sigma_c^{(m)}$ in $S \times [0,1] \setminus \eta_m(M_m)$ (for, since M_m is connected, there cannot be a brick obstructing the parallelism), and even if the strict inequality holds, $\mathrm{pr}_v(\Sigma_d^{(m)})$ is contained in $\mathrm{pr}_v(\Sigma_c^{(m)})$ (up to isotopy). Therefore, in particular if d lies on a side of c from which T_∞ accumulates to c, the strict inequality $\chi(\Sigma_d^{(m)}) > \chi(\Sigma_c^{(m)})$ holds. Since the only bricks that contribute to increase $\chi(\Sigma_c^{(m)})$ are those intersecting $S \times \{c\}$, and their fronts other than those on $S \times \{c\}$ lie outside the δ-region, we see that even for m smaller than n, we have the inequality $\chi(\Sigma_d^{(m)}) \geq \chi(\Sigma_c^{(m)})$. Thus we have shown the following claim.

Claim 9.4.2 *For $c \in T'_\infty$, there exists $\delta(c) > 0$ depending only on c such that $\chi_{\mathrm{stab}}(\Sigma_d) \geq \chi_{\mathrm{stab}}(\Sigma_c)$ if d lies in $[c-\delta(c), c+\delta(c)]$. In particular, if d lies on a side of c from which T_∞ accumulates to c, we have $\chi_{\mathrm{stab}}(\Sigma_d) > \chi_{\mathrm{stab}}(\Sigma_c)$.*

In general, for every n, the inequality $\chi(\Sigma_d^{(n)}) \geq \chi(\Sigma_c^{(n)})$ holds provided that d lies in $[c-\delta(c), c+\delta(c)]$, and $\mathrm{pr}_v(\Sigma_d^{(n)})$ is contained in $\mathrm{pr}_v(\Sigma_c^{(n)})$ up to isotopy.

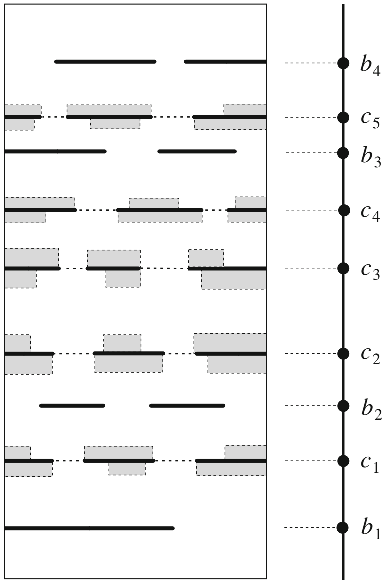


Fig. 9.3 $\bigcup_{i=1}^k [c_i - \delta(c_i), c_i + \delta(c_i)]$ covers $T_\infty \cup T'_\infty$ except for $b_1, \dots, b_u$

For an integer $s \geq 1$, we define $T'_{\infty,s}$ to be the subset of T'_∞ consisting of elements $c \in T'_\infty$ for which $-\chi_{\text{stab}}(\Sigma_c) = s$. Suppose that c is contained in $T'_{\infty,s}$. Then by the claim above, if d lies on a side of c from which T_∞ accumulates to c, and $|d - c| < \delta(c)$, then $-\chi_{\text{stab}}(\Sigma_d) < s$. Taking into account also the side from which T_∞ does not accumulate to c, we can take a possibly smaller $\delta(c)$ such that for *any* $\Sigma_d^{(n)}$ *with* $d \in T_\infty \cup T'_\infty$ *contained in* $Q_{\delta(c)}(\Sigma_c^{(n)}) \setminus \Sigma_c^{(n)}$, *we have* $-\chi_{\text{stab}}(\Sigma_d) < s$. This implies that $(c - \delta(c), c + \delta(c)) \cap T'_{\infty,s} = \{c\}$. It follows that $T'_{\infty,s}$ is a countable subset of $[0, 1]$ for every s, and hence so is T'_∞.

By making $\delta(c)$ smaller if necessary, we can assume that for any $c, c' \in T'_\infty$, either $[c - \delta(c), c + \delta(c)]$ and $[c' - \delta(c'), c' + \delta(c')]$ are disjoint or one of them contains the other. Since $T_\infty \cup T'_\infty$ is compact, there exists a finite subset $\{c_1, \dots, c_k\}$ of T'_∞ such that $\bigcup_{i=1}^k [c_i - \delta(c_i), c_i + \delta(c_i)]$ covers $T_\infty \cup T'_\infty$ except for finitely many elements $b_1, \dots, b_u$ of T_∞. See Fig. 9.3.

For a point $a \in T_\infty$ we define $c(a)$ to be a point in T'_∞ such that $[c(a) - \delta(c(a)), c(a)) \cup (c(a), c(a) + \delta(c(a))]$ contains a and is the smallest among such sets with respect to the inclusion. In the case when there is no such set, i.e. if a is among $b_1, \dots, b_u$, we define $c(a)$ to be 1 by convention.

Now we shall define maps called twists, which will be used to modify embeddings. Let F be a compact essential subsurface of $S \times \{a\}$ with $0 < a < 1$ and $\varphi : F \to F$ an orientation-preserving homeomorphism such that $\varphi|\partial F$ is the identity. Consider a 3-manifold N_φ obtained from $S \times [0, 1] \setminus \text{Int}F$ by identifying the $(\pm)$-sides $F^{(\pm)}$ of F by $\varphi : F^{(-)} \to F^{(+)}$ instead of the identity. The original $S \times [0, 1] \setminus \text{Int}F$ is naturally regarded as a subset of N_φ. We say that N_φ is the manifold obtained from $S \times [0, 1] \setminus \text{Int}F$ by the *φ-twist along F*. The manifold thus

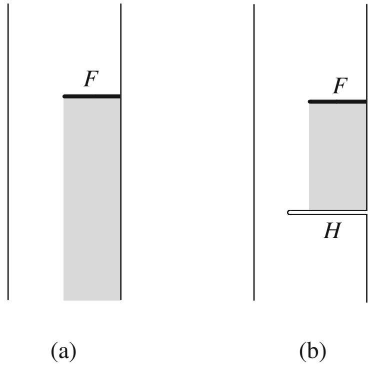


Fig. 9.4 The shaded parts represent the affected regions

obtained is homeomorphic to $S \times [0, 1]$, by a homeomorphism which we specify as follows. Let C_0 be either $F \times [0, a)$ or $F \times (a, 1]$. Then we have a homeomorphism $\xi_0 : N_\varphi \to S \times [0, 1]$ such that $\xi_0|(N_\varphi \setminus C_0)$ is the identity, whereas $\xi_0|C_0$ is $\varphi^{-1} \times \mathrm{id}_{[0,a)}$ if C_0 is $F \times [0, a)$, and $\varphi \times \mathrm{id}_{(a,1]}$ if C_0 is $F \times (a, 1]$. The part of N_ϕ where the homeomorphism is not the identity is called the *affected region* of the twist. In the present case, C_0 is the affected region. See Fig. 9.4a.

For the proof of Lemma 9.4.1, we need to reduce the affected region using the following trick. Let H be a non-peripheral horizontal essential subsurface in $S \times [0, 1]$ with $\mathrm{pr}_\mathrm{v}(H) \supset \mathrm{pr}_\mathrm{v}(F)$ which lies in $S \times \{b\}$ for some b with $F \times \{b\} \subset C_0$. Then there exists a homeomorphism $\xi_1 : N_\varphi \setminus H \to S \times [0, 1] \setminus H$ whose affected region is $C_1 = F \times \langle b, a\rangle$, where $\langle a, b\rangle$ denotes (a, b) if $b > a$ and (b, a) if $a > b$, i.e. $\xi_i|N_\varphi \setminus C_1$ is the identity. See Fig. 9.4b. In the following proof of Lemma 9.4.1, we shall use this trick letting b be $c(a) \in T'_\infty$ defined above.

Now we are ready to formally start the proof of Lemma 9.4.1.

Proof of Lemma 9.4.1 First we shall show part (i). We shall define inductively a leaf-preserving embedding $h_n : M_n \to S \times [0, 1]$ with $h_n^{-1}(S \times [a_{j-1}, a_j]) = \eta_n^{-1}(S \times [a_{j-1}, a_j])$ for $T_n = \{a_0, a_1, \ldots, a_t\}$. Here η_n denotes the one which we obtained after applying Rearrangements I and II for the original η_n. We set $h_1 = \eta_1$. We assume that h_{n-1} has already been defined, and define h_n inductively so that the h_n retain the properties obtained by Rearrangements I and II.

Recall that we defined R_n^j to be $\eta_n^{-1}(S \times [a_{j-1}, a_j])$. By Rearrangement I, we have $R_n^j \cap M_{n-1} = R_{n-1}^j$ for any $j = 1, \ldots, t$. By Rearrangement II for η_n and hence also for h_{n-1}, we see that there exists an embedding $\hat{h}_n^j : R_n^j \to S \times [a_{j-1}, a_j]$ such that $\hat{h}_n^j \circ \eta_n|R_n^j \cap M_{n-1} = h_{n-1}|R_n^j \cap M_{n-1}$. We note that the union of $\hat{h}_n^j \circ \eta_n$ does not necessarily match up on the boundaries of the R_n^h to define an entire embedding from M_n to $S \times [0, 1]$. Let $\hat{T}_n$ be the subset of T_n consisting of elements $a_j \in T_n$ for which $\chi(\Sigma_{a_j}^{(n-1)}) < \chi(\Sigma_{a_j}^{(n)})$, where $\Sigma_{a_j}^{(n-1)} = S \times \{a_j\} \setminus \mathrm{Int}(h_{n-1}(M_{n-1}))$ and $\Sigma_{a_j}^{(n)} = S \times \{a_j\} \setminus \mathrm{Int}(\bigcup_{j=1}^t \hat{h}_n^j \circ \eta_n(R_n^j))$ are slits for h_{n-1} and $\hat{h}_n^j \circ \eta_n$. In particular, $c \in \hat{T}_n$ implies that $\Sigma_c^{(n-1)}$ is unstable.

To construct an embedding on the entire M_n from this $\hat{h}_n^j \circ \eta_n$, we need to perform twist as defined before. For each $a_j \in \hat{T}_n$, we choose an orientation-preserving homeomorphism $\varphi_{a_j} : \Sigma_{a_j}^{(n-1)} \to \Sigma_{a_j}^{(n-1)}$ with $\varphi_{a_j}|\partial \Sigma_{a_j}^{(n-1)}$ being the identity so that $\bigcup_{j=1}^t \hat{h}_n^j \circ \eta_n$ extends to an embedding $\hat{h}_n : M_n \to N_n$, where N_n is the manifold obtained from $S \times [0,1] \setminus \bigcup_{a_j \in \hat{T}_n} \Sigma_{a_j}^{(n-1)}$ by the composition of the φ_{a_j}-twists. By our definition of $\hat{h}_n^j$, if we identify N_n with $S \times [0,1]$ so that the non-affected regions do not move as was explained before, then the difference between $\hat{h}_n|M_{n-1}$ and h_{n-1} is the composition of the φ_{a_j}-twists.

Now we consider reducing the affected region of the φ_{a_j}-twist, to make the embeddings stabilise on each brick. Recall that for a_j, there is a point $c(a_j) \in T'_\infty$ defined above such that $[c(a_j)-\delta(c(a_j)), c(a_j)) \cup (c(a_j), c(a_j)+\delta(c(a_j))]$ contains a_j and is the smallest among such sets. By Claim 9.4.2, we see that $\mathrm{pr}_v(\Sigma_{a_j}^{(n-1)})$ is contained in $\mathrm{pr}_v(\Sigma_{c(a_j)}^{(n-1)})$ for the embedding h_{n-1}. We then reduce the affected region of the φ_{a_j} region to $\mathrm{pr}_v(\Sigma_{a_j}^{(n-1)}) \times \langle a_j, c(a_j)\rangle$ in the way which we explained just before the beginning of the proof of Lemma 9.4.1. In general, there might be other a_k among $T_n = \{a_1, \ldots, a_t\}$ lying between a_j and $c(a_j)$. By our definition of the function c, in this case we have $\langle a_k, c(a_k)\rangle \subset \langle a_j, c(a_j)\rangle$. This implies that the φ_{a_k}-twist does not change the condition that $\mathrm{pr}_v(\Sigma_{a_j}^{n-1})$ is contained in $\mathrm{pr}_v(\Sigma_{c(a_j)}^{(n-1)})$. (This is valid even when $c(a_j) = 1$.) Therefore, there is a homeomorphism $\xi_n : N_n \setminus \bigcup_{a_j \in \hat{T}_n} \Sigma_{c(a_j)}^{(n)} \to S \times [0,1] \setminus \bigcup_{a_j \in \hat{T}_n} \Sigma_{c(a_j)}^{(n)\,\prime}$ such that the affected region of the φ_{a_j}-twist after composing ξ_n is contained in $\Sigma_{c(a_j)}^{(n)} \times \langle a_j, c(a_j)\rangle$, where $\Sigma_{c(a_j)}^{(n)\,\prime}$ is the horizontal essential subsurfaces in $S \times [0,1]$ corresponding to the slit $\Sigma_{c(a_j)}^{(n)}$ and we regard $\langle a_j, c(a_j)\rangle$ as $(a_j, 1]$ when $c(a_j) = 1$. Then $\xi_n \circ \hat{h}_n$ extends to a leaf-preserving embedding $h_n : M_n \to S \times [0,1]$, whose restriction to M_{n-1} coincides with h_{n-1} outside the affected regions, but h_n may not be an extension of h_{n-1} in the affected regions. For the sequence of embeddings h_n thus obtained, we shall show that the restriction $h_n|B$ to each brick $B \in \mathcal{K}$ is eventually the same map even though n varies.

Let B be a brick of $\mathcal{K}$. Then, there is a number m such that M_m contains B. Recall that $b_1, \ldots, b_u \in T_\infty$ are the points which are not contained in $\bigcup_{i=1}^k [c_i - \delta(c_i), c_i + \delta(c_i)]$. Take a sufficiently large $w_0 \in \mathbb{N}$ such that $w_0 > m$, and such that all the $\Sigma_j^{(w_0)}$ are stable for $j \in \{b_1, \ldots, b_u\}$. This also means that all the twists along the slits at $b_1, \ldots, b_u$ are already done by the w_0-th step. For $n > w_0$, consider a twist performed in the construction of h_n at a. If $S \times \langle a, c(a)\rangle$ is disjoint from $h_m(B)$, hence from $h_{n-1}(B)$, the image of B under h_n is the same as that of $h_{n-1}(B)$. Recall also that $\hat{T}_n$ consists of the points where the Euler number of the slit changes at n. Since for each $a \in T_\infty$, there are only finitely many n such that a is contained in $\hat{T}_n$, if there are only finitely many n and twists at a^n for which $S \times \langle a^n, c(a^n)\rangle$ intersects $h_m(B)$, then the image of B stabilises after finitely many steps.

Suppose that there are infinitely many regions $S \times \langle a_j^{n(j)}, c(a_j^{n(j)})\rangle$ $(n(j) \geq w_0)$ with $a_j^{n(j)} \in \hat{T}_{n(j)}$ which intersect $h_m(B)$. We claim that then the $\langle a_j^{n(j)}, c(a_j^{n(j)})\rangle$ are contained in $(\mathrm{pr}_h(\partial_- h_m(B)), \mathrm{pr}_h(\partial_+ h_m(B)))$ except for finitely many of them. Suppose, on the contrary, that infinitely many of them, which we denote again by $\langle a_j^{n(j)}, c(a_j^{n(j)})\rangle$, are not contained in $(\mathrm{pr}_h(\partial_- h_m(B), \mathrm{pr}_h(\partial_+ h_m(B)))$. Passing to a subsequence, we can assume that $\{a_j^{n(j)}\}$ converges to a point $b \in T'_\infty$. This implies that $[b - \delta(b), b + \delta(b)]$ contains $a_j^{n(j)}$ for sufficiently large j, and that $c(a_j^{n(j)})$ is not greater than b and converges to b as $j \to \infty$ since there are only finitely many $a_j^{n(j)}$ such that $\langle a_j^{n(j)}, c(a_j^{n(j)})\rangle$ contains b. Since $S \times \langle a_j^{n(j)}, c(a_j^{n(j)})\rangle$ intersects $h_{n-1}(B)$, the only possibility is that $a_j^{n(j)}$ is contained in $(\mathrm{pr}_h(\partial_- h_m(B), \mathrm{pr}_h(\partial_+ h_m(B)))$ for all large j. Therefore $\langle a_j^{n(j)}, c(a_j^{n(j)})\rangle$ must be contained in $(\mathrm{pr}_h(\partial_- h_m(B), \mathrm{pr}_h(\partial_+ h_m(B)))$, which is a contradiction.

Therefore we have only to consider $\phi_{a_j^{n(j)}}$-twists such that the $\langle a_j^{n(j)}, c(a_j^{n(j)})\rangle$ are contained in $(\mathrm{pr}_h(\partial_- h_m(B), \mathrm{pr}_h(\partial_+ h_m(B)))$. Then $\phi_{a_j}^{n(j)}$ is supported on $\Sigma_{a_j}^{n(j)}$, which is disjoint from $S \times \{a_j\} \cap h_{n(j)-1}(B)$. Therefore the embedding $h_{n(j)-1}(B)$ does not change after performing the $\phi_{a_j}^{n(j)}$-twist. Thus we have shown that the embedding of B stabilises after finite steps. It follows that a leaf-preserving embedding $\eta_\infty : M \to S \times [0, 1]$ is well defined by setting $\eta_\infty|B = h_n|B$ for large n. Since the rearranged η_n maps M_n into $S \times (0, 1)$, so does h_n. Hence the image of η_∞ lies in $S \times (0, 1)$. This completes the proof of (i).

If $B_j^{(m)} \in \mathcal{K}_m$ is peripheral with respect to h_n for all $n \geq m$, then either $\alpha_{j,n} = 0$ or $\beta_{j,n} = 1$ for all $n \geq m$, even after Rearrangement I. It follows from our definition of η_∞ that either $\alpha_{j,\infty} = 0$ or $\beta_{j,\infty} = 1$ holds. This shows part (iii).

Finally, we turn to part (ii). We consider the ends of the embedded image $\eta_\infty(M)$ instead of M itself. Fix a basepoint x_0 in $\eta_\infty(M)$. For an end e of $\eta_\infty(M)$, consider an arc α_e in $\eta_\infty(M)$ emanating from x_0 and tending to e which meets each horizontal leaf of every brick $\eta_\infty(B_j)$ $(B_j \in \mathcal{K})$ with $\alpha_e \cap \eta_\infty(B_j) \neq \emptyset$ transversely in a single point except for the one containing x_0. This implies that α_e meets each $S \times \{c\}$ at most at $-\chi(S)$ points. It follows that $\mathrm{pr_h}(\alpha_e)$ converges to a point $b(e)$ of T'_∞.

Now, for $c \in T'_\infty$, suppose that $e_1, \ldots, e_m$ are distinct m ends of $\eta_\infty(M)$ with $b(e_1) = \cdots = b(e_m) = c$. For a sufficiently large n, these ends are contained in distinct components of $\eta_\infty(M \setminus M_n)$. Therefore, for each $j = 1, \ldots, m$, we can choose a subarc β_{e_j} of α_{e_j} tending to e_j in such a way that β_{e_j} and $\beta_{e_{j'}}$ do not pass through the same brick of $\eta_\infty(M)$ if $j \neq j'$. If we take a sufficiently small $\delta > 0$, then each β_{e_j} passes through the δ-region $S \times [c - \delta, c) \cup S \times (c, c + \delta]$ transversely to the horizontal leaves. It follows that $m \leq -2\chi(S)$ since there are at most $-\chi(S)$ ends lying on $S \times \{c\}$ in each of $S \times [c - \delta, c)$ and $S \times (c, c + \delta]$. Since T'_∞ is a countable set as was seen before, this implies that the ends of $\eta_\infty(M)$ are countable. This completes the proof of part (ii). □

9.4.2 Conditions on Labelled Brick Manifolds

A labelled brick manifold is a brick manifold M in which every half-open brick has either a point in the Teichmüller space or an ending lamination attached to it as follows. Let B be a half-open brick in M which is homeomorphic to $F \times J$, where J is either $[0, 1)$ or $(0, 1]$. Half-open bricks are divided into two categories: geometrically finite bricks and simply degenerate bricks. If B is geometrically finite, then a point in $\mathcal{T}(\mathrm{Int}F)$ is given to B, otherwise an *ending lamination* of B, which is contained in $\mathcal{EL}(\mathrm{Int}F)$ is given. For a geometrically finite brick B, the interior of the ideal front of B is denoted by $\partial_\infty B$, and the point in $\mathcal{T}(\mathrm{Int}F)$ is regarded as a marked conformal structure on $\partial_\infty B$. Also for a simply degenerate brick, the given ending lamination is regarded as attached to the end corresponding to its ideal front.

As in Theorem A, we shall consider labelled brick manifolds M satisfying the following conditions.

A-(1) Every component of ∂M is either a torus or an open annulus.
A-(2) There is no properly embedded essential annulus whose boundary components lie in distinct boundary components of M.
A-(3) If there is an embedded, incompressible half-open annulus $S^1 \times [0, \infty)$ in M such that $S^1 \times \{t\}$ tends to a wild end e, then its core curve is homotopic into an open-annulus component of ∂M tending to e.
A-(4) M is embedded into $S \times (0, 1)$ preserving the horizontal and the vertical leaves in such a way that the ends of geometrically finite bricks are peripheral.
A-(5) Every geometrically finite half-open brick has real front which is an inessential joint: i.e. its real front is contained in the intersection with another brick.

We shall explain the meanings of these conditions briefly. We consider a model manifold M of a geometric limit of Kleinian surface groups, whose corresponding hyperbolic 3-manifold we denote by N. The boundary of M corresponds to the frontier of the non-cuspidal part N_0. This shows that condition A-(1) must be satisfied. Moreover by Margulis's lemma, no essential loops on two distinct components of $\mathrm{Fr}N_0$ can be homotopic to each other. This implies condition A-(2).

To illustrate the meaning of condition A-(3), we consider the situation where M is embedded in $S \times (0, 1)$ preserving the horizontal and vertical leaves, which is required by A-(4). A-(3) says that if M has a wild end e, there must be a sequence of complementary components of M in $S \times (0, 1)$ which tends to the image of e in $S \times (0, 1)$ in such a way that no closed curve can be homotoped to e without being obstructed by a complementary component, except those lying on an annulus boundary component tending to e. We note that the model manifolds of Kleinian surface groups (isomorphic to $\pi_1(S)$) constructed by Minsky can be regarded as labelled brick manifolds as will be explained later. Such brick manifolds can be embedded in $S \times (0, 1)$ preserving the horizontal and the vertical leaves. Lemma 9.4.1 implies that model manifolds of geometric limits can also be embedded in $S \times (0, 1)$ preserving the horizontal and the vertical leaves in such a way that the geometrically finite ends are peripheral. This implies condition A-(4).

The last condition A-(5) is just for convenience in defining a metric on a brick manifold later.

9.4.3 Tight Tube Unions

To construct model manifolds of Kleinian surface groups, Minsky considered a hierarchy of tight geodesics. In his construction, a tight geodesic is realised in the model manifold as a sequence of Margulis tubes. We shall consider a similar realisation of a tight geodesic in the model manifold, which we call a tight tube union.

Consider a brick $B = F \times [0, 1]$ with $\xi(F) > 4$ in a labelled brick manifold. Suppose that we are given a pair of multi-curves $\boldsymbol{I} \times \{0\}$ and $\boldsymbol{T} \times \{1\}$ lying on $\mathrm{Int}\partial_- B$ and $\mathrm{Int}\partial_+ B$ respectively, which represent simplices in $\mathcal{C}(\mathrm{Int}F)$ by identifying $\partial_- B$ and $\partial_+ B$ with F naturally. Let $g = \{v_i\}_{i=0}^n$ be a tight geodesic in $\mathcal{C}(\mathrm{Int}F)$ with $I(g) = \boldsymbol{I}$ and $T(g) = \boldsymbol{T}$. Then $\bigcup_{i=0}^n v_i \times [i/(n+1), (i+1)/(n+1)]$ is a disjoint union $\mathcal{A}_B$ of vertical annuli in B. We call the union $\mathcal{A}_B$ a *tight annulus union* in B connecting $\boldsymbol{I} \times \{0\}$ with $\boldsymbol{T} \times \{1\}$.

Next we consider the case when B is a half-open brick $F \times [0, 1)$ with $\xi(F) > 4$. Since we are not going to put an annulus union or a tube union for geometrically finite bricks, we assume that B is simply degenerate. Suppose then that $\boldsymbol{I} \times \{0\}$ is a multi-curve on $\mathrm{Int}\partial_- B = \mathrm{Int}F$, and that $\boldsymbol{T} \times \{1\}$ is an element of $\mathcal{EL}(\mathrm{Int}\partial_+ B) = \mathcal{EL}(\mathrm{Int}F)$, which is the ending lamination of B. Let $g = \{v_i\}_{i=0}^\infty$ be a tight geodesic ray in $\mathcal{C}(\mathrm{Int}F)$ with $I(g) = \boldsymbol{I}$ and $T(g) = \boldsymbol{T}$. Then the union $\mathcal{A}_B = \bigcup_{i=0}^\infty v_i \times [1 - 1/2^i, 1 - 1/2^{i+1}]$ of vertical annuli in B is called a *tight annulus union* in B connecting $\boldsymbol{I} \times \{0\}$ with $\boldsymbol{T} \times \{1\}$. We can consider a similar construction for a half-open brick $F \times (0, 1]$ when an ending lamination on $\mathrm{Int}\partial_- B$ and a multi-curve on $\mathrm{Int}\partial_+ B$ are given, and define $\mathcal{A}_B = \bigcup_{i=0}^\infty v_i \times [1/2^{i+1}, 1/2^i]$.

When $\xi(F) = 4$, we need to modify our definition to make annuli pairwise disjoint. In this case, we define a tight annulus union $\mathcal{A}_B$ by $\bigcup_{i=0}^n v_i \times [i/(n+1), (2i+1)/(2n+2)]$ if $B = F \times [0, 1]$, by $\bigcup_{i=0}^\infty v_i \times [1 - 1/2^i, 1 - 3/2^{i+2}]$ if $B = F \times [0, 1)$, and $\mathcal{A}_B = \bigcup_{i=0}^\infty v_i \times [3/2^{i+2}, 1/2^i]$ if $B = F \times (0, 1]$.

Let $\mathcal{A}_B = \bigcup_i v_i \times J_i$ be a tight annulus union in a brick B. Take a sufficiently thin annular neighbourhood R_i of v_i on F so that $R_i \times J_i$ are pairwise disjoint in B. Then $\mathcal{V}_B = \bigcup_i R_i \times J_i$ is called a *tight tube union* in B connecting $\boldsymbol{I} \times \{0\}$ with $\boldsymbol{T} \times \{1\}$.

9.4.4 Block Decompositions of Labelled Brick Manifolds

In this subsection, we shall show that a labelled brick manifold M admits a decomposition into blocks in the sense of Minsky provided that its brick decomposition $\mathcal{K}$ satisfies conditions A-(1)–(5) and the following additional condition (EL), which corresponds to the assumption on ending laminations of simply degenerate ends of M given in Theorem C.

(EL) For any two simply degenerate bricks B, B' in $\mathcal{K}$, their ending laminations $\mu(B)$ and $\mu(B')$ are not homotopic in M.

Under conditions A-(1)–(5), this condition is automatically satisfied unless M is homeomorphic to $F \times (0, 1)$ for a compact essential subsurface F of S as we can see in the following way. Let B_1 and B_2 be two simply degenerate bricks with $B_1 = F_1 \times J_1$ and $B_2 = F_2 \times J_2$, where J_1 and J_2 are half-open intervals. Note that each component of $\partial_v B_1$ and $\partial_v B_2$ lies in ∂M. Condition A-(2) shows that $F_1 \times \{t\}$ and $F_2 \times \{t'\}$ cannot be homotopic in M unless M is homeomorphic to $F_1 \times (0, 1)$. Since $\mu(B_1)$ is contained in $\mathcal{EL}(\mathrm{Int}F_1)$ whereas $\mu(B_2)$ lies in $\mathcal{EL}(\mathrm{Int}F_2)$, which means that they are filling on non-homotopic surfaces, they cannot be homotopic in M unless F_1 and F_2 are homotopic in M. Therefore M must be homeomorphic to $F_1 \times (0, 1)$ if B_1 and B_2 have homotopic ending laminations.

Let $\mathcal{K}_{\mathrm{gf}}$ be the subset of $\mathcal{K}$ consisting of geometrically finite bricks, and set $\mathcal{K}_{\mathrm{int}} = \mathcal{K} \setminus \mathcal{K}_{\mathrm{gf}}$. The union $\partial_\infty M = \bigcup_{B \in \mathcal{K}_{\mathrm{gf}}} \partial_\infty B$ is called *the boundary at infinity* of M. Bricks contained in $\mathcal{K}_{\mathrm{int}}$ are called *internal bricks*.

We modify the brick decomposition $\mathcal{K}$ of M by performing the following two operations.

(1) **Removing inessential joints:** Suppose that there is an inessential joint F of two bricks B, B' in $\mathcal{K}_{\mathrm{int}}$. Then we replace B, B' with the single brick $B \cup B'$. In the exceptional case when M is homeomorphic to $F \times (0, 1)$ and has two simply degenerate bricks, this may yield a 'brick' homeomorphic to $F \times (0, 1)$, which was not allowed in our definition. We still allow this operation and call a brick thus obtained an *open brick*.
(2) **Splitting bricks with non-overlapping annuli on the boundary:** Suppose that there is a brick $B = F \times [0, 1]$ in $\mathcal{K}_{\mathrm{int}}$ with a component A of $\partial M \cap \mathrm{Int}\partial_- B$ which does not overlap $\partial M \cap \mathrm{Int}\partial_+ B$. Here an annulus A_1 in B is said to *overlap* a union of annuli $\mathcal{A}$ in B when the vertical projections of A_1 and $\mathcal{A}$ to F intersect essentially. Then we remove $\mathrm{Int}A \times [0, 1]$ from B and split B into two bricks B_1, B_2. We can naturally identify M with $M \setminus A \times [0, 1)$ and regard $(\mathcal{K} \setminus \{B\}) \cup \{B_1, B_2\}$ as a new brick decomposition of M. We can perform the same operation also when there is an annulus in $\partial M \cap \mathrm{Int}\partial_+ B$ which does not overlap $\partial M \cap \mathrm{Int}\partial_- B$.

By repeating these two kinds of operations, we can assume

Assumption 9.4.3

(1) that there is no inessential joint for any two bricks in $\mathcal{K}$,
(2) that for any brick B both of whose fronts $\partial_- B$ and $\partial_+ B$ are real, each component of $\mathrm{Int}\partial_- B \cap \partial M$ overlaps $\mathrm{Int}\partial_+ B \cap \partial M$ and each component of $\mathrm{Int}\partial_+ B \cap \partial M$ overlaps $\mathrm{Int}\partial_- B \cap \partial M$.

By condition A-(1), ∂M is a union of tori and open annuli. Since M is a brick manifold, each of such tori and annuli consists of horizontal annuli and vertical annuli whose interiors are pairwise disjoint, and contains at least one horizontal annulus except for the case when it is a totally vertical annulus. Let H_A be the union of core curves of the horizontal annuli constituting the boundary components of M. (We take one core curve from each horizontal annulus.) For each geometrically finite brick B_i, we fix a multi-curve $s(B_i)$ on its real front F_i which is the shortest pants decomposition of F_i with respect to the hyperbolic structure given to B_i. Note that although we gave a conformal structure on the ideal front, we put the pants decomposition on the real front. Let $\boldsymbol{l}(\mathcal{K})$ be the union of H_A, the $s(B_i)$ for the geometrically finite bricks B_i, and the ending laminations $\mu(B_j)$ for all simply degenerate brick B_j in $\mathcal{K}_{\mathrm{int}}$, which we regard as lying on the ideal fronts. See Fig. 9.5a.

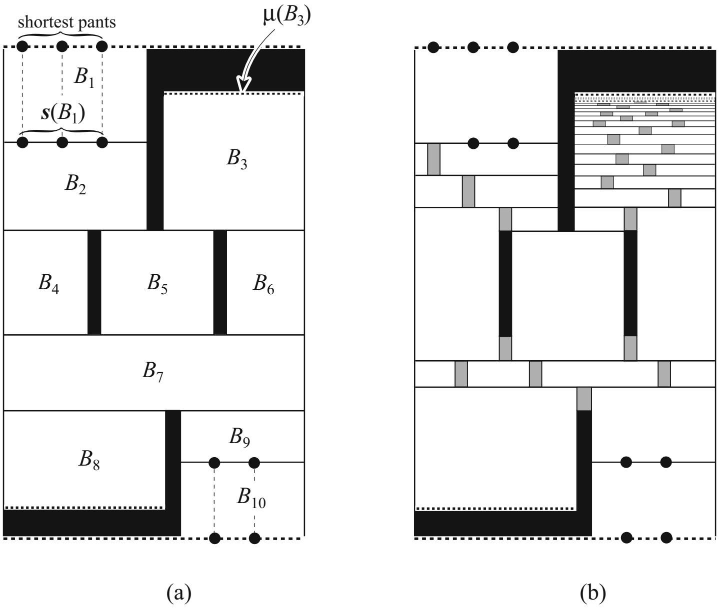


Fig. 9.5 (**a**) B_1, B_{10} is geometrically finite and B_3, B_8 are simply degenerate. The real fronts of B_1 and B_{10} are inessential joints. B_2, B_3, B_7 are connectable. (**b**) The union of shaded rectangles represents $\mathcal{V}^{(1)}$. The white rectangles are bricks in $\mathcal{K}_{\mathrm{int}}^{(1)} \cup \mathcal{K}_{\mathrm{gf}}$

We set $M_{\text{int}} = \bigvee \mathcal{K}_{\text{int}}$. A brick B in $\mathcal{K}_{\text{int}}$ is said to be *connectable* if neither $\boldsymbol{I}(B) = \partial_- B \cap \boldsymbol{l}(\mathcal{K})$ nor $\boldsymbol{T}(B) = \partial_+ B \cap \boldsymbol{l}(\mathcal{K})$ is empty. We use this term 'connectable' considering the fact that if both $\boldsymbol{I}(B)$ and $\boldsymbol{T}(B)$ are non-empty, then we can put a tube union in the brick B in question, which connects $\boldsymbol{I}(B)$ with $\boldsymbol{T}(B)$, at the first stage in the following construction. Notice that if B is a simply degenerate brick, although $\mu(B)$ does not lie inside M, either $\partial_- B$ or $\partial_+ B$ contains $\mu(B)$, and hence intersects $\boldsymbol{l}(\mathcal{K})$. It should be also noted that any brick B in $\mathcal{K}_{\text{int}}$ that has greatest $\xi(B)$ among the bricks in $\mathcal{K}_{\text{int}}$ is connectable unless $\xi(B) = 3$ since we removed inessential joints. We denote by ξ_0 the greatest $\xi(B)$, and suppose that $\xi_0 \geq 5$.

For any connectable brick B of $\mathcal{K}_{\text{int}}$ with $\xi(B) \geq 5$, we take a tight tube union in B connecting $\boldsymbol{I}(B)$ with $\boldsymbol{T}(B)$, and denote it by $\mathcal{V}_B$. In the case when B is an open brick, condition (EL) guarantees that there is a tight tube union connecting $\boldsymbol{I}(B)$ and $\boldsymbol{T}(B)$. We set $\mathcal{V}_B = \emptyset$ if either B is not connectable or $\xi(B) \leq 4$, and define $\bar{\mathcal{V}}^{(1)} = \bigcup_{B \in \mathcal{K}_{\text{int}}} \mathcal{V}_B$. See Fig. 9.5b. Now, if there are two tubes T_1, T_2 in $\bar{\mathcal{V}}^{(1)}$ which are homotopic in $M \setminus (\bar{\mathcal{V}}^{(1)} \setminus (T_1 \cup T_2))$ we merge them into one tube: we can assume that they are vertically isotopic, and by putting a tube between them which is also a thickened annulus, we can make them parts of a larger tube. Repeating this operation, we get a union of tubes $\mathcal{V}^{(1)}$ in which no two tubes are homotopic in the complement of the rest of the tubes.

Let $M_{\text{int}}^{(1)}$ be the closure of $M_{\text{int}} \setminus \mathcal{V}^{(1)}$ in M_{int}. Since $\mathcal{V}^{(1)}$ consists of tubes which are thickened vertical annuli in M_{int}, the 3-manifold $M_{\text{int}}^{(1)}$ has a local product structure (i.e. the one-dimensional vertical direction and two-dimensional horizontal direction) induced from that on M_{int} which in turn comes from the product structures of bricks. Thus $M_{\text{int}}^{(1)}$ has a brick decomposition $\mathcal{K}_{\text{int}}^{(1)}$ allowing a brick also to be an open one having the form $F \times (0, 1)$ such that each brick is the closure of a maximal union of vertically parallel horizontal leaves in $M_{\text{int}}^{(1)}$.

We shall next verify condition A-(2). Suppose, seeking a contradiction, that there is a properly embedded annulus in $M_{\text{int}}^{(1)}$ whose boundaries lie in distinct boundary components. This means that there are horizontal annuli A, A' lying on distinct boundary components whose core curves c, c' are homotopic in $M_{\text{int}}^{(1)}$. Since two simplices on a geodesic at distance greater than 1 have essential intersection, it is impossible that A and A' are contained in tubes in the same brick of $\mathcal{K}_{\text{int}}$. If they are contained in distinct bricks, by our operation modifying $\bar{\mathcal{V}}^{(1)}$ to $\mathcal{V}^{(1)}$, these annuli A and A' lie on the same boundary component, contradicting our assumption. Therefore, condition A-(2) holds also for $M_{\text{int}}^{(1)}$.

Now we consider condition A-(3). Let B be a half-open or open brick in $\mathcal{K}_{\text{int}}^{(1)}$. Suppose that B meets infinitely many original internal bricks $\hat{B}_p$ of $\mathcal{K}_{\text{int}}$. Note that $B \cap \hat{B}_p$ is homeomorphic to $\partial_h B \times [0, 1]$ for each p, except possibly for one which contains an ideal end of B. Then we can take an essential simple closed curve on the horizontal surface of B which is not homotopic into an annulus component of ∂M, and is vertically isotopic into each of the $\hat{B}_p$. This gives rise to an incompressible half-open annulus with core curve not homotopic into an annulus component of ∂M.

This half-open annulus tends to the same wild end of M that the bricks $\hat{B}_p$ tend to. This contradicts condition A-(3) for M. (This end cannot be simply degenerate since each simply degenerate end is contained in one brick of $\mathcal{K}$.) Therefore, any brick in $\mathcal{K}_{\text{int}}^{(1)}$ meets only finitely many bricks of $\mathcal{K}_{\text{int}}$. Also, we can see that an ideal front F of B cannot be contained in the ideal front F' of some simply degenerate brick $B' = F' \times J$ of $\mathcal{K}_{\text{int}}$ since $\mu(B')$ is contained in $\mathcal{EL}(F')$, and hence there is no open annulus in B' disjoint from the tight union of tubes which we extracted to construct $M_{\text{int}}^{(1)}$. Thus we have shown that $M_{\text{int}}^{(1)}$ contains neither half-open nor open bricks. We should note that the greatest $\xi(B)$ for the bricks B in $M_{\text{int}}^{(1)}$, which we denote by ξ_1, is less than ξ_0 since bricks in M_{int} with $\xi = \xi_0$ are all connectable.

Suppose next that $\xi_1 \geq 5$, and consider the union $\mathcal{V}^{(2)}$ of tubes which we obtained by modifying the union of all tight tube unions $\mathcal{V}_B$ for all $B \in \mathcal{K}_{\text{int}}^{(1)}$ in the same way as we defined $\mathcal{V}^{(1)}$ in $\mathcal{K}$ merging homotopic tubes, and the closure $M_{\text{int}}^{(2)}$ of $M_{\text{int}}^{(1)} \setminus \mathcal{V}^{(2)}$ in $M_{\text{int}}^{(1)}$. For the same reason as before, the greatest $\xi(B)$ for the bricks B in $M_{\text{int}}^{(2)}$ is less than ξ_1. Therefore, repeating the same procedure at most $\xi(S) - 4$ times, we reach a brick decomposition $\mathcal{K}_{\text{int}}^{(k)}$ on $M_{\text{int}}^{(k)}$ such that $\xi(B)$ is either 3 or 4 for every brick $B \in \mathcal{K}_{\text{int}}^{(k)}$. In the special case when $\xi_0 = 4$, we are in this situation from the beginning, and hence we set $k = 0$.

Let $\mathcal{V}^{(k+1)}$ be the union of tubes obtained by modifying in the same way as before the union of tight tube unions $\mathcal{V}_B$ for bricks $B \in \mathcal{K}_{\text{int}}^{(k)}$ with $\xi(B) = 4$, and let $\mathcal{K}_{\text{int}}^{(k+1)}$ be the brick decomposition on the closure $M_{\text{int}}^{(k+1)}$ of $M_{\text{int}}^{(k)} \setminus \mathcal{V}^{(k+1)}$ such that each brick is a maximal union of parallel leaves with respect to the horizontal foliation on $M_{\text{int}}^{(k+1)}$. Moving components of $\mathcal{V}^{(k+1)}$ vertically by an ambient isotopy of $M_{\text{int}}^{(k)}$ if necessary, we can assume that for every brick B of $\mathcal{K}_{\text{int}}$, its fronts $\partial_\pm B$ do not go though the gaps of tubes of $\mathcal{V}^{(k+1)}$, i.e. the following holds.

(BB) For any $B \in \mathcal{K}_{\text{int}}$ and $B' \in \mathcal{K}_{\text{int}}^{(k)}$ with $H = (\partial_+ B \cup \partial_- B) \cap B' \neq \emptyset$, each component of $H \setminus \text{Int}\mathcal{V}_{B'}$ is homeomorphic to $\Sigma_{0,3}$.

We set $\mathcal{B}_{\text{int}} = \mathcal{K}_{\text{int}}^{(k+1)}$, $\mathcal{B} = \mathcal{K}_{\text{int}}^{(k+1)} \cup \mathcal{K}_{\text{gf}}$, $M[0]_{\text{int}} = M_{\text{int}}^{(k+1)}$, $M[0] = M[0]_{\text{int}} \cup (\bigvee \mathcal{K}_{\text{gf}})$, and $\mathcal{V} = \bigcup_{m=1}^{k+1} \mathcal{V}^{(m)}$. We call $\mathcal{B}$ a *block decomposition* of $M[0]$ and each element of $\mathcal{B}$ a *block*. Note that each block in $\mathcal{B}_{\text{int}}$ is homeomorphic to either $\Sigma_{0,3} \times J$ or $\Sigma_{1,1} \times J$ or $\Sigma_{0,4} \times J$, where J is a closed or half-open or open interval, since every brick in $\mathcal{K}_{\text{int}}^{(k+1)}$ has ξ at most 4. Also by our definition of bricks for $M_{\text{int}}^{(k+1)}$ no two blocks meet at an inessential joint.

Remark 9.4.4 It may appear that our definition of blocks is slightly different from that of Minsky in [35] as we allow blocks homeomorphic to $\Sigma_{0,3} \times J$. Still the difference is just a minor point since we can convert our block decomposition into that à la Minsky just by cutting a block of the form $\Sigma_{0,3} \times J$ into halves and pasting one of them to the block above it and the other to the one below it.

Each component (i.e. tube) of $\mathcal{V}$ is a solid torus which is foliated by vertically parallel horizontal annuli. For each solid torus V in $\mathcal{V}$, its boundary ∂V is contained

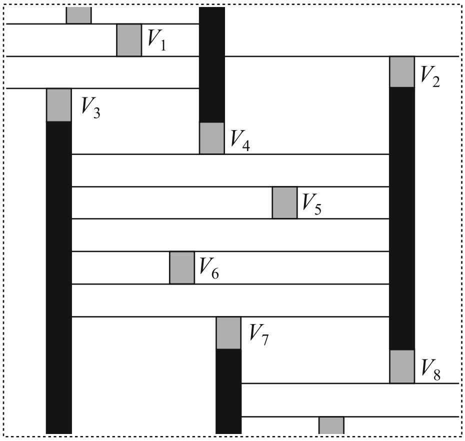


Fig. 9.6 A local picture of M in the case of $k = 0$. The white region is $M[0]$. $V_1 \cup V_5 \cup V_6 \subset \mathcal{V}[0]$ and $V_2 \cup V_3 \cup V_4 \cup V_7 \cup V_8 \subset \mathcal{V} \setminus \mathcal{V}[0]$

in $\partial M[0] \cup \partial M$. If $M[0] \cap V$ consists of two vertical annuli A_1, A_2 for some $V \in \mathcal{V}$, then $\partial V \setminus \mathrm{Int}(A_1 \cup A_2)$ is a union of two horizontal annuli contained in ∂M, and hence each of A_1, A_2 is a properly embedded essential annulus in M. (These annuli cannot be boundary-parallel since, by definition, a brick is not allowed to be a solid torus.) Since each component of ∂M is either an open annulus or a torus, this is possible only when each of A_1, A_2 connects two distinct components of ∂M. This contradicts condition A-(2). Therefore, for any component V of $\mathcal{V}$, the intersection $M[0] \cap V$ is either a torus or an annulus consisting of two vertical annuli and one horizontal annulus. See Fig. 9.6.

Let $\mathcal{V}[0]$ be the union of all components V of $\mathcal{V}$ such that $M[0] \cap V$ is a torus, and set $M^0 = M[0] \cup \mathcal{V}[0]$. Then M^0 is obviously a deformation retract of M and there exists a homeomorphism $\eta_M : M^0 \to M$ homotopic to the inclusion such that the restriction $\eta_M|_{\mathcal{V}[0]}$ is the identity. We often identify the original brick manifold M with M^0 via the map η_M.

9.4.5 Model Metrics on Brick Manifolds

Now we shall define a metric on a brick manifold induced from its decomposition into blocks. We note that our metric is obtained by pasting metrics given on the blocks, and does not depend on the embedding in $S \times [0, 1]$ which we constructed in Sect. 9.4.1.

We shall put a standard metric on each block. However, our metric is slightly different from that of Minsky in [35]. Fix $\varepsilon_1 > 0$ less than the three-dimensional Margulis constant, and a hyperbolic metric on the three-holed sphere $\Sigma_{0,3}$ with respect to which each component of $\partial \Sigma_{0,3}$ is a closed geodesic of length ε_1. Let $B_{0,3}$ be $\Sigma_{0,3} \times [0, 1]$ endowed with the product metric of the hyperbolic metric on $\Sigma_{0,3}$ and the standard metric on $[0, 1]$.

Consider two essential simple closed curves l_0, l_1 on $\Sigma_{0,4}$ (resp. $\Sigma_{1,1}$) with geometric intersection number $i(l_0, l_1) = 2$ (resp. $i(l_0, l_1) = 1$) and set B_α to be a brick of the form $\Sigma_\alpha \times [0, 1]$ for $\alpha \in \{(0, 4), (1, 1)\}$. Let A_- and A_+ be annular neighbourhoods of $l_0 \times \{0\}$ and $l_1 \times \{1\}$ in $\partial_- B_\alpha$ and $\partial_+ B_\alpha$ respectively. We define a piecewise Riemannian metric on B_α such that each component of $\partial_- B_\alpha \setminus \mathrm{Int} A_-$ and $\partial_+ B_\alpha \setminus \mathrm{Int} A_+$ is isometric to $\Sigma_{0,3}$ with the hyperbolic metric given above, all of A_-, A_+ and $\partial_{\mathrm{v}} B_\alpha$ are isometric to the product annulus $S^1(\varepsilon_1) \times [0, 1]$ and $\mathrm{dist}_{B_\alpha}(\partial_- B_\alpha, \partial_+ B_\alpha) = 1$, where $S^1(\varepsilon_1)$ is a round circle in the Euclidean plane of circumference ε_1.

For any brick $B \in \mathcal{B}_{\mathrm{int}}$ of type $\beta \in \{(0, 3), (0, 4), (1, 1)\}$, consider a diffeomorphism $h_B : B_\beta \to B$ such that $h_B(\partial_{\mathrm{v}} B_\beta) = \partial_{\mathrm{v}} B$ and moreover $h_B(A_\pm) = \partial_\pm B \cap \mathcal{V}[0]$ when $\xi(B) = 4$. We can choose these homeomorphisms in such a way that for any B, B' of types $\beta, \beta' \in \{(0, 3), (0, 4), (1, 1)\}$ in $\mathcal{B}_{\mathrm{int}}$ with $F = \partial_+ B \cap \partial_- B' \neq \emptyset$, $((h_{B'}|_F)^{-1} \circ h_B|_{h_B^{-1}(F)})$ is an isometry with respect to the metrics on B_β and $B_{\beta'}$ defined above. Then $M[0]_{\mathrm{int}}$ has a piecewise Riemannian metric induced from those on $B_{0,3}, B_{0,4}, B_{1,1}$ via embeddings $h_B : B_\beta \to M[0]_{\mathrm{int}}$.

We shall next define metrics on geometrically finite bricks. Each geometrically finite brick B of $\mathcal{B}$ is identified with $F \times [-1, \infty)$ preserving the horizontal and the vertical leaves for a compact core F of some open essential subsurface $\mathring{F}$ of S with $\xi(F) \geq 3$. Since $\mathring{F}$ can be identified with $\mathrm{Int} F$, by our definition of geometrically finite bricks, $\mathring{F} = \mathring{F} \times \{\infty\}$ is given a conformal structure. Let $\sigma(B)$ be a complete hyperbolic metric on $\mathring{F}$ which is compatible with the given conformal structure. We regard F as obtained from $\mathring{F}(\sigma(B))$ by deleting the cusp neighbourhoods which are components of $\mathring{F}(\sigma)_{(0,\varepsilon_1)}$. Consider a piecewise Riemannian metric $\tau(B)$ on $\mathring{F}$ obtained by rescaling $\sigma(B)$ on the points of $\mathring{F}$ in such a way that $\tau(B)/\sigma(B)$ is continuous and is equal to 1 on $\mathring{F}(\sigma(B))_{[\varepsilon_1,\infty)}$, and each component of $\mathring{F}(\sigma(B))_{(0,\varepsilon_1]}$ is a Euclidean cylinder with respect to the $\tau(B)$-metric. On the other hand, we put another piecewise Riemannian metric $\upsilon(B)$ on F such that each component of $F(\upsilon(B))_{(0,\varepsilon_1]}$ is a Euclidean cylinder, $F(\upsilon(B))_{[\varepsilon_1,\infty)} \times \{-1\}$ coincides with $M[0]_{\mathrm{int}} \cap B$, and each component of $F(\upsilon(B))_{[\varepsilon_1,\infty)}$ is isometric to $\Sigma_{0,3}$. We choose such a metric so that the identity $F(\tau(B)) \to F(\upsilon(B))$ is uniformly bi-Lipschitz (i.e. the bi-Lipschitz constant is bounded by a constant independent of B and F). We call such a metric as $\upsilon(B)$ a *cylinder*-$\Sigma_{0,3}$ metric on $F \times \{-1\}$. We note that our $\upsilon(B)$ corresponds to the metric $\sigma^{m'}$ given in [35, §8.3].

We put a piecewise Riemannian metric on $F \times [-1, 0]$ such that its restriction to $F \times \{-1\}$ is equal to $F(\upsilon(B))$, its restriction to $F \times \{0\}$ is equal to $F(\tau(B))$, and the induced metric on $F \times \{t\}$ is uniformly bi-Lipschitz to $\tau(B)$ via the identification of $F \times \{t\}$ with F. Recall that F is a compact core of an open surface $\mathring{F}$. We take a diffeomorphism $\eta : F \times [0, \infty) \to \mathring{F} \times [0, \infty)$ such that the restriction $\eta|F \times \{0\}$ is the identity and $\eta(\partial F \times [0, \infty))$ lies on $\mathring{F} \times \{0\}$ so that $\eta|\partial F \times [0, 1]$ is isometric with respect to the metric on $F \times [0, 1]$ defined above and $\tau(B)$ on $\mathring{F} \times \{0\}$. We put on $F \times [0, \infty)$ the induced metric $\eta^*(ds^2)$, where ds^2 is a piecewise Riemannian metric on $\mathring{F} \times [0, \infty)$ defined by

$$ds^2 = \tau(B)e^{2r} + dr^2 \qquad (r \in [0, \infty)). \tag{9.4.1}$$

We define a piecewise Riemannian metric on B by pasting the metrics on $F \times [-1, 0]$ and $F \times [0, \infty)$ along $F \times \{0\}$, which has the metric $\tau(B)$ on both sides. We may assume that the metric on $M[0]_{\rm int}$ and that on B are equal on $M[0]_{\rm int} \cap B = F(\upsilon(B))_{[\varepsilon_1,\infty)} \times \{-1\}$ deforming the map attaching B to $M[0]_{\rm int}$ by an ambient isotopy if necessary. Thus we have obtained a piecewise Riemannian metric on $M[0]$, which we call the *model metric* on $M[0]$. By our construction, each component C of $\partial M[0]$ is either a Euclidean torus or a Euclidean cylinder which has a foliation $\mathcal{F}_C$ whose leaves consist of closed geodesics of length ε_1.

9.4.6 Meridian Coefficients

For a complex number z with $\mathrm{Im}(z) > 0$ and a real number $\eta > 0$, we denote the covering map $\mathbf{C} \to \mathbf{C}/\eta(\mathbb{Z} + z\mathbb{Z})$ by $\pi_{z,\eta}$. For any component V of $\mathcal{V}[0]$, its boundary ∂V has a Euclidean metric induced from the model metric on $M[0]$ as above. Then there is a unique $\omega \in \mathbf{C}$ with $\mathrm{Im}(\omega) > 0$ for which we have an orientation-preserving isometry from the quotient space $\mathbf{C}/\varepsilon_1(\mathbb{Z}+\omega\mathbb{Z})$ to ∂V taking $\pi_{\omega,\varepsilon_1}(\mathbb{R})$ (resp. $\pi_{\omega,\varepsilon_1}(\omega\mathbb{R})$) to a longitude (resp. a meridian) of V. (Here a longitude of V is defined to be a horizontal essential simple closed curve on ∂V.) We denote this ω by $\omega_M(V)$ and call it the *meridian coefficient* of ∂V.

Minsky showed that any meridian coefficient can be realised by a "Margulis tube" as the Euclidean structure of its boundary. For $r > 0$ and $\lambda \in \mathbb{C}$, we consider the loxodromic transformation $z \mapsto e^{\lambda}z$ for $z \in \hat{\mathbb{C}}$. The quotient of the r-neighbourhood of the axis, which connects $0 \in \mathbb{C}$ and ∞ in $\mathbb{H}^3$, by the $\mathbb{Z}$-action generated by the loxodromic transformation is denoted by $\mathrm{V}(\lambda, r)$, and is called the Margulis tube with coefficients (λ, r). Here, abusing the terminology, we always call hyperbolic equidistance tubular neighbourhoods of simple closed geodesics Margulis tubes even when the lengths of the core curves are not less than the Margulis constant. We note that the shortest longitudes give a foliation on $\partial V(\lambda, r)$, which we call the *longitude foliation*.

Lemma 9.4.5 (Lemma 3.2 in [35]) *For any $\omega \in \mathbb{C}$ with $\mathrm{Im}(\omega) > 0$, there is a Margulis tube $\mathrm{V}(\lambda, r)$ as above, determined uniquely up to isometry, whose boundary has Euclidean metric isometric to ∂V with the given ω in such a way that the meridian and the longitude of $\mathrm{V}(\lambda, r)$ correspond to those of V.*

The following lemma proved in [12] shows that the meridian coefficients give quasi-isometric control of Margulis tubes.

Lemma 9.4.6 (Lemma 8.5 in [12]) *For any constant $K \geq 1$, $R > 0$ and $L > 0$, there exists a constant K' as follows. Suppose that $r, r' \geq R$ and $|\lambda|, |\lambda'| \leq L$. Let*

$\upsilon : \partial \mathrm{V}(\lambda, r) \to \partial \mathrm{V}(\lambda', r')$ be a K-bi-Lipschitz homeomorphism taking a meridian to a meridian and the longitude foliation to the longitude foliation. Then υ can be extended to a K'-bi-Lipschitz homeomorphism from $\mathrm{V}(\lambda, r)$ to $\mathrm{V}(\lambda', r')$.

For any integer $k > 0$, consider the union $\mathcal{V}[k]$ of components V of $\mathcal{V}[0]$ with $|\omega_M(V)| \geq k$ and set

$$M[k] = M[0] \cup (\mathcal{V}[0] \setminus \mathcal{V}[k]).$$

By definition, we have $M^0 = M[k] \cup \mathcal{V}[k]$. We put on each component V of $\mathcal{V}[0]$ a hyperbolic metric induced from the Margulis tube whose boundary has exactly the Euclidean metric induced from the model metric on $M[0]$, which is guaranteed to exist by Lemma 9.4.5. In this way, we extend the model metric on $M[0]$ to a metric on M^0 whose restriction on $\mathcal{V}[0]$ is hyperbolic. The brick manifold M has a metric induced from that on M^0 via the homeomorphism η_M. We also call these metrics on M^0 and M the model metrics.

9.5 The Bi-Lipschitz Model Theorem for Brick Manifolds

Minsky constructed in [35] model manifolds for hyperbolic 3-manifolds homeomorphic to $S \times (0, 1)$ and proved that for any such hyperbolic manifold, there is a Lipschitz map, called a model map, from its model manifold, whose Lipschitz constant is uniformly bounded. Furthermore, in Brock–Canary–Minsky [12], it was shown that such a model map can be taken to be a bi-Lipschitz homeomorphism, still with uniformly bounded bi-Lipschitz constant. Using and generalising these results, we shall show that a homeomorphism from a labelled brick manifold satisfying conditions A-(1)–(5) and (EL) to a hyperbolic 3-manifold preserving end invariants can be homotoped to a bi-Lipschitz homeomorphism with uniformly bounded bi-Lipschitz constant. Let us recall that for any hyperbolic 3-manifold N and a constant $\varepsilon_1 > 0$ less than the Margulis constant, $N_0 = N_0^{\varepsilon_1}$ denotes the ε_1-non-cuspidal part, i.e. the union of $N_{[\varepsilon_1, \infty)}$ and all Margulis tube components of $N_{(0,\varepsilon_1]}$ as defined in Sect. 9.3.2.

Theorem 9.5.1 (Bi-Lipschitz Model Theorem) *Let M be a labelled brick manifold satisfying conditions A-(1)–(5) and (EL), and N a hyperbolic 3-manifold with a homeomorphism $f : M \to N_0$ preserving the end invariants. Then f is properly homotopic to a homeomorphism $g : M \to N_0 = N_0^{\varepsilon_1}$ satisfying the following conditions, where $k \in \mathbb{N}$, $K \geq 1$ and ε_1 less than the Margulis constant depend only on $\xi(S)$.*

(i) *The image $g(\mathcal{V}[k]) = \mathrm{T}[k]$ is a union of ε_1-Margulis tubes of N_0.*
(ii) $g(M[k]) = N_0 \setminus \mathrm{IntT}[k]$.
(iii) *The restriction $g|_{M[k]} : M[k] \to N_0 \setminus \mathrm{IntT}[k]$ is a K-bi-Lipschitz map.*
(iv) *The homeomorphism g extends continuously to a conformal map from $\partial_\infty M$ to $\partial_\infty N$.*

9.5.1 Minsky's Arguments

Since we need to use Minsky's results and arguments contained in [35] to prove Theorem 9.5.1, we shall review them in this subsection. Let $\mathbf{N}$ be a complete hyperbolic 3-manifold homotopy equivalent to S. The end invariants of $\mathbf{N}$ correspond to two geodesic laminations ν_-, ν_+ on S, obtained by converting each conformal structure at infinity defined on an essential subsurface of S to a shortest pants decomposition. There is a hierarchy $H_{\mathbf{N}}$ of (possibly infinite) tight geodesics connecting ν_- with ν_+ in $\mathcal{C}(S)$. Minsky constructed in [35] a model manifold $\mathbf{M}$ of $\mathbf{N}$ homotopy equivalent to S, and a Lipschitz model map $f_{\mathbf{N}} : \mathbf{M} \to \mathbf{N}$ based on $H_{\mathbf{N}}$. His model manifold $\mathbf{M}$ consists of "internal blocks", "boundary blocks", "exterior blocks", and Margulis tubes. The union of the former three types, forgetting Margulis tubes, is denoted by $\mathbf{M}[0]$.

For each Margulis tube V, in the same way as in Sect. 9.4.6, the meridian coefficient $\omega_V(\mathbf{M})$ is defined. We use the symbols $\mathcal{V}[k]$ and $\mathbf{M}[k]$ in the same way as in Sect. 9.4.6.

Exterior models were introduced to give Lipschitz models of the components of the complement of the "augmented convex core". Here an augmented convex core is the union of the 1-neighbourhood of the convex core and the ε_1-thin part, where ε_1 is a universal constant less than the three-dimensional Margulis constant which we introduced at the beginning Sect. 9.5. Each component U of the complement of the augmented convex core of $\mathbf{M}$ is homeomorphic to $\Sigma \times (0, \infty)$ or $\Sigma \times (-\infty, 0)$ preserving the orientation for an essential subsurface Σ of S, where $\Sigma \times \{0\}$ corresponds to a boundary component of the augmented convex core. The end corresponding to $\Sigma \times \{\infty\}$ or $\Sigma \times \{-\infty\}$ has a conformal structure at infinity, denoted by $\sigma(U)$. The corresponding exterior block E_U is homeomorphic to $\Sigma \times [0, \infty)$ or $\Sigma \times (-\infty, 0]$ preserving the orientation. The conformal structure at infinity $\sigma(U)$ is assigned to the end of E_U. The exterior block E_U has a Riemannian metric which we defined in Eq. (9.4.1) setting $\tau(B)$ in Sect. 9.4.5 to be $\sigma(U)$ when it is identified with $\Sigma \times [0, \infty)$. When E_U is identified with $\Sigma \times (-\infty, 0]$, we replace t in Eq. (9.4.1) with $-t$. The following was proved as Lemma 3.4 in [35].

Lemma 9.5.2 (Minsky [35] Lemma 3.4) *For each component U of the complement of the augmented convex core of $\mathbf{N}$, there is uniform bi-Lipschitz homeomorphism from E_U to the closure of U which induces the identity on Σ.*

Apart from the exterior blocks, $\mathbf{M}[0]$ consists of internal blocks and boundary blocks. An internal block B has the form of $\Sigma \times [-1, 1]$, with $A(\alpha) \times [-1, -1/2)$ and $A(\beta) \times (1/2, 1]$ deleted for an essential subsurface Σ of S with $\xi(\Sigma) = 4$, where α, β are essential simple closed curves on Σ intersecting at fewest possible points (at one point when $\Sigma = \Sigma_{0,4}$, and at two points when $\Sigma = \Sigma_{1,1}$). We fix a hyperbolic metric $\sigma_{\Sigma_{(0,3)}}$ on $\Sigma_{(0,3)}$ with respect to which the length of every boundary component is ε_1. We put a metric on B so that $B \cap \Sigma \times \{t\}$ is isometric to $(\Sigma_{(0,3)}, \sigma_{\Sigma_{(0,3)}})$ for $t \in [-1, -1/2) \cup (1/2, 1]$ via the canonical identification. For $t \in [-1/2, 1/2]$ we put on $\Sigma \times \{t\}$ a metric which is a union of one or two copies of $(\Sigma_{(0,3)}, \sigma_{\Sigma_{(0,3)}})$ and a flat annulus with circumference ε_1. A boundary block B' has the form of $\Sigma \times [-1, 0]$ for an essential subsurface Σ of S with $\xi(\Sigma) \geq 4$, from which $A(P_\Sigma) \times [-1, -1/2)$ is deleted, where P_Σ is a multicurve decomposing Σ into pairs of pants, and $A(P_\Sigma)$ denotes its regular neighbourhood. For every $t \in [-1, -1/2) \cup (1/2, 1]$, the surface $B' \cap \Sigma \times \{t\}$ is a union of pairs of pants each of which has the metric $\sigma_{\Sigma_{(0,3)}}$ independent of t. For $t = -1/2$, the surface $\Sigma \times \{t\}$ is a union of pairs of pants, which has the same metric, and $A(P_\Sigma)$ on which a flat metric with circumference ε_1 is put. The top boundary $\Sigma \times \{0\}$ has a hyperbolic metric within a uniformly bounded distortion from the hyperbolic metric $\sigma(U)$ for the geometrically finite end U facing B', with respect to which every component of P_Σ has length less than ϵ_1. For $t \in (0, 1/2)$, we put a metric on $\Sigma \times \{t\}$ so that outside $A(P_\Sigma) \times (0, 1/2)$, the metric is uniformly bi-Lipschitz to the product of $\sigma_{\Sigma_{(0,3)}}$ and dt. The metric on $\mathbf{M}[0]$ is given by gluing the metrics on the blocks by pasting them by isometries on pairs of pants.

To construct a uniform Lipschitz model map, starting from any continuous map f between $\mathbf{M}$ and $\mathbf{N}$ in a given homotopy class, Minsky gave in §10 of [35] a homotopy to deform f to a Lipschitz map in several steps as listed below. Throughout these steps, we continue to use the symbol f to denote the map obtained in each modification.

(Step 0) *We homotope the restriction of f to each exterior block to a uniformly Lipschitz map.* This can be done by Lemma 9.5.2.

(Step 1) Let P be any gluing surface, a pair of pants along which blocks of $\mathbf{M}[0]$ are glued. Let V_1, V_2 and V_3 be the Margulis tubes on whose boundaries ∂P lies, and γ_1, γ_2 and γ_3 be the closed geodesics which are core curves of $f(V_1)$, $f(V_2)$ and $f(V_3)$ respectively. *We homotope f so that $f|P$ is a restriction of a pleated surface bounded by γ_1, γ_2 and γ_3 for every gluing surface P.* This makes the restriction of f to the union of the exterior blocks and gluing surfaces uniformly Lipschictz.

(Step 2) For each block $b \cong \Sigma \times [-1, 1]$ of $\mathbf{M}[0]$, we consider the middle surface Σ_m corresponding to $\Sigma \times \{0\}$. *We homotope f so that $f|\Sigma_m$ is the "halfway surface" between the two pleated surfaces $f|\Sigma \times \{-1\}$ and $f|\Sigma \times \{1\}$.* This step makes the restriction of f to the union of the exterior blocks, the gluing surfaces and the middle surfaces uniformly Lipschitz.

(Step 3) *We homotope f by straightening the part between the gluing surfaces and the middle surface for each internal block, and the part between the gluing surface and the outer boundary.* This step makes the restriction of f to $\mathbf{M}[0]$ uniformly Lipschitz.

(Step 4) *We homotope f so that its image is contained in the augmented convex core of* $\mathbf{N}$. This can be done preserving the uniform Lipschitzness.

(Step 5) *We show that for any k, there exists $\varepsilon > 0$ such that the image of $f|\mathbf{M}[k]$ is disjoint from the ε-Margulis tubes in* $\mathbf{N}$.

(Step 6) *We homotope f so that its restriction to each Margulis tube in $\mathcal{V}[k]$ is $L(k)$-Lipschitz,* where $L(k)$ is a constant depending only on S and k.

(Step 7) *We homotope f so that there is a constant k_0 depending only on $\xi(S)$ such that for any $k > k_0$, the Lipschitz constant $L(k)$ can be taken to be independent of k.*

9.5.2 Length Bound

The rest of this section is devoted to the proof of Theorem 9.5.1. We should note that by Lemma 9.4.1, there is a proper embedding ι_M of our model manifold into $S \times (0, 1)$. Accordingly, we have an embedding $\iota_N : N_0 \to S \times (0, 1)$ such that $\iota_N \circ f = \iota_M$. As in the previous section, we modify the brick decomposition of M so that Assumption 9.4.3 holds.

By the same argument as Lemma 9.5.2, we can deform f to a map f_1 by a proper homotopy so that for any geometrically finite half-open brick $B' \in \mathcal{K}_{\mathrm{gf}}$, the restriction $f_1|_{B'} : B' \to f_1(B')$ is a uniformly bi-Lipschitz homeomorphism which extends continuously to a conformal map from $\partial_\infty B'$ to $\partial_\infty f_1(B')$ and its real front is mapped into the boundary of the convex core of N.

We shall first show that f_1 can be properly homotoped to a K-Lipschitz map with a constant K depending only on $\xi(S)$. For that, we shall follow the line of Minsky's argument in [35]. Recall that we have a union of tubes $\mathcal{V}$ in M which we constructed in Sect. 9.4.4 inducing a decomposition of M into blocks, and that for each tube V in $\mathcal{V}[0]$, its meridian coefficient $\omega_M(V)$ is defined. The first step is to prove the following lemma.

Lemma 9.5.3 *There is a universal constant L depending only on $\xi(S)$ such that for the core curve v of each tube V in $\mathcal{V}$, the length of the closed geodesic in N homotopic to $f(v)$ is less than L.*

Proof This lemma corresponds to Lemma 7.9 in Minsky [35]. We shall also use its generalisation by Bowditch, Theorem 1.3 in [9].

Recall that we constructed a block decomposition of M repeating the process of putting tight tube unions in bricks, starting from the decomposition of M into bricks of $\mathcal{K}$. First, we note that the geodesic lengths of the images under f of core curves of tubes in $\mathcal{V} \setminus \mathcal{V}[0]$ are bounded by a constant depending only on $\xi(S)$: for each of them corresponds to a curve in the shortest pants decomposition of the hyperbolic structure at infinity on a geometrically finite end whose length is bounded by Bers's lemma, and by Sullivan's theorem [16], the corresponding closed geodesic in N has also length bounded by a constant depending only on $\xi(S)$.

At the first stage, for each connectable internal brick $B = F \times J$, we connected a component $\partial_- B \cap \boldsymbol{l}(\mathcal{K})$ with a component of $\partial_+ B \cap \boldsymbol{l}(\mathcal{K})$ by a tight geodesic. Since f takes $\boldsymbol{l}(\mathcal{K})$ to either an ending lamination or a parabolic element in N or a closed geodesic corresponding to a simple closed curve in a shortest pants decomposition in the conformal structure at infinity, by applying Lemma 7.9 in Minsky [35] or Theorem 1.3 in Bowditch [9] to the covering of N associated to $f_\# \pi_1(B)$, we see that there is a constant L_0 depending only on $\xi(S)$ such that each curve in the simplices constituting the tight geodesic has length in N bounded by L_0.

At the n-th stage, we have bricks $\mathcal{K}_{\text{int}}^{(n)}$ constituting $M_{\text{int}}^{(n)}$ which is the complement of $\mathcal{V}_n = \bigcup_{m=1}^{n} \mathcal{V}^{(m)}$ in M_{int}. Let $\boldsymbol{l}_n$ be the union of $\boldsymbol{l}(\mathcal{K})$ and the core curves of $\mathcal{V}_n$ that are not homotopic to simple closed curves in $\boldsymbol{l}(\mathcal{K})$. In each brick $B^{(n)}$ of $\mathcal{K}_{\text{int}}^{(n)}$, we constructed a tight tube union connecting $\partial_- B^{(n)} \cap \boldsymbol{l}_n$ and $\partial_+ B \cap \boldsymbol{l}_n$. Therefore using Bowditch's Theorem 1.3 inductively, we see that if the geodesic lengths in N of curves in $\boldsymbol{l}_n$ are bounded by L_n, then there is L_{n+1} depending only on L_n bounding the lengths in N of $\boldsymbol{l}_{n+1}$. Since we reached the block decomposition within $\xi(S) - 3$ steps, we see that there is a constant L depending only on $\xi(S)$ which bounds the lengths of the closed geodesics corresponding to the core curves of $\mathcal{V}$.

□

9.5.3 Homotoping f to a Lipschitz Map Preserving the Thin Part

Moving $\mathcal{V}$ by an ambient isotopy of M_{int} without changing the structure of block decomposition, we may assume that for any $B \in \mathcal{K}_{\text{int}}$, every component of $\partial_+ B \setminus \mathcal{V}$ and $\partial_- B \setminus \mathcal{V}$ is homeomorphic to a thrice-punctured sphere. Let F be a compact essential subsurface of S such that B is homeomorphic to $F \times J$ for an interval J. If $\partial_+ B$ is a real front, then $\partial_+ B \cap \mathcal{V}$ determines a simplex in $\mathcal{C}(F)$ inducing a pants decomposition of F. We now homotope f_1 so that each core curve of $\mathcal{V}[0]$ is mapped to a closed geodesic. By Lemma 9.5.3, all such closed geodesics have length bounded by L. In this situation, we can apply Minsky's construction which we explained in Sect. 9.5.1 to get a map $f_2 : M \to N$ for which the following hold. Recall that we have fixed a constant ε_1 less than the three-dimensional Margulis constant.

(1) We have $f_2|B' = f_1|B'$ for every $B' \in \mathcal{K}_{\text{gf}}$. This corresponds to Lemma 9.5.2.
(2) For each block B of $M[0]_{\text{int}}$, the $f_2|\partial_\pm B$ lies on a pleated surface with totally geodesic boundary each of whose components is a closed geodesic homotopic to $f_2(v)$ for a core curve v of some $V \in \mathcal{V}$. This corresponds to Step 1 in Sect. 9.5.1.
(3) There exists a constant $\varepsilon_0 > 0$ depending only on $\xi(S)$ such that for a core curve v of a solid torus component V of $\mathcal{V}$, if the geodesic length of $f_2(v)$ is less than ε_0, then $f_2(V)$ is contained in the ε_1-Margulis tube with core curve $f_2(v)$. This corresponds to Step 3.
(4) The image of f_2 is contained in the union of the 1-neighbourhood of the convex core of N and the ε_1-Margulis tubes of N. This corresponds to Step 4.

To modify f_2 further to get a Lipschitz map, we need the following lemma, which corresponds to Step 7 in Sect. 9.5.1.

Lemma 9.5.4 *Let V be a tube in $\mathcal{V}[0]$, and v its core curve. For any $\delta > 0$, there exists k which depends on δ and $\xi(S)$ but is independent of M and N such that if $|\omega_M(V)| > k$ then the closed geodesic homotopic to $f(v)$ has length smaller than δ.*

Proof This lemma corresponds to Lemma 10.1 in Minsky [35]. In our situation, V may be shared by blocks contained in distinct bricks. Therefore, we cannot apply Minsky's result directly. Instead, we use an argument which can also be found in Soma [46]. Our argument is by contradiction. Suppose that there exist $\delta > 0$ and tubes V_j with core curves v_j such that $|\omega_M(V_j)| \to \infty$ whereas the closed geodesics homotopic to $f_2(v_j)$ have length greater than δ.

Since $|\omega_M(V_j)| \to \infty$, by passing to a subsequence, we can assume that either $\Im\omega_M(V_j) \to \infty$ or $\Re\omega_M(V_j) \to \infty$ holds. We shall first consider the case when $\Im\omega_M(V_j) \to \infty$. By the definition of $\omega_M(V_j)$, there are $(\Im\omega_M(V_j) - 2)$ blocks which intersect ∂V_j along their vertical sides. This implies that there are at least $\Im\omega_M(V_j)$ gluing surfaces, which are homeomorphic to $\Sigma_{(0,3)}$, having boundary components lying on ∂V_j. We should also note that no two distinct gluing surfaces are homotopic in M. Since we assumed that $\Im\omega_M(V_j)$ goes to ∞, there are k_j pairwise non-homotopic gluing surfaces with boundary components on ∂V_j with $k_j \to \infty$. The image of each gluing surface lies on a pleated surface with totally geodesic boundary one of whose components is the closed geodesic γ_j homotopic to $f(v_j)$. Therefore, there are k_j pairwise non-homotopic pleated surfaces from $\Sigma_{(0,3)}$ which have γ_j as a boundary component.

Now, we put a basepoint x_j on γ_j, and consider the geometric limit (N_∞, x_∞) of (N, x_j), passing to a subsequence if necessary. Since the length of γ_j is bounded from above by Lemma 9.5.3 and from below by $\delta > 0$ by our assumption, the geometric limit exists (as a hyperbolic 3-manifold) if we take a subsequence. The geometric limit does not depend on the choice of x_j as long as it lies on γ_j once we fix some geometrically convergent subsequence. Let $\rho_i : B_{R_i}(N, x_j) \to B_{K_i R_i}(N_\infty, x_\infty)$ be an approximate isometry associated to the geometric convergence with $R_i \to \infty$ and $K_i \to 1$. In the geometric limit N_∞, we have the limit γ_∞ of γ_j, which is a closed geodesic since the lengths of the γ_j are bounded away from 0. The geometric limit of pleated surfaces with boundary components on γ_j are pleated surfaces with a boundary component on γ_∞. We should also note that if we fix a positive constant ϵ smaller than δ and ε_1, then all the pleated surfaces intersect the ϵ-thin part of N only near their boundary components other than γ_j, and hence that the limit pleated surfaces can intersect the ϵ-thin part only near their boundary components other than γ_∞. Since $k_j \to \infty$, we can find among the limit pleated surfaces, two limit pleated surfaces F_1, F_2 such that F_2 is homotopic to F_1 in a small regular neighbourhood F_1 whereas $\rho_i^{-1}(F_1)$ and $\rho_i^{-1}(F_2)$ are not homotopic. This is a contradiction.

It remains to deal with the case when $\Re\omega_M(V_j) \to \infty$ with $\Im\omega_M(V_j)$ bounded. Fix a horizontal simple closed curve c_i on ∂V_j. We let d_i be a simple closed curve on ∂V_j intersecting c_i at one point and having shortest length among all simple closed curves intersecting c_i at one point. Let m_j be a meridian of V_j. Since each d_j intersects c_j at one point, as elements of the first homology group of ∂V_j, we have

$[d_j] = [m_j] + \alpha_j[c_j]$ with $\alpha_j \in \mathbb{Z}$ if we fix orientations on c_j, m_j and d_j. Since we assumed that $\Re\omega_M(V_j) \to \infty$, we have $|\alpha_j| \to \infty$, and in particular, we can assume that $\alpha_j \neq 0$ by taking a subsequence. Since the length of d_j is shortest among the simple closed curves intersecting c_j at one point, we have $\text{length}_{\partial V_j}(d_j) \leq (\Im\omega_M(V_j) + 1)\epsilon_1$. Now, since ∂V_j is contained in $M[0]$, by condition (6) above, we have $\text{length}(f_2(d_j)) \leq L(0)(\Im\omega_M(V_j) + 1)\epsilon_1$. The right hand side is bounded above since we have already proved that $\Im\omega_M(V_j)$ is bounded as $j \to \infty$. Since $[d_j] = [m_j] + \alpha_j[c_j]$, the curve $f_2(d_j)$ with an appropriate orientation is homotopic to the $|\alpha_j|$-time iteration of γ_j in N. This implies $\text{length} f_2(d_j) \geq |\alpha_j| \text{length}(\gamma_j)$. The right hand side goes to ∞, whereas the left hand side is bounded as we have seen above. This is a contradiction. □

Having proved Lemma 9.5.4, the rest of the modification to get *a proper, degree-1 map* $f_3 : M \to N_0$ such that $f_3|M[k]$ is K_3-Lipschitz with K_3 depending only on $\xi(S)$ consists of the following two steps.

(5) For any k, there exists a positive number $\epsilon(k) < \varepsilon_1$ such that $f_2(M[k])$ is disjoint from the ε_1-Margulis tubes of N whose core curves have length less than $\epsilon(k)$. This corresponds to Step 5 of Sect. 9.5.1.
(6) For any k, there exists a constant $L(k)$ such that $f_2|M[k]$ is $L(k)$-Lipschitz. This corresponds to Step 6.

We state one more property of f_3 which is derived from our construction.

(7) Since f_3 has degree 1, there exist constants k_2 and $\epsilon(k_2)$ as in condition (5) depending only on $\xi(S)$ such that any ε_1-Margulis tube in N whose core curve has length less than $\epsilon(k_2)$ is contained in the image of a component of $\mathcal{V}[k_2]$.

9.5.4 Preliminary Steps to Homotope f_3 to a Bi-Lipschitz Map

We now turn to modify f_3 to a bi-Lipschitz homeomorphism. This was done in Brock–Canary–Minsky [12] for the case of surface Kleinian groups. Recall that we moved $\mathcal{V}$ so that for each brick B in $\mathcal{K}_{\text{int}}$, if its upper or lower front $\partial_\pm B$ is real, then every component of $\partial_\pm B \setminus \mathcal{V}$ is a thrice-punctured sphere. We parametrise B as $F \times J$ with a closed or half-open interval J. We define $\boldsymbol{i}(B)$ to be a simplex in $\mathcal{C}(F)$ with empty transversals such that $\boldsymbol{i}(B) \times \{\min J\}$ is homotopic to the union of core curves of $\partial_- B \cap \mathcal{V}$ if $\partial_- B$ is real, and to be the ending lamination of the end corresponding to $F \times \{\inf J\}$ if $\partial_- B$ is ideal. Similarly we define $\boldsymbol{t}(B)$ for the upper boundary of B. We shall first show that in this setting, the block decomposition of B induced by $\mathcal{V}$ corresponds to a hierarchy in the sense of Masur–Minsky [30].

Lemma 9.5.5 *Let B be a brick in $\mathcal{K}_{\text{int}}$, homeomorphic to $F \times J$ with a closed or half-open interval J. Then there is a* 4*-complete hierarchy h of tight geodesics on F with $I(h) = \boldsymbol{i}(B)$ and $T(h) = \boldsymbol{t}(B)$ whose* 4*-sub-hierarchy gives rise to the same block decomposition of B as the one induced by $\mathcal{V}$ converted as in Remark 9.4.4 to Minsky's decomposition.*

Proof In the construction of $\mathcal{V}$ in the previous section, we began with putting tight tube unions in all connectable bricks in M_{int} whose initial and terminal vertices are in $\boldsymbol{l}(\mathcal{K})$. After that, we merged homotopic tubes into one and let the obtained tube union be $\mathcal{V}^{(1)}$. Then we considered the brick manifold $M^{(1)}$ which is the complement of $\mathcal{V}^{(1)}$ and repeated the same procedure until we got a block decomposition. Now, we shall look more closely how tubes are put (and merged) in B during this construction and define tight geodesics which constitute h. We define $I(B) = \boldsymbol{i}(B) \times \inf J$ and $T(B) = \boldsymbol{t}(B) \times \sup J$. (These may be larger than $\boldsymbol{I}(B)$ and $\boldsymbol{T}(B)$ defined in the previous section.)

If B is connectable in the first step of the construction of $\mathcal{V}$, then we get a tube union $\mathcal{V}_B$ on B in the first step, which corresponds to a tight geodesic g_B in $\mathcal{C}(F)$ connecting a component of $\boldsymbol{l}(\mathcal{K}) \cap \partial_- B$ with a component of $\boldsymbol{l}(\mathcal{K}) \cap \partial_+ B$. (If one of them is an ending lamination, the geodesic g_B refers to a tight geodesic ray tending to it.) Since $\boldsymbol{l}(\mathcal{K}) \cap \partial_- B \subset I(B)$ and $\boldsymbol{l}(\mathcal{K}) \cap \partial_+ B \subset T(B)$, we can assume that g_B has $\boldsymbol{i}(B)$ as initial marking, and $\boldsymbol{t}(B)$ as terminal marking. We let g_B the main geodesic of h.

We next consider how the merging of tubes is reflected in the construction of geodesics in the hierarchy still under the assumption that B is connectable at the first step. If there is a tube V in B which is merged with another homotopic tube V' in another brick B', then a core curve v of V must be in either $\boldsymbol{i}(B)$ or $\boldsymbol{t}(B)$ since $\partial_\pm B \setminus \mathcal{V}$ consists of thrice-punctured spheres. This can occur only when the core curve is contained in either the first, or the second, or the second but last, or the last simplex of the geodesic g_B, for its core curve regarded as a curve on $\partial_\pm B$ cannot have non-zero intersection number with $\mathbf{I}(B)$ or $\mathbf{T}(B)$. If v is contained in the first or the last vertex of g_B, this procedure of merging does not affect tubes in B. Otherwise, v is contained in either the second or the second to last simplex of g_B. In this case, we regard the procedure as corresponding to putting a geodesic consisting of only one vertex, i.e. of length 0, which is subordinate to g_B at the first or the last vertex.

Next, we shall consider the case when B is not connectable in the first step. In the second step, either (a) B is contained in another brick $\bar{B}$ constituting $M^{(1)}$ or (b) B is split into two (or more) in the process of merging two homotopic tubes of $\mathcal{V}^{(1)}$, one lying above B and the other below B. In the latter case, let $V_1, \ldots, V_p$ be tubes in $\mathcal{V}^{(1)}$ which split B. We should note that these tubes have core curves which are homotopic to curves both in $I(B)$ and $T(B)$. Let $v_1, \ldots, v_p$ be the curves on F corresponding to their core curves. Then we define geodesics $g_1, \ldots, g_p$ each of which consists of only one vertex, such that $D(g_1) = F$, $D(g_j)$ is a component of $F \setminus \cup_{s=1}^{j-1} v_s$ for $j = 2, \ldots p$, $I(g_j) = \boldsymbol{i}(B) \cap D(g_j)$, $T(g_j) = \boldsymbol{t}(B) \cap D(g_j)$, and $g_{j-1} \overset{d}{\swarrow} g_j \overset{d}{\searrow} g_{j-1}$, and let them be geodesics contained in h setting the main geodesic g_B to be g_1. In case (a), if $\bar{B}$ is connectable in $M^{(1)}$, then we consider $\mathcal{V}_{\bar{B}} \cap B$, where as explained above B is assumed to be in a position such that $\mathcal{V} \cap \partial_- B$ is a regular neighbourhood of $I(B)$ and $\mathcal{V} \cap \partial_+ B$ is that of $T(B)$, and define the main geodesic g_B to be the tight geodesic in $\mathcal{C}(F)$ corresponding to $\mathcal{V}_{\bar{B}} \cap B$. As before, we define $I(g_B) = \boldsymbol{i}(B)$ and $T(g_B) = \boldsymbol{t}(B)$. If $\bar{B}$ is not connectable, we proceed to the following step and repeat the same procedure depending on whether there is a brick containing $\bar{B}$ or $\bar{B}$ is split by merging of homotopic tubes. Thus we have defined the main geodesic g_B, together with some more geodesics in h in the case when B is split. We shall now turn to the subsequent steps.

In the subsequent steps, we put a tight tube union $\mathcal{V}_{B'}$ into a brick B' constituting a brick decomposition of $M_{\text{int}} \setminus \mathcal{V}^{(k)}$. We shall show that the intersection with B of each tube union in a connectable brick B' in the $(k+1)$-th step gives rise to a tight geodesic on F which is subordinate to the ones obtained up to the k-th step. This implies that in the final step, we shall get a hierarchy on F connecting $\boldsymbol{i}(B)$ and $\boldsymbol{t}(B)$. To show that, we shall analyse what a tube union in B' brings about to B, dividing the argument into subcases depending on the location of B' with regard to B. (Again, B is in a position where $\partial_- B \cap \mathcal{V}$ is a regular neighbourhood of $I(B)$ and $\partial_+ B \cap \mathcal{V}$ is that of $T(B)$.) We parametrise B' as $F' \times J'$ with $F' \subset F$, in such a way that horizontal leaves and vertical leaves are contained in those of bricks in $\mathcal{K}_{\text{int}}$. Since $F' \times \{x\}$ for $x \in \text{Int}J'$ is a horizontal leaf whose boundary lies on $\partial \mathcal{V}_k$, the surface F' is a component domain of a geodesic corresponding to a tube union which was already put into M up to the k-th step. Now we divide our argument into three, depending on an inclusion relation between J and J'.

First, suppose that B' is contained in B, which means that both $\partial_- B'$ and $\partial_+ B'$ lie in B and J' is contained in J. Recall that core curves of $\partial_\pm B' \cap \mathcal{V}_k$ are $\partial_\pm B' \cap \boldsymbol{l}_k$. We define $I(B') = \partial_- B' \cap (\boldsymbol{l}_k \cup (\boldsymbol{i}(B) \times \inf J))$ and $T(B') = \partial_+ B' \cap (\boldsymbol{l}_k \cup (\boldsymbol{t}(B) \times \sup J))$. In this definition, we need to add $\boldsymbol{i}(B) \times \inf J$ and $\boldsymbol{t}(B) \times \sup J$ to deal with the case when $\inf B' = \inf B$ or $\sup B' = \sup B$. Note that $\boldsymbol{I}(B')$ is $\boldsymbol{l}_k \cap \partial_- B'$ and $\boldsymbol{T}(B')$ is $\boldsymbol{l}_k \cap \partial_+ B'$, which are contained in $I(B')$ and $T(B')$ respectively. By our construction of $\mathcal{V}^{(k+1)}$, the tube union $\mathcal{V}_{B'}$ in B' connects a component of $\boldsymbol{I}(B')$ to that of $\boldsymbol{T}(B')$. We define $g_{B'}$ to be the tight geodesic corresponding to $\mathcal{V}_{B'}$ whose initial and terminal markings are simplices corresponding to $I(B')$ and $T(B')$ respectively. Since we defined $I(B') = \partial_- B' \cap (\boldsymbol{l}_k \cup (\boldsymbol{i}(B) \times \inf J))$ and $T(B') = \partial_+ B' \cap (\boldsymbol{l}_k \cup (\boldsymbol{t}(B) \times \sup J))$, and F' is a component domain of some simplex of a tight geodesic of h already defined as was explained above, the tight geodesic $g_{B'}$ is both forward and backward subordinate to a geodesic in h which was obtained up to the k-th step.

Next suppose that, one of $\partial_- B'$ and $\partial_+ B'$ is contained in B whereas the other is not. This means that one of the endpoints of J' lies in J whereas the other does not. Now, we assume that $\partial_- B'$ is the one contained in B: for the other case, we can argue in the same way, just interchanging the directions. In this situation, $\boldsymbol{I}(B')$ coincides with $\partial_- B \cap \boldsymbol{l}_k$ which is contained in $\boldsymbol{l}_k \cap B$. On the other hand, $\boldsymbol{T}(B')$ may not lie in B. Now, by our definition of $\mathcal{V}$, the tube union $\mathcal{V}_{B'}$ put into B' is contained in $\mathcal{V}$. Therefore, $\mathcal{V}_{B'}$ intersects $\partial_+ B$ by components of $\partial_+ B \cap \mathcal{V}$ since we moved $\mathcal{V}$ so that every component of $\partial_+ B \setminus \mathcal{V}$ is a thrice-punctured sphere. (Recall that unless $\xi(B') = 4$, the upper front of each tube of $\mathcal{V}_{B'}$ lies on the same horizontal level as the lower front of the subsequent tube. In the case when $\xi(B') = 4$, there is a gap between them, but we moved $\mathcal{V}$ so that $\partial_\pm B$ avoid such gaps.) Therefore, if we consider a sub-tube union $\bar{\mathcal{V}}_{B'}$ of $\mathcal{V}_{B'}$ starting from the first tube and ending at a tube in $\partial_+ B \cap \mathcal{V}$, then it is exactly what $\mathcal{V}_{B'}$ brings about to B. If $\mathcal{V}_{B'} \cap \partial_+ B$ consists of only one component, then we let $g_{\bar{B}'}$ be the tight geodesic corresponding to $\bar{\mathcal{V}}_{B'}$ defining $I(g_{\bar{B}'})$ to be a simplex consisting of curves corresponding to $\partial_- B' \cap (\boldsymbol{l}_k \cup (\boldsymbol{i}(B) \times \inf J))$ and $T(g_{\bar{B}'})$ to be $\boldsymbol{t}(B) \cap F'$. Otherwise, we choose one component of $\mathcal{V}_{B'}$, denoted by $V^0_{B'}$ and remove the others, denoted by $V^1_{B'}, \dots, V^u_{B'}$, from $\bar{\mathcal{V}}_{B'}$, and then define $g_{\bar{B}'}$ in the same way. Since the last tube of $\bar{\mathcal{V}}_{B'}$ intersects $\partial_+ B$, it has a core curve contained in $T(B)$. This implies that $g_{\bar{B}'}$ is forward subordinate to one of the geodesics obtained up to the k-th step. We also know that $g_{\bar{B}'}$ is also backward subordinate to such a geodesic by the argument in the previous case.

In the latter case when $\mathcal{V}_{B'} \cap \partial_+ B$ is not connected, we further define tight geodesics $g^1_{B'}, \dots, g^u_{B'}$ inductively as follows. Let $v^j_{B'}$ be a core curve of $V^j_{B'}$ for $j = 0, \dots, u$. Let D be a component of $F \setminus v^0_{B'}$ containing $v^1_{B'}$, and let $v^{-1}_{B'}$ be the simplex of $g_{\bar{B}'}$ which precedes $v^0_{B'}$. Then we define $g^1_{B'}$ to be a tight geodesic of length 1 supported on D with $I(g^1_{B'})$ equal to $v^{-1}_{B'} \cap D$ and $T(g^1_{B'})$ equal to $\boldsymbol{t}(B) \cap D$. (The intersection $v^{-1}_{B'} \cap D$ is not empty since $g_{\bar{B}'}$ is tight.) In the same way, we define D^k to be the component of $F \setminus (v^0_{B'} \cup \cdots \cup v^{k-1}_{B'})$ containing $v^k_{B'}$ and $g^k_{B'}$ to be a tight geodesic of length 1 supported on D^k with $I(g^k_{B'})$ equal to $v_{B'} \cap D^k$ and $T(g^k_{b'})$ equal to $\boldsymbol{t}(B) \cap D^k$. Then all these geodesics $g^0_{B'}, \dots, g^u_{B'}$ are subordinate to $g_{\bar{B}'}$ in both directions. Thus in either case, we get tight geodesics which are both backward and forward subordinate to geodesics in h obtained up to the previous step.

Finally suppose that neither $\partial_- B'$ nor $\partial_+ B'$ is contained in B. Then $\mathcal{V}_{B'}$ intersects $\partial_- B$ by a union of solid tori V_1 contained in $\partial_- B \cap \mathcal{V}$ and $\partial_+ B$ by a union of solid tori V_2 contained in $\partial_+ B \cap \mathcal{V}$. We define $\bar{\mathcal{V}}_{B'}$ to be the sub-tube union of $\mathcal{V}_{B'}$ starting from a component of V_1 and ending at a component of V_2. Then $\bar{\mathcal{V}}_{B'}$ is the union of tubes which $\mathcal{V}_{B'}$ brings about to B. We define $g_{\bar{B}'}$ to be the corresponding tight geodesic supported on F', setting its initial and terminal markings to be $\boldsymbol{i}(B) \cap F'$ and $\boldsymbol{t}(B) \cap F'$ respectively. In the same way as in the previous paragraph, we define geodesics of length 1 corresponding to the components of V_1 and V_2 which are not chosen. These geodesics are both forward and backward subordinate to geodesics which are obtained up to the k-th step by the same reason as in the previous case.

Now, recall that in each step, we also merge homotopic tubes into one. As was analysed in the first step, this procedure corresponds to putting a geodesic consisting of only one vertex which is subordinate to a geodesic which was already constructed in the previous steps. Thus we have shown that at each step we get a geodesic subordinate to those which we have already had and at the final step, we get non-annular geodesics in h.

We shall next define annular geodesics of h. Let V be a tube in $\mathcal{V}$ intersecting B along its entire boundary. We parametrise V as $A \times [0, 1]$ preserving leaves. Since $A \times \{0\}$ lies on $\partial_+ b$ for some block b of the form $\Sigma_{0,4} \times J$ or $\Sigma_{1,1} \times J$, the core curve of the annulus on $\partial_- b$ which is the complement of the other blocks intersecting $\partial_- b$ defines an arc on $A \times \{0\}$ which is regarded as a vertex v_- in $\mathcal{C}(A)$. Similarly we can define an arc on $A \times \{1\}$ regarded as a vertex v_+ in $\mathcal{C}(A)$ from the block on whose bottom $A \times \{1\}$ lies. We define a geodesic g_B supported on A connecting v_- and v_+, and let it be contained in h. By our construction, this geodesic is both forward and backward subordinate to 4-geodesics already contained in h.

It remains to show that the block decomposition of B induced from $\mathcal{V}$ is compatible with h. This means that we have a resolution of the 4-sub-hierarchy of h which gives rise to the block decomposition induced from $\mathcal{V}$. We consider the family of horizontal surfaces $F \times \{t\}$ in B. Then outside countably many intervals corresponding to gaps which we introduced for 4-geodesics, $F \times \{t\} \cap \mathcal{V}$ induces a pants decomposition of F. We should also note that if t is contained in a gap interval, then $F \times \{t\}$ passes a block of the form either $\Sigma_{0,4} \times J$ or $\Sigma_{1,1} \times J$. Passing each interval of gap, the configuration of pants decomposition changes by elementary moves which may take place at finitely many disjoint places at the same time. This must come from stepping forward on 4-geodesics which we defined above. Therefore, this induces a resolution of the 4-sub-hierarchy of h. Since each elementary move also corresponds to a block of the form $\Sigma_{0,4} \times J$ or $\Sigma_{1,1} \times J$ in our decomposition, the block decomposition induced from this resolution is obtained by converting the one induced from $\mathcal{V}$ as in Remark 9.4.4. □

In a hierarchy, a curve can appear at most once. Since our tube union $\mathcal{V}$ itself does not correspond to a hierarchy (Lemma 9.5.5 only says that its restriction to a brick corresponds to a hierarchy.), we need to show the same kind of property for $\mathcal{V}$.

Lemma 9.5.6 *There are no two distinct tubes in $\mathcal{V}$ which are homotopic in M.*

Proof Suppose, seeking a contradiction, that there are tubes V_1, V_2 in $\mathcal{V}$ which are homotopic to each other in M. Let k_1, k_2 be the numbers such that $V_1 \in \mathcal{V}^{(k_1)}$ and $V_2 \in \mathcal{V}^{(k_2)}$, and set $k = \max\{k_1, k_2\}$. (When we say $V_s \in \mathcal{V}^{(k_s)}$ for $s = 1, 2$, we take the smallest k_s such that $\mathcal{V}^{(k_s)}$ contains a tube constituting V_s. We follow the same convention throughout the proof.) Then longitudes of V_1 and V_2 pushed out to their boundaries are not homotopic in $M^{(k)}$, for otherwise they should have been merged into one in our construction. Let $\mathcal{U}$ be the union of tubes in $\mathcal{V}_k = \cup_{j=1}^{k} \mathcal{V}^{(j)}$ which intersect essentially an embedded annulus A bounded by the pushed-out longitudes of V_1 and V_2 in M. (These are determined independently of the choice of an annulus since M is atoroidal.)

Let $U \in \mathcal{U}$ be a tube which appears in the earliest step among the tubes in $\mathcal{U}$, and suppose that $U \in \mathcal{V}^{(l)}$. Note that we have $l \leq k$ by our definition of $\mathcal{U}$. Let $B \cong F \times J$ be a brick in $\mathcal{K}^{(l-1)}$ where U appears as a tube in the tight tube union. If either the interior of V_1 or V_2, say V_1, intersects a front of B, then by replacing $V_1 \cap B$ with V_1, we can assume that both V_1 and V_2 have interiors disjoint from the fronts of B. We also note that by our choice of l, the annulus A can be regarded as lying in $M^{(l-1)}$.

First suppose that both V_1 and V_2 are contained in B. In the following argument, for two tubes $U, V \in \mathcal{V}$, we write $U \approx V$ if $U = A_1 \times J_1$, $V = A_2 \times J_2$ and $\mathrm{Int}J_1 \cap \mathrm{Int}J_2 \neq \emptyset$ for the parametrisation on $S \times (0, 1)$ in which M is embedded by ι_M. In the k-th step, a tight tube union $\mathcal{V}_B$ corresponding to a tight geodesic g_B on $\mathcal{C}(F)$ is given. Then there are tubes U_1, U_2 in the tight tube union of B such that $V_1 \approx U_1$ and $V_2 \approx U_2$. Let u_1, u_2, u, v be vertices of $\mathcal{C}(F)$, which correspond to U_1, U_2, U, V_1 respectively. Since U intersects A, we have $u_1 < u < u_2$ or $u_2 < u < u_1$ with respect to the ordering on the simplices of g_B. Since $V_1 \approx U_1$ and

$V_2 \approx U_2$, we see that $i(u_1, v) = 0$ and $i(u_2, v) = 0$. Then, we have $u_1, u_2 \in \phi_{g_B}(v)$ whereas $u \notin \phi_{g_B}(v)$. This contradicts the fact that $\phi_{g_B}(u)$ consists of contiguous simplices, which is described in Lemma 9.3.2.

Next suppose that one of V_1, V_2, say V_1, lies outside B whereas V_2 is contained in B. In this case, A passes through a joint contained in the upper or the lower front of B. We only consider the case when A passes through a joint in the upper front. The other case can be dealt with in the same way just by turning the figure upside down. Since A passes through a joint in the upper front, the vertical projection of the core curve of $\boldsymbol{T}(B)$ is disjoint from that of the core curve of V_1, which implies that the last vertex u_∞ of g_B is contained in $\phi_{g_B}(v)$. As in the previous paragraph, we have $u_1 < u < u_\infty$ and $u_1, u_\infty \in \phi_{g_B}(v)$ whereas $u \notin \phi_{g_B}(v)$, which contradicts Lemma 9.3.2 as before. Also in the case when both V_1 and V_2 lie outside B, we can argue in the same way considering joints which A passes contained in the upper and the lower fronts. Then we see that the first and the last vertices are contained in $\phi_{g_B}(v)$ whereas u is not. This again contradicts Lemma 9.3.2, which completes the proof. □

The next lemma is obtained from Otal [44] for hyperbolic 3-manifolds homeomorphic to $S \times \mathbb{R}$ for a hyperbolic surface S. Since the only condition that is necessary for the proof is the fact that the manifold can be filled up by incompressible pleated surfaces (with bounded genus), his argument also works in our setting.

Lemma 9.5.7 *There is a constant k_0 depending only on $\chi(S)$ such that for any $k \geq k_0$ and the tubes $V \in \mathcal{V}[k]$, the union $\boldsymbol{c}$ of core curves of the V, taken one on each tube, is mapped by f_3 to unknotted and unlinked closed geodesics, i.e. there is a isotopy of* $S \times (0, 1)$ *which takes $\iota_N(\boldsymbol{c})$ to a disjoint collection of simple closed curves lying on horizontal surfaces.*

Take k_2 in condition (7) so that $\epsilon(k_2)$ is less than our fixed ε_1 (less than the Margulis constant). By Lemma 9.5.4. there exists k_1 such that if $|\omega(V)| \geq k_1$, then $f(v)$ has length less than $\epsilon(k_2)$. We define $k_u = \max\{k_0, k_1, k_2\}$ for k_0 in the above lemma, and let ϵ_u be $\epsilon(k_u)$.

Now, we recall the following definition of topological order introduced in Brock–Canary–Minsky [12], which we shall apply for surfaces in M or N_0.

Definition 9.5.8 (Brock–Canary–Minsky [12]) Let $j_1 : F_1 \to M$ and $j_2 : F_2 \to M$ be maps from essential subsurfaces $F_1, F_2 \subset S$ such that $\iota_M \circ j_i$ is homotopic to the inclusion $F_i \to F_i \times \{t\}$ for $i = 1, 2$. We write $j_1 \prec_{\text{top}} j_2$ if and only if $\iota_M \circ j_1$ can be homotoped to $S \times \{0\}$ without touching $\iota_M \circ j_2(F_2)$ and $\iota_M \circ j_2$ can be homotoped to $S \times \{1\}$ without touching $\iota_M \circ j_1(F_1)$. We call the relation $\prec_{\text{top}}$ the topological order. We define the topological order on maps from surfaces to N_0 in the same way just replacing M with N_0 and ι_M with ι_N.

We should also recall that two embedded surfaces F_1, F_2 in $S \times (0, 1)$ are said to overlap if their projections to S have essential intersection. We use this term also for surfaces in M or N_0, for they can be embedded in $S \times (0, 1)$ by ι_M and ι_N.

9.5.5 Homotoping f_3 to a Homeomorphism

We shall next consider homotoping f_3 so that its restriction to the union of the joints of the bricks is an embedding. Let F be a joint of B with another brick. Recall that F intersects $\mathcal{V}$ in such a way that each component of $F \setminus \mathcal{V}$ is a thrice-punctured sphere. We define $\check{F}[k]$ to be an embedded surface in $M[k]$ obtained from F by isotoping annuli in $F \cap \mathcal{V}[k]$ to those on $\partial\mathcal{V}[k]$. There are two choices for an annulus for each component of $F \cap \mathcal{V}[k]$. We take an annulus on $\partial\mathcal{V}[k]$ situated below the other one with respect to the embedding ι_M.

Recall that the images of $\mathcal{V}[k_u]$ are unknotted and unlinked ε_1-Margulis tubes whose core curves have lengths less than $\epsilon(k_2)$. Conversely, every ε_1-Margulis tube whose core curve has length less than $\epsilon(k_2)$ is the image of a component of $\mathcal{V}[k_2]$ by f_3. Recall that we denote the union of the Margulis tubes which are the images under f_3 of the tubes in $\mathcal{V}[k_u]$ by $T[k_u]$. We denote $N_0 \setminus \mathrm{Int}T[k_u]$ by $N[k_u]$. By Lemma 9.5.6, f_3 induces a bijection between the components of $\mathcal{V}[k_u]$ and those of $T[k_u]$. Moreover, the image of $M[k_2]$ is disjoint from $T[k_u]$ by condition (5). Again by Lemma 9.5.6, no tubes in $\mathcal{V}[k_2] \setminus \mathcal{V}[k_u]$ are mapped to $T[k_u]$. Therefore f_3 induces a Lipschitz map $f_3 : M[k_u] \to N[k_u]$.

Proposition 9.5.9 *The Lipschitz map $f_3 : M[k_u] \to N[k_u]$ can be properly homotoped to a homeomorphism $f_4 : M[k_u] \to N[k_u]$, which extends to a homeomorphism between M and N_0. This map f_4 may not be Lipschitz.*

Proof Let B be a brick of $\mathcal{K}_{\mathrm{int}}$. We denote by $F_1^+, \ldots, F_\mu^+$ its joints contained in the upper front, and by $F_1^-, \ldots, F_\nu^-$ those contained in the lower front. (One of the fronts may be ideal; hence one of these families may be empty.) We consider $\check{F}_1^+[k_u], \ldots, \check{F}_\mu^+[k_u]$ and $\check{F}_1^-[k_u], \ldots, \check{F}_\nu^-[k_u]$ as defined above, and denote their unions by $\check{F}_+$ and $\check{F}_-$. Note that both $\check{F}_+$ and $\check{F}_-$ are incompressible in M. By changing each joint F to $\check{F}$, we get a brick decomposition of M which is isotopic to the original one. From now on until the end of the proof of this proposition, when we refer to a brick B, we mean the one in this new decomposition, which is isotopic to the original B. Let $p_B : M_B \to M$ be the covering associated to the image of $\pi_1(B)$ in $\pi_1(M)$. Similarly, we consider the covering N_B of N_0 associated to $(f_3)_\#\pi_1(B)$. Let $\tilde{f}_3 : M_B \to N_B$ be the lift of f_3 which is uniformly Lipschitz outside the preimages of $\mathcal{V}[k_u]$, and $\tilde{f} : M_B \to N_B$ that of f, which is a homeomorphism. The surfaces $\check{F}_+, \check{F}_-$ lift homeomorphically to surfaces $\tilde{F}_+, \tilde{F}_-$ lying on the boundary of a submanifold $\tilde{B}$ homeomorphic to B under p_B. We use the symbols $\partial_-\tilde{B}$ and $\partial_+\tilde{B}$ to denote the fronts of $\tilde{B}$ corresponding to $\partial_- B$ and $\partial_+ B$ respectively. Let $\tilde{\mathcal{V}}[k_2]$ and $\tilde{T}[k_u]$ be the preimages of $\mathcal{V}[k_2]$ and $T[k_u]$ respectively. We denote by $M_B[k_u]$ the complement of $\mathrm{Int}\tilde{\mathcal{V}}[k_u]$ in M_B, and by $N_B[k_u]$ the complement of $\mathrm{Int}\tilde{T}[k_u]$ in N_B.

Note that $\tilde{f}_3|(\tilde{F}_+ \sqcup \tilde{F}_-)$ is properly homotopic to $\tilde{f}|(\tilde{F}_+ \sqcup \tilde{F}_-)$ which is an embedding. We can assume that $\tilde{f}_3|(\tilde{F}_+ \sqcup \tilde{F}_-)$ is an immersion from the start by perturbing it. Then, by Theorem 9.3.6, we see that $\tilde{f}_3|(\tilde{F}_+ \sqcup \tilde{F}_-)$ can be properly homotoped to an embedding by a homotopy which passes through only relatively compact components of $N_B \setminus \tilde{f}_3(\tilde{F}_+ \sqcup \tilde{F}_-)$. We note that each of such relatively compact components is homeomorphic to a trivial I-bundle whose associated ∂I-bundle can be identified with a compact subsurface of $\tilde{F}_+ \sqcup \tilde{F}_-$.

Suppose that a component W of $N_B \setminus \tilde{f}_3(\tilde{F}_+ \sqcup \tilde{F}_-)$ intersects a component T of $\tilde{T}[k_u]$. This means that W contains T since $f_3(\check{F})$ is disjoint from $T[k_u]$. We shall now prove the following claim.

Claim 9.5.10 *The surfaces $\tilde{f}_3(\tilde{F}_+)$ and $\tilde{f}_3(\tilde{F}_-)$ are homotopic to disjoint embeddings by proper homotopies which do not touch T.*

Proof Because f_3 satisfies conditions (3), (5) and (7), there is a unique component V of $\tilde{\mathcal{V}}[k_u]$, which is a solid torus, such that $\tilde{f}_3(V) = T$. Since $M[k_u]$ is mapped to $N[k_u]$ and $\mathcal{V}[k_u]$ bijects to $T[k_u]$, we see that $\tilde{f}_3(M_B \setminus V) \subset N_B \setminus T$. Since every Kleinian surface group is tame, the interior of N_B is homeomorphic to $\partial_- B \times (0, 1)$, and hence so is M_B. Let $V_1, \ldots, V_p$ be the components of ∂M whose longitudes (in the case of torus boundary) or core curves (in the case of open annulus boundary) are homotopic into $\check{F}_1^+ \cup \cdots \cup \check{F}_\mu^+$ in $M \setminus \mathrm{Int}B$. We renumber them in such a way that $V_1, \ldots, V_r$ are disjoint from B whereas $V_{r+1}, \ldots, V_p$ intersect B along annuli. By the annulus theorem and a standard cut-and-paste technique, we see that there are disjoint embedded annuli $\alpha_1, \ldots, \alpha_r$ realising homotopies between longitudes or core curves on $V_1, \ldots, V_p$ and simple closed curves on $\check{F}_1^+ \cup \cdots \cup \check{F}_p^+$ with $\partial\alpha_j \subset V_j \cup \check{F}_1^+ \cup \cdots \cup \check{F}_p^+$. We can lift $V_1, \ldots, V_p$ and $\alpha_1, \ldots, \alpha_r$ to open annuli $A_1, \ldots, A_p$ on ∂M_B and annuli $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_r$ such that A_j and $\tilde{\alpha}_j$ intersect at a core curve of A_j for $j = 1, \ldots, r$. Similarly, we consider the components $V_1', \ldots, V_q'$ of ∂M whose longitudes or core curves are homotopic into $\check{F}_1^- \cup \cdots \cup \check{F}_\nu^-$ in $M \setminus \mathrm{Int}B$, among which $V_1', \ldots, V_s'$ lie outside B, and take annuli $\alpha_1', \ldots, \alpha_s'$ realising homotopies between longitudes or core curves and simple closed curves on $\check{F}_1^- \cup \cdots \cup \check{F}_\nu^-$. We lift $V_1', \ldots, V_q'$ to open annuli $A_1', \ldots, A_q'$ and $\alpha_1', \ldots, \alpha_s'$ to annuli $\tilde{\alpha}_1', \ldots, \tilde{\alpha}_s'$ in the same way as the previous case.

Let $\bar{A}_1, \ldots, \bar{A}_p$ and $\bar{A}_1', \ldots, \bar{A}_q'$ be core annuli of $A_1, \ldots, A_p$ and $A_1', \ldots, A_q'$ such that $\bar{A}_j$ contains $\tilde{\alpha}_j \cap A_j$ for $j \leq r$ whereas $\bar{A}_j = \tilde{B} \cap A_j$ for $j > r$, and $\bar{A}_j'$ contains $\alpha_j' \cap A_j'$ for $j \leq s$ whereas $\bar{A}_j' = \tilde{B} \cap A_j'$ for $j > s$. By identifying $\partial_-\tilde{B}$ and $\partial_+\tilde{B}$ by vertical translation and $\partial_-\tilde{B}$ with $\partial_- B$ by p_B, we regard $\bar{A}_1, \ldots, \bar{A}_p; \bar{A}_1', \ldots, \bar{A}_q'$ as lying on $\partial_- B$. To construct parts corresponding to the $\mathbb{Z}$-cusps in $\partial_- B \times (0, 1)$, we set $U_+ = (\bar{A}_1 \cup \cdots \cup \bar{A}_r) \times (7/8, 1)$, $U_+' = (\bar{A}_{r+1} \cup \cdots \cup \bar{A}_p) \times (3/4, 1)$, $U_- = (\bar{A}_1' \cup \cdots \cup \bar{A}_s') \times (0, 1/8)$, and $U_-' = (\bar{A}_{s+1}' \cup \cdots \cup \bar{A}_q') \times (0, 1/4)$ and denote the union of these four parts by U. We parametrise M_B by a proper homeomorphism $I_M : M_B \to \partial_- B \times (0, 1) \setminus U$, in such a way that $\tilde{F}_-$ is identified with the horizontal surface $\partial_- B \times \{1/4\} \setminus \mathrm{Int}U_-'$ whereas $\tilde{F}_+$ is identified with the horizontal surface $\partial_+ B \times \{3/4\} \setminus \mathrm{Int}U_+'$.

Similarly, we parametrise N_B by a homeomorphism $I_N : N_B \to \partial_- B \times (0, 1) \setminus U$ in such a way that $I_N(\tilde{f}_3(\tilde{B}))$ lies in $\partial_- B \times [1/4, 3/4]$ and $I_N(W)$ lies in $\partial_- B \times [1/8, 7/8]$. Note that each component of ∂U corresponds to the boundary of a $\mathbb{Z}$-cusp neighbourhood of N_B. Since N_B is the covering of the non-cuspidal part N_0, we can extend N_B to a hyperbolic 3-manifold $\hat{N}_B$ which is the covering of N associated to $\pi_1(B)$ by attaching cusp neighbourhoods. Therefore, the parametrisation I_N extends to a homeomorphism $\hat{I}_N : \hat{N}_B \to \partial_- B \times (0, 1)$.

Since both $\tilde{F}_+$ and $\tilde{F}_-$ are disjoint from $\tilde{\mathcal{V}}[k_u]$, the solid torus $I_M(V)$ is contained in either $\partial_- B \times (0, 1/4)$ or $\partial_- B \times (1/4, 3/4)$ or $\partial_- B \times (3/4, 1)$. We shall first consider the case when $I_M(V)$ is contained in $\partial_- B \times (1/4, 3/4)$. Take a sufficiently small number s_0 so that both $\hat{I}_N^{-1}(\partial_- B \times (1 - s_0, 1))$ and $\hat{I}_N^{-1}(\partial_- B \times (0, s_0))$ are disjoint from the 1-neighbourhood of W. Since f_3 is proper and has degree 1, for sufficiently small $t_0 > 0$, the surfaces $I_N \circ \tilde{f}_3 \circ I_M^{-1}(\partial_- B \times \{t_0\} \setminus U)$ and $I_N \circ \tilde{f}_3 \circ I_M^{-1}(\partial_- B \times \{1 - t_0\} \setminus U)$ are contained in $\partial_- B \times (0, s_0)$ and $\partial_- B \times (1 - s_0, 1)$ respectively. Denote $I_M^{-1}(\tilde{F}_+ \times \{1 - t_0\} \setminus U)$ by $F\prime_+$, and $I_M^{-1}(\tilde{F}_- \times \{t_0\} \setminus U)$ by $F\prime_-$.

We can enlarge $F\prime_-$ and $F\prime_+$ to surfaces $\check{F}\prime_-$ and $\check{F}\prime_+$ homeomorphic to $\check{F}_-$ and $\check{F}_+$ respectively by joining pairs of parallel boundary components of $F\prime_-$ lying on ∂U_- by annuli on ∂U_- bounded by them, and those of $F\prime_+$ lying on ∂U_+ by annuli on ∂U_+ bounded by them. On the other hand, since $\tilde{f}_3(F\prime_-)$ and $\tilde{f}_3(F\prime_+)$ are disjoint from the 1-neighbourhood of W, we can enlarge $\tilde{f}_3(F\prime_-)$ and $\tilde{f}_3(F\prime_+)$ by joining each pair of parallel boundary component on $I_N \circ \tilde{f}_3(\partial U_- \cup \partial U_+) \subset \partial N_0$ by an annulus embedded in the closure of an ε-cusp neighbourhood which is a component of $N \setminus \mathrm{Int} N_0$ so that their images under $\hat{I}_N$ are contained in $\partial_- B \times (0, s_0)$ and $\partial_- B \times (1 - s_0, 1)$ respectively. These surfaces, which are homeomorphic to $\check{F}_-$ and $\check{F}_+$, are homotopic to embeddings $\bar{F}_-$ and $\bar{F}_+$ respectively outside the 1-neighbourhood of W by our choice of s_0, again using Theorem 9.3.6. Then by our choice of t_0 in the previous paragraph, we see that $\tilde{f}_3(\check{F}\prime_+)$ and $\tilde{f}_3(\check{F}\prime_-)$ are homotopic to $\bar{F}_-$ and $\bar{F}_+$ respectively by homotopies disjoint from $W \supset T$. Since $\tilde{f}_3(\tilde{F}_-)$ is homotopic to $\tilde{f}_3(\check{F}\prime_-)$ outside T and $\tilde{f}_3(\tilde{F}_+)$ is homotopic to $\tilde{f}_3(\check{F}\prime_+)$ outside T (for $I_M(V)$ is contained in $\partial_- B \times (1/4, 3/4)$ and $\tilde{f}_3(M_B \setminus V) \subset N_B \setminus T$), the surfaces $\tilde{f}_3(\tilde{F}_-)$ and $\tilde{f}_3(\tilde{F}_+)$ are homotopic to disjoint embeddings by homotopies disjoint from T.

Next suppose that $I_M(V)$ is contained in $\partial_- B \times (0, 1/4)$. In this case, we shall consider moving both $\tilde{F}_-$ and $\tilde{F}_+$ in the +-direction. As in the previous case, there are sufficiently small $s_0, t_0 > 0$ such that $\hat{I}_N^{-1}(\partial_- B \times (1 - s_0))$ is disjoint from the 1-neighbourhood of W, and such that $I_N \circ \tilde{f}_3 \circ I_M^{-1}(\partial_- B \times \{1 - t_0\})$ is contained in $\partial_- B \times (1 - s_0, 1)$. Then, by the same argument as in the previous case, we can see that both $\tilde{f}_3(\tilde{F}_-)$ and $\tilde{f}_3(\tilde{F}_+)$ are homotopic to an embedding contained in $\hat{I}_N^{-1}(\partial_- B \times (1 - s_0, 1))$ by a homotopy outside T. They can be homotoped to disjoint embeddings just by considering parallel copies of the embedding. Thus we are done also in this case. The argument for the case when $I_M(V)$ is contained in $\partial_- B \times (3/4, 1)$ is similarly just by changing the $+$-direction to the $-$-direction. □

The above claim says that a homotopy from $f_3(\tilde{F}_+ \sqcup \tilde{F}_-)$ to an embedding can be taken to be disjoint from W since any essential homotopy passing through W must intersect T. We can repeat the same argument for every relatively compact component of $N_0 \setminus \tilde{f}_3(\tilde{F}_+ \sqcup \tilde{F}_-)$ containing a component of $\tilde{T}[k_u]$ and show that $\tilde{f}_3(\tilde{F}_+ \sqcup \tilde{F}_-)$ can be homotoped to an embedding by a homotopy within $N_B[k_u]$.

Now, we consider a new hyperbolic metric m_N on Int$N[k_u]$ which makes every component of $T[k_u]$ a torus cusp preserving the original cusps of N. Pull back this metric to Int$N_B[k_u]$ and denote it by m_B. We consider a least area map $h_3 : \tilde{F}_- \sqcup \tilde{F}_+ \to (\mathrm{Int}N_B[k_u], m_B)$ homotopic to $\tilde{f}_3|\tilde{F}_- \sqcup \tilde{F}_+$. By Theorem 9.3.6, h_3 is an embedding.

In the following argument, we shall use the notion of topological order due to Brock–Canary–Minsky [12] which we explained in Definition 9.5.8.

Claim 9.5.11 *Let B be a brick in $\mathcal{K}_{\mathrm{int}}$ neither of whose fronts lies on the boundary of M_{int}. Then the embedding h_3 can be extended to an orientation-preserving embedding of $\tilde{B} \cap \mathrm{Int}M_B[k_u]$ to $(\mathrm{Int}N_B[k_u], m_B)$ taking $\tilde{B} \cap \tilde{\mathcal{V}}[k_u]$ to cusps corresponding to $\tilde{T}[k_u]$ and the homotopy classes of meridians of tube components of $\tilde{B} \cap \tilde{\mathcal{V}}[k_u]$ to those of $\tilde{T}[k_u]$.*

Proof Recall that there is a homeomorphism $I_N : N_B \to \partial_- B \times (0, 1) \setminus U$. By our definition of k_u, the images of the tube components of $\tilde{T}[k_u]$ under I_N are unknotted and unlinked in $\partial_- B \times (0, 1)$. Since ends of $h_3(\tilde{F}_- \sqcup \tilde{F}_+)$ other than those tending to cusps of N_B tend to $I_N(\partial\tilde{T}[k_u])$, the surfaces $I_N \circ h_3(\tilde{F}_-) \sqcup I_N \circ h_3(\tilde{F}_+)$ together with annuli on $I_N(\partial\tilde{T}[k_u])$ bound a submanifold homeomorphic to $\partial_- B \times [1/4, 3/4] \cong \tilde{B}$. We shall first prove that $I_N \circ h_3(\tilde{F}_+)$ is situated above $I_N \circ h_3(\tilde{F}_-)$. This trivially holds by definition when one of $\tilde{F}_+$ and $\tilde{F}_-$ is empty. Therefore, we assume that neither of them is empty. Since we assumed that neither $\partial_- B$ nor $\partial_+ B$ lie on the boundary of M_{int}, both $\partial_- \tilde{B} \cap \partial M_B$ and $\partial_+ \tilde{B} \cap \partial M_B$ are non-empty.

By Assumption 9.4.3, every component of $\partial_- B \cap \partial M$ overlaps some component of $\partial_+ B \cap \partial M$. Therefore, we can take components X and X' of $\tilde{\mathcal{V}}[k_u]$ on which boundary components of $\tilde{F}_+$ and of $\tilde{F}_-$ are respectively such that $X \cap \tilde{B}$ and $X' \cap \tilde{B}$ overlap. It follows that we have $X \cap \tilde{B} \prec_{\mathrm{top}} X' \cap \tilde{B}$. Since $\tilde{f}_3$ is a proper degree-1 map and $\tilde{f}_3|\tilde{\mathcal{V}}[k_u]$ is a homeomorphism onto its image, this implies that $\tilde{f}_3(X \cap \tilde{B}) \prec_{\mathrm{top}} \tilde{f}_3(X \cap \tilde{B})$. On the other hand, if $I_N \circ h_3(\tilde{F}_+)$ is situated under $I_N \circ h_3(\tilde{F}_-)$, then we should have $\tilde{f}_3(X' \cap \tilde{B}) \prec_{\mathrm{top}} \tilde{f}_3(X \cap \tilde{B})$, which is a contradiction. Thus we have proved that $I_N \circ h_3(\tilde{F}_+)$ is situated above $I_N \circ h_3(\tilde{F}_-)$ and h_3 extends to an orientation-preserving homeomorphism from $\tilde{B}$ to a submanifold B_N bounded by $h_3(\tilde{F}_- \sqcup \tilde{F}_+)$.

We shall next show that this homeomorphism induces one between $\tilde{B} \cap \mathrm{Int}M_B[k_u]$ to $B_N \cap \mathrm{Int}N_B[k_u]$. For that, it suffices to show that for the components of $\tilde{\mathcal{V}}[k_u]$ in $\tilde{B}$, the corresponding components of $\tilde{T}[k_u]$ are contained in B_N preserving the topological order since all such components in B_N are unknotted and unlinked. Let V be a component of $\tilde{\mathcal{V}}[k_u]$ contained in $\tilde{B}$. Then we have $\tilde{F}_- \prec_{\mathrm{top}} V \prec_{\mathrm{top}} \tilde{F}_+$. Let T be a component of $\tilde{T}[k_u]$ with $T = \tilde{f}_3(V)$. Since $\tilde{f}_3$ is a proper degree-1 map and takes $M_B \setminus V$ to $N_B \setminus T$, we see that $\tilde{f}_3(\tilde{F}_-) \prec_{\mathrm{top}} T \prec_{\mathrm{top}} \tilde{f}_3(\tilde{F}_+)$. Since h_3, defined on $\tilde{F}_- \sqcup \tilde{F}_+$, is homotopic to $f_3|(\tilde{F}_- \sqcup \tilde{F}_+)$ in $N_B[k_u]$, we also have $h_3(\tilde{F}_-) \prec_{\mathrm{top}} T \prec_{\mathrm{top}} h_3(\tilde{F}_+)$. Therefore any tube component of $\mathcal{V}[k_u]$ in $\tilde{B}$ has its corresponding Margulis tube in B_N. Now suppose that we have two such tube components V_1, V_2 with $V_1 \prec_{\mathrm{top}} V_2$. Let T_1, T_2 be the components of $\tilde{T}[k_u]$ with $\tilde{f}_3(V_1) = T_1$ and $\tilde{f}_3(V_2) = T_2$. Then by the same argument as above using the bijective correspondence between the components of $\tilde{\mathcal{V}}[k_u]$ and $\tilde{T}[k_u]$, we have $T_1 \prec_{\text{top}} T_2$. Thus we have shown that we have a homeomorphism $\bar{h}_3$ from $\tilde{B} \cap \text{Int}M_B[k_u]$ onto $B_N \cap \text{Int}N_B[k_u]$ which is an extension of h_3.

It remains to show that a meridian of a solid torus component of $\tilde{\mathcal{V}}[k_u]$ contained in $\tilde{B}$ is taken to a meridian of $\tilde{T}[k_u]$ by $\bar{h}_3$. This is rather obvious from our construction: for $\tilde{f}_3$ takes meridians of solid torus components of $\tilde{\mathcal{V}}[k_u]$ to those of $\tilde{T}[k_u]$. □

Now, for each brick B of $\mathcal{K}_{\text{int}}$ neither of whose fronts lies on the boundary of M_{int}, we consider $B \cap \text{Int}M[k_u]$, its lift $\tilde{B}$ in $\text{Int}M_B[k_u]$, and its embedding into $\text{Int}N_B[k_u]$ by an extension of the least-area map h_3 given above, which we denote by B_N as above. We denote the map taking $B \cap \text{Int}M[k_u]$ to B_N obtained in this way by f_B. We regard B_N as a hyperbolic 3-manifold with boundary by restricting the metric m_B, and call B_N with this metric the *least-area realisation of* B. In the case when B is a brick one of whose front lies on the boundary of M_{int}, we define B_N to be a submanifold of $\text{Int}N_B[k_u]$ homeomorphic to $\partial_- B \times (0, 1)$ obtained by cutting $\text{Int}N_B[k_u]$ along the embedding of one of the boundary components of $\tilde{B}$ whose projection in M does not lie on the boundary of M_{int}, defined using the least area map in the same way as above.

For a brick in $B \cong F \times [-1, \infty)$ in $\mathcal{K}_{\text{gf}}$, in the same way as bricks in $\mathcal{K}_{\text{int}}$, we consider homotoping $\tilde{f}_3|(F \times \{-1\} \cap M[k_u])$ to a least area surface in $N_B[k_u]$ and constructing a realisation B_N of $B \cap \text{Int}M[k_u]$ in $\text{Int}N_B[k_u]$. Since the least area surface is contained in the convex core of N_B, we can assume that for the part $F \times [0, \infty)$, the realisation B_N is just a lift of $f_3(F \times [0, \infty))$ to N_B.

Suppose that two bricks B^1 and B^2 of $\mathcal{K}$ share a joint F. We can assume F is a component of $\partial_+ B^1$ and $\partial_- B^2$ by interchanging B^1 and B^2 if necessary. Construct least-area realisations B^1_N and B^2_N as above. Then both of their boundaries contain a least area surface corresponding to F as components. We denote by F^j the one contained in ∂B^j_N for $j = 1, 2$. Since the projections of F^1 and F^2 in $(\text{Int}N[k_u], m_N)$ are least-area surfaces homotopic to $f_3(F)$ (which might not be embeddings), they must coincide. Therefore, F^1 is isometric to F^2. Then we can consider the hyperbolic 3-manifold homeomorphic to $(\text{Int}B^1 \cup \text{Int}B^2 \cup F) \cap \text{Int}M[k_u]$ by pasting B^1_N and B^2_N along F^1 and F^2 by an isometry.

Repeating this procedure for every joint on B^1 and B^2, then again for all the bricks, we get a hyperbolic 3-manifold $N'[k_u]$ homeomorphic to $\text{Int}M[k_u]$. We denote the homeomorphism obtained by identifying $B \cap \text{Int}M[k_u]$ with B_N in $N'[k_u]$ by $h : \text{Int}M[k_u] \to N'[k_u]$. We shall show that this manifold is isometric to $(\text{Int}N[k_u], m_N)$.

Claim 9.5.12 *There is an isometry* $f' : N'[k_u] \to (\text{Int}N[k_u], m_N)$*, whose restriction to* B_N *for each brick* B *is an isometric embedding homotopic to* $f_3 \circ f_B^{-1}$ *in* N_0.

Proof For each brick B, by Claim 9.5.11, there is an (extended) embedding $h_3 : \tilde{B} \cap \text{Int}M_B[k_u] \to \text{Int}N_B[k_u]$ homotopic to $\tilde{f}_3|B \cap \text{Int}M_B[k_u]$. If we lift $f_B^{-1}(B_N)$

to Int$M_B[k_u]$, and embed it by h_3 into Int$N_B[k_u]$, then the map is isometric by our definition of the metric on $N'[k_u]$. By projecting it to $N[k_u]$, we get a locally isometric map from B_N, which was defined above and is bounded by least area surfaces, into Int$N[k_u]$. Since for two bricks sharing a joint, such maps induce the same map on the joint, we can glue this map at joints and get a local isometry $f' : N'[k_u] \to \text{Int}N[k_u]$. (Note that if two bricks share a joint, then their images by h_3 lie on the opposite sides of the image of the joint by our way of extending h_3 in Claim 9.5.11, which guarantees that the map is also a local isometry at joints.) Since h_3 is homotopic to $f_3|B$, we see that $f' \circ h_3$ is homotopic to f_3.

Since f' induces an isomorphism between fundamental groups, to show that it is an isometry, it is sufficient to show that f' is proper. Suppose, seeking a contradiction, that f' is not proper. Then, there exists a sequence of distinct bricks B^i of $\mathcal{K}_{\text{int}}$ and points $x_i \in B^i_N$ such that $\{f'(x_i)\}$ converges in Int$N[k_u]$. Since $\{f'(x_i)\}$ converges, the injectivity radius at $f'(x_i)$ is bounded below by a positive constant independent of i, hence so is the injectivity radius at x_i. We divide our argument depending on whether the distance between x_i and ∂B^i_N is bounded or not as $i \to \infty$.

First we consider the case when the distance from x_i and ∂B^i_N is bounded as $i \to \infty$. Let F^i be a least-area boundary component of B^i_N from which x_i is within uniformly bounded distance. Since $\xi(F^i) \leq \xi(S)$, the diameter of the thick part of F^i is uniformly bounded. Since x_i lies in the thick part, it is within uniformly bounded distance from either an ε_1-Margulis tube or an ε-cusp neighbourhood touching F^i which corresponds to a component $\tilde{V}^i$ of $\tilde{T}[k_u]$. We denote by V^i a component of $T[k_u]$ which is the projection of $\tilde{V}^i$.

We can show that in (Int$N[k_u], m_N$), for each component V of $T[k_u]$ there are only finitely many images of joints by f' touching V as follows. For any $R > 0$, there are a finitely many components of $T[k_u]$ and original cusp neighbourhoods of N which can be reached from V within the distance R modulo the ϵ_0-thin part. Since joints are homeomorphic to subsurfaces of S and the boundaries of their images in $N[k_u]$ lie in $T[k_u] \cup \partial N_0$ as longitudes or core curves, there are only finitely many possibilities for the homotopy classes of the boundaries of their images in $N[k_u]$. This implies that there are only finitely many joints up to homotopy whose images can touch V since there are at most two kinds of homotopy classes of horizontal surfaces if we fix a boundary. Since no two distinct joints are homotopic as we removed inessential joints, it follows that there are only finitely many joints whose images touch V.

Since our joints F^i are all distinct, we can assume that all the V^i are distinct by taking a subsequence. Since f_3 takes the components of $\mathcal{V}[k_u]$ to those of $T[k_u]$ one-to-one, and no other part of M is mapped to $T[k_u]$, we see that f' takes the V^i to distinct components of $T[k_u]$. Therefore $f'(x_i)$ is within bounded distance from infinitely many distinct components of $T[k_u]$. Since the $f'(x_i)$ are assumed to converge, this contradicts the fact that there are only finitely many components of $T[k_u]$ within a bounded distance.

Thus, it only remains to consider the case when the distance from x_i to the boundary of B^i_N goes to ∞ as $i \to \infty$. Recall that B^i_N was originally a submanifold in $\mathrm{Int}N_{B^i}[k_u]$. Therefore, we can regard x_i also as a point in $\mathrm{Int}N_{B^i}[k_u]$. Since B^i_N is bounded by least-area surfaces, it is contained in the convex core of $(\mathrm{Int}N_{B^i}[k_u], m_{B^i})$. Therefore, there is a pleated surface $k_i : \partial_- B^i \to \mathrm{Int}N_{B^i}[k_u]$ which is within bounded distance from x_i and is homotopic to the inclusion of $\partial_- B^i$ as $\partial_- B^i \times \{t\}$ with respect to the parametrisation $N_{B^i} \cong \partial_- B^i \times (0, 1)$. Since the distance from x_i to ∂B^i_N goes to ∞, we can assume that the image of k_i is contained in B^i_N. Hence we can regard k_i as a pleated surface in $N'[k_u]$. Also since the cuspidal part of $N'[k_u]$ consists of those of N_0 and rank-2 cusps corresponding to $T[k_u]$, we can take cusp neighbourhoods small enough to be disjoint from all the images of k_i.

We consider the pleated surfaces $f' \circ k_i$. Since $\{f'(x_i)\}$ converges and $f' \circ k_i$ is disjoint from the cusp neighbourhoods which are images of those taken above, the sequence of pleated surfaces $\{f' \circ k_i\}$ converges geometrically inside $(\mathrm{Int}N[k_u], m_N)$, passing to a subsequence. This implies in particular that there are distinct i_1, i_2 such that $f' \circ k_{i_1}$ and $f' \circ k_{i_2}$ are properly homotopic. Since f' induces an isomorphism between fundamental groups, it follows that k_{i_1} and k_{i_2} are properly homotopic. This is a contradiction since no two horizontal surfaces of distinct bricks are properly homotopic. (Recall that $N'[k_u]$ and $\mathrm{Int}M[k_u]$ are homeomorphic.) Thus we have established that f' is an isometry. By our construction, it is evident that $f'|B_N$ is homotopic to $f_3 \circ f_B^{-1}$ in N_0. □

Thus $\mathrm{Int}N[k_u]$ is isometric to $N'[k_u]$ which is the union of the B_N each of which is homeomorphic to $B \cap \mathrm{Int}M[k_u]$. This shows that there is a homeomorphism $h : \mathrm{Int}M[k_u] \to N'[k_u]$ such that $f' \circ h$ is homotopic to $f_3|\mathrm{Int}M[k_u]$. By setting f_4 to be the natural extension of $f' \circ h$ to $M[k_u]$, we get a homeomorphism as we wanted.

It only remains to show that f_4 extends to a homeomorphism between M and N_0. To show this, it suffices to show that for each component V of $\mathcal{V}[k_n]$, its meridian is sent to a meridian of a component of $T[k_u]$. If V is contained in some brick B, then this follows from Claim 9.5.11. Since we isotoped the original brick decomposition to a new one by moving each joint F to $\check{F}$, we see that every component of $\mathcal{V}[k_u]$ is contained in some brick.

This completes the proof of Proposition 9.5.9. □

Having proved that $M[k_u]$ and $N[k_u]$ are homeomorphic, we shall next show that the Lipschitz map f_3 can be homotoped so as to embed the joints preserving the Lipschitzness. For that, it is more convenient to consider a brick decomposition of $M[k_u]$ rather than that of M. As in Sect. 9.4.4, we define a brick of $M[k_u]$ to be a maximal union of vertically parallel horizontal leaves which are inherited from the horizontal foliation of M. By the same argument as in Sect. 9.4.4, we can check conditions A-(1)–(5) are satisfied. (In reality, only A-(2) and A-(3) need to be checked.) *We denote this brick decomposition of* $M[k_u]$ by $\mathcal{K}[k_u]$.

Before changing the images of joints to Lipschitz embeddings, we shall first move f_3 so that it preserves the order of joints on the boundary except for parallel ones. Let $\mathcal{F}$ be the union of joints of pairs of bricks in $\mathcal{K}[k_u]$. We introduce an equivalence relation $\sim$ in the set of components of $\mathcal{F}$ such that $F_1 \sim F_2$ if they are parallel. By our definition of brick decomposition, there are no three distinct joints in $\mathcal{F}$ which are all parallel. Therefore each equivalence class consists of at most two joints. We define the reduced union of joints to be the union of joints taken one from each equivalence class, and denote it by $\hat{\mathcal{F}}$.

Lemma 9.5.13 *There is a uniform constant K_3' as follows. We can homotope f_3 to a proper, degree-1 map $f_3' : M[k_u] \to N[k_u]$ with the following properties.*

- (i) *f_3' coincides with f_3 outside small pairwise disjoint neighbourhoods of the components of $\partial M[k_u]$.*
- (ii) *f_3' is K_3'-Lipschitz.*
- (iii) *On each component T of $\partial M[k_u]$, distinct components of $\mathcal{F} \cap T$ have disjoint images under f_3'.*
- (iv) *On each component T of $\partial M[k_u]$, the restriction $f_3'|T$ maps the components of $\hat{\mathcal{F}} \cap T$ disjointly preserving the orientation of $\hat{\mathcal{F}} \cap T$ and the order of $\{F \cap T \mid F$ is a component of $\hat{\mathcal{F}}\}$. (When T is a torus the order means the cyclic order.)*
- (v) *For a component F of $\mathcal{F} \setminus \hat{\mathcal{F}}$, let $\hat{F}$ be the other component of $\mathcal{F}$ equivalent to F and contained in $\hat{\mathcal{F}}$. Then f_3' also preserves the order of $((\hat{\mathcal{F}} \setminus \hat{F}) \cup F) \cap T$ for any component F of $\mathcal{F} \setminus \hat{\mathcal{F}}$.*
- (vi) *The order of $F \cap T$ and $\hat{F} \cap T$ as in (v) may be reversed only when $f_3'(F) \cap f_3'(\hat{F}) = \emptyset$.*
- (vii) *For each small $\delta > 0$, there is a universal number n_0 such that for any component F of $\mathcal{F}$, there are at most n_0 joints F_i such that $f_3'(F_i \cap T)$ are within distance δ from $f_3'(F \cap T)$.*

Proof Let T be a component of $\partial M[k_u]$, which is either a torus or an open annulus. As was shown before, T consists of horizontal annuli and vertical annuli, and the joints can intersect only vertical annuli, their boundaries included. We shall show that we can homotope $f_3|T$ to a uniformly Lipschitz map with desired properties by a homotopy moving each point at a uniformly bounded distance. We should note that $f_3|T$ is a degree-1 map to a boundary component T' of $N[k_u]$. The foliation of M by horizontal leaves induces a foliation on T whose leaves are parallel horizontal circles. By our definition of the model metric, each leaf has length ε_1. We can extend this foliation also to horizontal annuli so that they are also foliated by parallel circles with length ε_1. We let γ be a simple closed geodesic with respect to the induced metric intersecting each leaf at one point when T is a torus, and a geodesic ray intersecting each leaf at one point when T is an open annulus.

Since f_3 is K_3-Lipschitz, the homotopy class in T' of the images of the leaves has geodesic length bounded by $K_3\varepsilon_1$ with respect to the Euclidean metric on T'. We also note that this length is also bounded below by ε_1 since T' lies on the boundary of an ε_1-Margulis tube. We first homotope $f_3|T$ fixing $f_3|\gamma$ so that for each leaf l of the foliation on T, the simple closed curve $\bar{f}_3(l)$ is a closed geodesic with respect to the Euclidean metric on T'. Also, if there are distinct components of $\mathcal{F} \cap T$ which have the same image, we can perturb the map by a homotopy to make their images disjoint moving them within a very small distance. Thus we get a map $\bar{f}_3$ hoomotopic to f_3. We can take a homotopy $H_3 : A \times [0, 1] \to T'$ from f_3 to $\bar{f}_3$ as a $\bar{K}_3$-Lipschitz map, where $\bar{K}_3$ depends only on ε_1 and K, since the length of each closed curve $f_3(l)$ is between ε_1 and $K\varepsilon_1$ and the perturbation moves the images at uniformly bounded distances.

Now, the map from γ to $f_3(\gamma) = \bar{f}_3(\gamma)$ may not proceed in the positive direction monotonously. (As we shall see below, since $f_3|T$ has degree 1, the orientations of T and T' determine the positive direction to which $\bar{f}_3(\gamma)$ should proceed.) This may cause a permutation of the order of $\hat{\mathcal{F}} \cap T$ by $\bar{f}_3$. We fix an orientation of the foliation on T, which, together with the orientation of T, induces a transverse orientation of the leaves and an orientation of γ. This also defines a transverse orientation of the foliation on T' induced by the closed geodesics which are images of the leaves on T, since $f_3'|T$ has degree 1. We number the simple closed curves constituting $\hat{\mathcal{F}} \cap T$ as $F_1, F_2, \ldots$ in accordance with the order determined by the orientation of γ. In the case when T is a torus, we fix a leaf on the lower horizontal annulus, and let its intersection with γ, which we denote by a_0, be the starting point. The transverse orientation of the foliation on T' gives an order on the images $\bar{f}_3(F_1 \cap T), \ldots$, which may be different from the order on T. (We allow some of them to go beyond $\bar{f}_3(a_0)$ in the negative direction. As long as $\bar{f}_3(\gamma)$ moves in the negative direction, we regard it as receding with respect to the order on T'.) Let σ be a permutation such that $\bar{f}_3(F_{\sigma(1)}), \ldots$ is the right order on T'; in other words F_i is mapped to the $\sigma^{-1}(i)$-th curve with respect to the order on T'. Now, we first look at $\bar{f}_3(F_1 \cap T)$ which is the $\sigma^{-1}(1)$-th curve on T', and consider the curves $\bar{f}_3(F_{\sigma(1)} \cap T), \ldots, \bar{f}_3(F_{\sigma(\sigma^{-1}(1)-1)} \cap T)$ which are those situated before $\bar{f}_3(F_1 \cap T)$ on T'. Set $j = \max\{\sigma(1), \ldots, \sigma(\sigma^{-1}(1) - 1)\}$. We shall consider moving $\bar{f}_3(F_1 \cap T), \ldots, \bar{f}_3(F_j \cap T)$ to correct their order. The point in the following argument is that this can be done by a homotopy with bounded displacement.

Using Theorem 9.3.6, we shall bound uniformly the distance between any two of $\bar{f}_3(F_1 \cap T), \ldots, \bar{f}_3(F_j \cap T)$. Let k be a number among $2, \ldots, j$. First consider the case when $\bar{f}_3(F_k \cap T)$ comes before $\bar{f}_3(F_1 \cap T)$ on T'. Recall that by Proposition 9.5.9, f_3 is homotopic in $M[k_u]$ to a homeomorphism $f_4 : M[k_u] \to N[k_u]$. By the same procedure as we used to construct $\bar{f}_3$ from f_3, we can assume that f_4 also maps each leaf on T to a closed geodesic with respect to the induced Euclidean metric on T'. Then, since both F_1 and F_k are incompressible, by Theorem 9.3.6, we can homotope $\bar{f}_3|F_1$ and $\bar{f}_3|F_k$ fixing the boundaries to embeddings g_3^1 and g_3^k in $N[k_u]$ which are contained in small regular neighbourhoods of $\bar{f}_3(F_1)$ and $\bar{f}_3(F_k)$ respectively. By perturbing g_3^1 and g_3^k, we can assume that they are transverse to $f_4(F_1)$ and $f_4(F_k)$ at their interiors. Then $(g_3^1(F_1) \cup f_4(F_1)) \cap T'$ bounds an annulus A_1' which may degenerate to a circle. When T is a torus, there are two choices for A_1'. We choose one which bounds a compact region with subsurfaces on $g_3^1(F_1)$ and $f_4(F_1)$ (possibly together with annuli on other components of $\partial N[k_u]$) which is disjoint from $g_3^k(F_k)$ ($k \neq 1$) with $f_4(F_1) \cap g_3^k(F_k) = \emptyset$. Similarly, we define an annulus A_k' for $g_3^k(F_k)$ and $f_4(F_k)$. Since $g_3^1(F_1 \cap T)$ comes after $g_3^k(F_k \cap T)$ whereas $f_4(F_1 \cap T)$ is situated before

$f_4(F_k \cap T)$, we see that A'_1 and A'_k must intersect. Since $f_4(F_1) \cap f_4(F_k) = \emptyset$, both F_1 and F_k are connected, and by our definition of $\hat{\mathcal{F}}$, F_1 and F_k are not parallel, we see that $g_3^1(F_1)$ and $g_3^k(F_k)$ must intersect at their interiors. By our construction of g_3^1 and g_3^k, this implies that $\bar{f}_3(F_1)$ and $\bar{f}_3(F_k)$ also intersect at their interiors. Next suppose that $\bar{f}_3(F_k \cap T)$ comes after $\bar{f}_3(F_1 \cap T)$. By our definition of j, we see that $\bar{f}_3(F_j \cap T)$ comes before $\bar{f}_3(F_1 \cap T)$, hence also before $\bar{f}_3(F_k \cap T)$. Since $k < j$, the order of $F_j \cap T$ and $F_k \cap T$ is reversed under $\bar{f}_3$, and we can argue in the same way as above to conclude that $\bar{f}_3(F_j)$ and $\bar{f}_3(F_k)$ intersect at their interiors.

Recall that the diameters of the joints $F_1, \ldots$ are uniformly bounded from above by a constant depending only on $\xi(S)$. Since $\bar{f}_3$ is uniformly Lipschitz, their images $\bar{f}_3(F_1), \ldots$ also have diameters bounded from above by a constant λ depending only on $\xi(S)$. This implies that for any $k = 2, \ldots, j$, the distance between $\bar{f}_3(F_k \cap T)$ and either $\bar{f}_3(F_1 \cap T)$ or $\bar{f}_3(F_j \cap T)$ is bounded by 2λ. Therefore the distance between any two of $\bar{f}_3(F_1 \cap T), \ldots, \bar{f}_3(F_j \cap T)$ is bounded by 4λ. Recall that $\bar{f}_3(F_1 \cap T), \ldots, \bar{f}_3(F_p \cap T)$ are parallel closed geodesics on T' with respect to the induced Euclidean metric. By the uniform quasi-convexity of horoballs, we see that there is a number λ_0 depending only on $\xi(S)$ which bounds the distance on T' between any two of $\bar{f}_3(F_1 \cap T), \ldots, \bar{f}_3(F_j \cap T)$ with respect to the induced Euclidean metric. Then we can homotope $\bar{f}_3|T$ so that $\bar{f}_3(F_1 \cap T), \ldots, \bar{f}_3(F_j \cap T)$ lie in the right order on T' and near the original position of $\bar{f}_3(F_{\sigma(1)} \cap T)$ so that all $\bar{f}_3(F_i \cap T)$ with $i > j$ come after them, without changing the condition that every leaf is mapped to a closed geodesic preserved, by moving the image by $\bar{f}_3$ of thin neighbourhoods of $F_1 \cap T, \ldots, F_j \cap T$ only at distance at most $\lambda_0 + 1$. The map which we get after this homotopy is also uniformly Lipschitz since the displacement of the points by the homotopy is uniformly bounded.

We now forget about $F_1, \ldots, F_j$ and only consider $F_{j+1}, \ldots$. If $\sigma(j+1) = j+1$, we also forget about F_{j+1} and proceed to the first $j_0 > j$ with $\sigma(j_0) \neq j_0$. Otherwise we let j_0 be $j+1$. Regarding $\bar{f}_3(F_{j_0} \cap T)$ instead of $\bar{f}_3(F_1 \cap T)$ as the first one, we repeat the same argument. Then we can correct the order of $\bar{f}_3(F_{j_0} \cap T), \ldots, \bar{f}_3(F_{j_1} \cap T)$ for $j_1 > j_0$ which is defined to be $\max\{\sigma(j_0), \ldots, \sigma(\sigma^{-1}(j_0) - 1)\}$ and make them come after $\bar{f}_3(F_{j_0 - 1})$ by moving $\bar{f}_3$ in thin neighbourhoods of $F_{j_0} \cap T, \ldots, F_{j_1} \cap T$ only at distance less than $\lambda_0 + 1$. We note that we do not touch the components $F_k \cap T$ with $k < j_0$ at this stage. We repeat the same process, and eventually we can homotope $\bar{f}_3|T$ to a uniformly Lipschitz map $f_3^T : T \to T'$ which preserves the order of $F_1 \cap T, \ldots$ by a homotopy moving every point within the distance $\lambda_0 + 1$. (To be more precise, we need to define the homotopy inductively in the case when there are infinitely many components of $\hat{\mathcal{F}} \cap T$.)

Having moved $\bar{f}_3|T$ to f_3^T which preserves the order of $\hat{\mathcal{F}} \cap T$, we now turn to considering a component F of $\mathcal{F} \setminus \hat{\mathcal{F}}$. Suppose that f_3^T does not preserve the order of $\hat{\mathcal{F}} \cap T$ if we replace $\hat{F}$ in $\hat{\mathcal{F}}$ with F. Then for each component F' of $\hat{\mathcal{F}} \setminus \hat{F}$ such that the order between F and F' is reversed by f_3^T, we see that by the same argument, F must intersect F', and we can move $\bar{f}_3$ in a thin neighbourhood of $F \cap T$ within the distance $\lambda_0 + 1$ to correct the order. Moreover, in the same way, we can correct the order of the images $F \cap T$ and $\hat{F} \cap T$ under f_3^T by moving f_3^T within the distance λ_0+1 if $\bar{f}_3(F)$ and $\bar{f}_3(\hat{F})$ intersect. We note that during this homotopy, each component of $\mathcal{F}$ is moved at most twice; hence the displacement is bounded independently of the number of components of $\mathcal{F}$. Thus we have shown that if we construct a uniform Lipschitz map whose restriction to T is f_3^T, then conditions (iii), (iv) and (v) in the statement are satisfied. We denote a homotopy on T by H_3'. This homotopy H_3' is uniformly Lipschitz since the homotopy only passes through uniformly Lipschitz maps and its displacement function is uniformly bounded.

We next show that f_3^T thus obtained satisfies condition (vii) (with f_3' in the statement replaced by f_3^T). Fix some component F of $\mathcal{F}$; and we shall bound the number of components F' of $\mathcal{F}$ such that $f_3^T(F\cap T)$ and $f_3^T(F'\cap T)$ are within the distance δ. Since f_3^T is obtained from $\bar{f}_3|T$ by moving at most the distance λ_0+1, if $f_3^T(F'\cap T)$ is within distance δ from $f_3^T(F\cap T)$, then $f_3(F')$ is within distance $\delta+2(\lambda_0+1)$ from $f_3(F)$. By our construction of the brick decomposition of $M[k_u]$, for each joint, there is at most one other joint to which it is parallel. Therefore, for any natural number ν there exists n such that if there are n distinct joints, then there are at least ν boundary components of $M[k_u]$ which these joints intersect or which is contained in a region cobounded by two among these joints which is foliated by horizontal surfaces. Since f_3 is uniformly Lipschitz, the diameter of the image of each horizontal surface under f_3 is bounded by a constant λ_1 depending only on $\xi(S)$. We note that if a component T of $\partial N[k_u]$ is contained in a region cobounded by two joints, then it is within distance $2\lambda_1$ from every point on each of the joints. Now, since there is a bound for the number of components of $\partial N[k_u]$ which can be reached from F within distance $\delta+2(\lambda_0+1)+2\lambda_1$, we get n_0 bounding the number of components of $\mathcal{F}\cap T$ whose images by f_3^T are within distance δ from $f_3^T(F\cap T)$.

We finally show that the map f_3^T defined above can be extended to a uniform Lipschitz map f_3'. We can take $r>0$ depending only on ε_1 and $\xi(S)$ such that the boundary components of $M[k_u]$ have product r-neighbourhoods in $M[k_u]$ which are pairwise disjoint. Let $\mathcal{N}_r(T)$ denote the r-neighbourhood in $M[k_u]$ of a boundary component T of $M[k_u]$; and we parametrise $\mathcal{N}_r(T)$ by $T\times[0,r]$ so that $T\times\{t\}$ is at distance t from T. We modify f_3 only inside $\cup\mathcal{N}_r(T)$ to get f_3'. We first define $f_3'|T\times[2r/3,r]$ to be rescaled $f_3|\mathcal{N}_r(T)$ so that $f_3'|\partial V\times\{2r/3\}$ is naturally identified with $f_3|T$. Next we define $f_3'|T\times[r/3,2r/3]$ to realise the homotopy H_3 so that $f_3'|T\times\{t\}$ corresponds to $H_3(\ ,2-3t/r)$. Finally we define $f_3'|T\times[0,r/3]$ to realise the homotopy H_3', so that $f_3'|T\times\{t\}$ corresponds to $H_3'(\ ,1-3t/r)$. Since H_3 and H_3' are uniformly Lipschitz, we see that there is a uniform constant K_3' such that f_3' is K_3'-Lipschitz. □

Lemma 9.5.14 *Let $\mathcal{F}$ be the union of the joints of pairs of bricks in $\mathcal{K}[k_u]$ as defined above. Then, there exists a K'-Lipschitz homotopy $H:\mathcal{F}\times[0,1]\to N[k_u]$ fixing the boundary of $\mathcal{F}$ with K' depending only on $\xi(S)$ such that the following hold.*

(i) $H|\mathcal{F}\times\{0\}$ *coincides with* $f_3'|\mathcal{F}$.
(ii) $H(x,t)=f_3'(x)$ *for every* $x\in\partial M[k_u]\cap\mathcal{F}$.

(iii) $H|\mathcal{F} \times [1/2, 1]$ *is a* C^2*-embedding.*
(iv) *For each component* F *of* $\mathcal{F}$*, the restriction* $H|F \times [1/2, 1]$ *is* K'*-bi-Lipschitz.*

Proof Let F be a component of $\mathcal{F}$. Since by condition (5) in Sect. 9.5.3, the geodesic lengths of core curves in $\mathcal{V}[0] \setminus \mathcal{V}[k_u]$ are bounded below by ϵ_u, and $F \setminus \mathcal{V}$ consists of thrice-punctured spheres, the modulus of F is uniformly bounded. By condition (7) in Sect. 9.5.3 and our choice of k_u, we see that there is no essential closed curve with length less than ϵ_u in $N[k_u]$. This shows that the map $f_3'|F$ is a uniformly bi-Lipschitz map onto its image. (We should note that $f_3'|F$ may not be injective. The bi-Lipschitzness here means that the metric on F induced from $M[k_u]$ and the one induced from $N[k_u]$ by f_3' are bi-Lipschitz equivalent.) We can approximate $f_3'|\mathcal{F}$ by an immersion fixing the boundary and preserving the uniform bi-Lipschitzity. Now, by Proposition 9.5.9, $f_3'|\mathcal{F}$ is properly homotopic to an embedding in $N[k_u]$ (not fixing the boundary).

We shall first show that each component F of $\mathcal{F}$ can be homotoped fixing the boundary to a uniformly bi-Lipschitz embedding. Suppose, seeking a contradiction, that this is not the case. Then there exist sequences of labelled brick manifolds M^i, homeomorphisms $f^i : M^i \to N^i$, Lipschitz maps $f_3^i : M^i[k_u] \to N^i[k_u]$ corresponding to f_3' constructed above, and joints F^i in $M^i[k_u]$ such that an embedding g_3^i as above within the δ-neighbourhood of $f_3^i(F^i)$ cannot be made K_i-bi-Lipschitz, with $K_i \to \infty$. We put the superscript i for all the symbols related to M^i and N^i. By taking a subsequence we can assume that all the $\hat{F}^i$ are homeomorphic to some fixed surface F. As shown before, by our definition of the model metric, the moduli of the F^i are bounded. Therefore, we can choose a homeomorphism $\kappa_i : F \to F^i$ so that the pullback of the metric on F^i by κ_i converges as $i \to \infty$. Take a basepoint x on F, and consider geometric limits of $(F^i, \kappa_i(x))$, $(M^i[k_u], \kappa_i(x))$, and $(N^i[k_u], f_3^i \circ \kappa_i(x))$. Since f_3^i is uniformly Lipschitz, it converges to a Lipschitz map $f_3^\infty : M^\infty[k_u] \to N^\infty[k_u]$, where $M^\infty[k_u]$ and $N^\infty[k_u]$ are the geometric limits of $(M^i[k_u], \kappa_i(x))$ and $(N^i[k_u], f_3 \circ \kappa_i(x))$ respectively. Since the metrics induced from the F^i on F are bounded, the homeomorphism k_i converges to a homeomorphism $\kappa_\infty : F \to F^\infty$, where F^∞ is embedded in $M^\infty[k_u]$.

As before, we can assume that both $f_3^i \circ \kappa_i$ and $f_3^\infty \circ \kappa_\infty$ are immersions. By Theorem 9.3.6 as was used in the proof of Proposition 9.5.9, $f_3^i \circ \kappa_i$ is homotopic to a least-area embedding relative to the boundary by a homotopy passing through only relatively compact components of $N^i[k_u] \setminus f_3^i \circ \kappa_i(F)$. Since $N[k_u]$ contains no Margulis tubes whose core curves have lengths less than ϵ_u and the diameters of the $f_3^i \circ \kappa_i(F)$ are bounded, these components have uniformly bounded diameters and converge geometrically to relatively compact components of $N^\infty[k_u] \setminus f_3^\infty \circ \kappa_\infty(F)$ through which $f_3^\infty \circ \kappa_\infty$ can be homotoped to an embedding (after perturbation if necessary). Therefore, the geometric limit $f_3^\infty \circ \kappa_\infty$ can be homotoped to a bi-Lipschitz embedding in $N^\infty[k_u]$. By pulling back this embedding and using a homotopy, we can homotope $f_3^i \circ \kappa_i$ to a uniformly bi-Lipschitz embedding, contradicting our assumption. Thus we have shown that $f_3'|F$ can be homotoped to a uniformly bi-Lipschitz least-area embedding, which we shall let be $H(\cdot\,, 3/4)$

on F. The above argument also shows that we can choose a homotopy H on F between $H(\cdot\,, 3/4)|F$ and $H(\cdot\,, 0) = f_3'|F$ to be uniformly Lipschitz.

Since f_3' preserves the order of $\hat{\mathcal{F}} \cap T$ as was shown in Lemma 9.5.13-(iii), f_3' is homotopic to a homeomorphism from $M[k_u]$ to $N[k_u]$ fixing $\hat{\mathcal{F}} \cap T$. Therefore the least area surfaces homotopic to the restrictions of f_3' to the components of $\hat{\mathcal{F}}$ fixing the boundary must be pairwise disjoint. The same holds even if we put F of $\mathcal{F} \setminus \hat{\mathcal{F}}$ into $\hat{\mathcal{F}}$ removing its counterpart $\hat{F}$ instead. Therefore, to show the disjointness of the least-area images of the components of $\mathcal{F}$, it suffices to show that the least area surfaces homotopic to $f_3'(F)$ and $f_3'(\hat{F})$ are disjoint for each component F of $\mathcal{F} \setminus \hat{\mathcal{F}}$. This follows immediately from Theorem 9.3.6 when $f_3'(F)$ and $f_3'(\hat{F})$ are already disjoint. If $f_3'(F)$ and $f_3'(\hat{F})$ intersect, then condition (vi) of Lemma 9.5.13 implies that the order of $F \cap T$ and $\hat{F} \cap T$ is preserved under $f_3'|T$. Therefore, by considering $\hat{\mathcal{F}} \cup F$ instead of $\hat{\mathcal{F}}$, we see that the least area surfaces are disjoint.

It remains to show that we can take disjoint regular neighbourhoods of the components. (Since the restriction of $H(\cdot\,, 3/4)$ to each component of $\mathcal{F}$ is uniformly bi-Lipschitz, the uniform bi-Lipschitzness on $\mathcal{F} \times [1/2, 1]$ follows immediately once we prove that we can take regular neighbourhoods to be disjoint. Combined with the fact shown above that a homotopy between f_3' and $H(\cdot\,, 3/4)$ can be made uniformly Lipschitz, the uniform Lipschitzness of H also follows.) Recall that by Lemma 9.5.13-(vii), we can assume that there is a uniform positive lower bound for the distances between the images of distinct boundary components of $\mathcal{F}$ under f_3', hence also under $H(\cdot\,, 3/4)$. To get disjoint regular neighbourhoods, what we need is a lower bound for the distances between the images of distinct components of $\mathcal{F}$ under $H(\cdot\,, 3/4)$, not only for their boundaries but for the entire surfaces. Suppose that such a lower bound does not exist. Then by taking a geometric limit, we get two least-area surfaces which are tangent to each other at their interiors. This contradicts the maximal principle of minimal surfaces. Thus, we have shown that there is a lower bound, and that we can take disjoint regular neighbourhoods. □

9.5.6 Topological Ordering of Joints

Next we shall show that the obtained embedding $H(\cdot\,, 1) : \mathcal{F} \to N[k_u]$ preserves the topological order of joints.

Lemma 9.5.15 *Let F_1 and F_2 be joints in $\mathcal{F}$ such that $\iota_M(F_1)$ and $\iota_M(F_2)$ are not homotopic in $S \times (0, 1)$. If $F_1 \prec_{\text{top}} F_2$, then $H(F_1, 1) \prec_{\text{top}} H(F_2, 1)$.*

Proof Suppose that $F_1 \prec_{\text{top}} F_2$, and that $\iota_M(F_1)$ is not homotopic to $\iota_M(F_2)$ in $S \times (0, 1)$. Let c be a boundary component of F_2 which overlaps F_1 if there are any. There is a component T of $\partial M[k_u]$ on which c lies. Then, Lemma 3.3 in Brock–Canary–Minsky [12] implies that $F_1 \prec_{\text{top}} c$. Recall from Proposition 9.5.9 that

f_4 extends to a homeomorphism $\hat{f}_4 : M \to N_0$ properly homotopic to f. Since $F_1 \prec_{\text{top}} c$, the surface $\iota_M(F_1)$ can be homotoped to $S \times \{0\}$ without touching $\iota_M(c)$. Since $\iota_M = \iota_N \circ f$, we see that $\iota_N \circ f(F_1)$ can be homotoped to $S \times \{1\}$ without touching $\iota_N \circ f(c)$, which by Lemma 3.18 in [12] implies $f(F_1) \prec_{\text{top}} f(c)$. Because f is properly homotopic to $\hat{f}_4$, we also have $f_4(F_1) \prec_{\text{top}} f_4(c)$. Since c lies on a component T of $\partial M[k_u]$, the homeomorphism f_4 is homotopic to f_3' in $N[k_u]$, and H is a proper homotopy in $N[k_u]$, this topological order is preserved by $H(\cdot, 1)$, and we have $H(F_1, 1) \prec_{\text{top}} H(c, 1)$. By the same argument and changing the direction of order, we see that for any boundary component c' of F_1 that overlaps F_2, we have $H(c', 1) \prec_{\text{top}} H(F_2, 1)$. Since F_1 and F_2 are assumed not to be homotopic, by Lemma 3.17 in [12], this implies that $H(F_1, 1) \prec_{\text{top}} H(F_2, 1)$. □

We next consider the case when two joints F_1 and F_2 have homotopic images under ι_M.

Lemma 9.5.16 *Suppose that F_1 and F_2 are joints in $\mathcal{F}$ such that $\iota_M(F_1)$ is homotopic to $\iota_M(F_2)$. We further assume that $F_1 \cup F_2$ does not bound a brick in $M[k_u]$. If $F_1 \prec_{\text{top}} F_2$, then we have $H(F_1, 1) \prec_{\text{top}} H(F_2, 1)$.*

Proof Since $\iota_M(F_1)$ is homotopic to $\iota_M(F_1)$, for each component c of ∂F_1, there is a unique component c' of ∂F_2 such that $\iota_M(c)$ is freely homotopic to $\iota_M(c')$ in $S \times (0, 1)$. Suppose first that c and c' are homotopic in $M[k_u]$ for all components c of ∂F_1. Then c and c' lie on the same boundary component of $\partial M[k_u]$ by condition A-(2) in Sect. 9.4.2 and the definition of $\mathcal{V}[k_u]$. Since this holds for every boundary component of F_1, we see that $F_1 \cup F_2$ bounds a submanifold W in $M[k_u]$. If W is homeomorphic to $F_1 \times [0, 1]$, then by our definition of the brick decomposition of $M[k_u]$, we see that W consists of only one brick. This contradicts our assumption that $F_1 \cup F_2$ does not bound a brick.

Therefore, $F_1 \cup F_2$ bounds a submanifold W in $M[k_u]$, which is not homeomorphic to $F_1 \times [0, 1]$. Since $\iota_M(F_1 \cup F_2)$ bounds a region homeomorphic to $F_1 \times [0, 1]$ in $S \times [0, 1]$, this implies that there is a component T of $\partial M[k_u]$ which is contained in W. We take a horizontal curve d contained in T. Then d overlaps both F_1 and F_2 and $F_1 \prec_{\text{top}} d \prec_{\text{top}} F_2$. This implies that $f_4(F_1) \prec_{\text{top}} f_4(d) \prec_{\text{top}} f_4(F_2)$. Since $\iota_N \circ \hat{f}_4$ is homotopic to ι_M, we see that $\iota_N \circ f_4(F_1)$ is homotopic to $\iota_N \circ f_4(F_2)$, and the boundary of $\iota_N \circ f_4(F_1)$, which lies on $\mathcal{V}[k_u]$, is unknotted and unlinked. Therefore, applying Lemma 3.16 in [12], we have $f_4(F_1) \prec_{\text{top}} f_4(F_2)$, which implies that $H(F_1, 1) \prec_{\text{top}} H(F_2, 1)$ as before.

Thus it only remains to consider the case when there is a component c of ∂F_1 which is not homotopic to c' in $M[k_u]$. Since $\iota_M(c)$ and $\iota_M(c')$ are homotopic, and $\iota_M(c)$ and $\iota_M(c')$ are horizontal, there is an embedded annulus A bounded by $\iota(c) \cup \iota(c')$ in $S \times (0, 1)$. Since c and c' are not homotopic in $M[k_u]$, there is a boundary component T of $M[k_u]$ such that $\iota_M(T)$ intersects A essentially. Take a longitude or a core curve c'' of T. Then we have $F_1 \prec_{\text{top}} c'' \prec_{\text{top}} F_2$. Now since f_4 is a homeomorphism from M to N_0, we see that $f_4(F_1) \prec_{\text{top}} f_4(c'') \prec_{\text{top}} f_4(F_2)$, and as before, we have $H(F_1, 1) \prec_{\text{top}} H(c'', 1) \prec_{\text{top}} H(F_2, 1)$. Since $\iota_N \circ H(F_1, 1)$ and

$\iota_N \circ H(F_2, 1)$ are homotopic, Lemma 3.16 in [12] again implies that $H(F_1, 1) \prec_{\text{top}} H(F_2, 1)$. Thus we have completed the proof. □

The remaining case is when F_1 and F_2 are homotopic in $M[k_u]$ and cobound a brick in $M[k_u]$. Let B be a brick bounded by $F_1 \cup F_2$, and $h(B)$ the hierarchy on B which we obtain by applying Lemma 9.5.5 to $M[k_u]$. We now introduce the term a "deep-seated geodesic" for a tight geodesic in the hierarchy $h(B)$.

Definition 9.5.17 We say that a tight geodesic $g \in h(B)$ is *deep-seated* if there is a component of $\mathrm{Fr}D(g)$ whose corresponding tube in $\mathcal{V}$ is disjoint from either $\partial_+ B$ or $\partial_- B$. In the case when $D(g)$ is an annulus, we regard a core curve of $D(g)$ as a component of $\mathrm{Fr}D(g)$.

We shall first show that $h(B)$ cannot have a long deep-seated geodesic.

Lemma 9.5.18 *Let B be a brick in $M[k_u]_{\text{int}}$. Then every deep-seated geodesic in $h(B)$ has length less than A, where A is a constant depending only on $\xi(S)$.*

Proof By Theorem 9.1 in [35], we can take a constant A such that if $g \in h(B)$ has length at least A, then every component c of $\mathrm{Fr}D(g)$ lies on either ∂M or a boundary component ∂V for $V \in \mathcal{V}$ such that $|\omega_M(V)| > k_u$. (Since we are considering tight geodesics in $h(B)$ consisting of simplices on the curve complex of $\mathcal{C}(\partial_- B)$, we can apply Minsky's result on Kleinian surface groups.) Therefore in this situation, every component of $\mathrm{Fr}D(g)$ lies on $\partial M[k_u]$. If g is deep-seated, then some component c of $\mathrm{Fr}D(g)$ must lie on ∂V which is disjoint either from $\partial_+ B$ or $\partial_- B$. Thus we see that if $h(B)$ has a deep-seated geodesic with length at least A, then there is a component of $\partial M[k_u]$ which intersects B but not at least one of $\partial_- B$ and $\partial_+ B$. This contradicts the assumption that $\partial_- B$ and $\partial_+ B$ are homotopic and bound B in $M[k_u]$. □

Fix A as in Lemma 9.5.18 so that all the deep-seated geodesics in $h(B)$ have length less than A. We further divide our argument into two cases: the first is when the number of blocks constituting B is large and the other is when it is small. It will turn out later that we do not need to show that the topological order is preserved in the latter case for the proof of Theorem 9.5.1.

Lemma 9.5.19 *There exists a constant C depending only on $\xi(S)$ (and A) such that the following holds. If $|h(B)| > C$, then we have $H(\partial_- B, 1) \prec_{\text{top}} H(\partial_+ B, 1)$. (Recall that $|h(B)|$ denotes the sum of the lengths of all the geodesics constituting $h(B)$.)*

Proof We can take a constant C so that if $|h(B)| > C$, then there must be a geodesic g in $h(B)$ whose length is greater than A. By Lemma 9.5.18, then g cannot be deep-seated. If g is not deep-seated, then since every frontier component of $D(g)$ lies in $\partial M[k_u]$, the only possibility is that g is the main geodesic of $h(B)$. Then we can apply Theorem 7.1 in [12] to our hierarchy $h(B)$. The same argument as in the case 1b of the proof of Lemma 8.4 in [12] implies that $H(\partial_- B, 1) \prec_{\text{top}} H(\partial_+ B, 1)$. □

The remaining case when $|h(B) \le C$ is included in the "short" case, which will be defined in the following subsection.

9.5.7 Deformation to a Bi-Lipschitz Map

Having obtained the results in the previous subsection, we are now in a position to show that we can further homotope $H(\cdot\ , 1)$ to make it bi-Lipschitz on the region between joints, applying the arguments of §§8.2-8.4 in Brock–Canary–Minsky [12].

For a brick B in $M[k_u]_{\text{int}}$ which is not short, we shall construct a *cut system*, following §4 and §8.2 in Brock–Canary–Minsky [12]. Our cut system C_B is a set of slices of $h(B)$ having the following properties with a constant $d_1 > 5$ depending only on $\xi(S)$ which will be specified later.

(1) For a geodesic $g \in h(B)$, let $C_B|g$ denote the subset of C_B consisting of slices with bottom geodesic g. Then, for any geodesic $g \in h(B)$, the bottom simplices $\{v_\tau \mid \tau \in C_B|g\}$ cut g into intervals all of whose lengths are between d_1 and $3d_1$.
(2) Two distinct slices in $C_B|g$ cannot have the same bottom simplex.
(3) For each $\tau \in C_B$ and any (k, v) in τ other than the bottom one, v is the first vertex of k.
(4) For every non-annular geodesic g, any slice in $C_B|g$ is a non-annular saturated slice.
(5) For an annular geodesic g, there is at most one slice in $C_B|g$.

We take a constant $d_1 \geq C$ for C in Lemma 9.5.19 so that for any geodesic g in the hierarchy $h(B)$, if g has length at least d_1, the geodesic length of each component of $f_3(\partial D(g))$ in N is less than ϵ_u. (By Lemmas 9.5.4 and 9.5.19, such a d_1 exists and depends only on $\xi(S)$.) We say that the brick B is *short* if $|h(B)| \leq 4^{\xi(S)} d_1$. (Here we choose $4^{\xi(S)}$ as a function of $\xi(S)$ bounding the number of geodesics constituting a hierarchy.) If B is not short, then at least one of the geodesics in $h(B)$ has length greater than d_1. Therefore, C_B is non-empty in this case. When B is short, we define C_B to be an empty set.

For each slice τ in C_B, we define *extended split level surfaces* as below following [12]. Suppose that the bottom pair (g_τ, v_τ) of τ is not supported in an annulus. Since τ is a non-annular saturated slice and $h(B)$ is 4-complete, **base**(τ) defines a pants decomposition of $D(g_\tau)$. For each pair of pants Y constituting the pants decomposition, there is a horizontal boundary of two adjacent blocks à la Minsky of the form $Y \times \{t\}$ with respect to the parametrisation of $S \times (0, 1)$, along which the two are glued. (This lies at the middle of a block of the form $\Sigma_{0,3} \times J$ in our block decomposition in Sect. 9.4.) This horizontal surface is denoted by F_Y. We consider the union $F_\tau = \cup F_Y$ for all Y constituting the pants decomposition, and call it the split level surface corresponding to τ. By joining two parallel boundaries of F_τ using annuli in solid tori of $\mathcal{V}[0] \setminus \mathcal{V}[k_u]$, we get a surface homeomorphic to $D(g_\tau)$. By condition (3), such annuli can be chosen to be uniformly close to horizontal annuli. We call this surface the extended split level surface corresponding to τ and denote it by $\mathcal{F}_\tau$. For a cut system C_B as above, the split level surface $\mathcal{F}_\tau$ for $\tau \in C_B$ is called a *cut* in B. Let $\mathcal{F}_B$ be the union of $\mathcal{F}_\tau$ for all $\tau \in C_B$ and $\mathcal{F}_b$ the union of $\mathcal{F}_B$ for all bricks $B \in \mathcal{K}[k_u]_{\text{int}}$. Let V be a component of $\mathcal{V}$ on which a boundary component of $\mathcal{F}_\tau$ lies. By condition (1) of the definitions of C_B and d_1, we see that $\omega(V) > k_u$, hence $V \in \mathcal{V}[k_u]$. Therefore, by adding $\mathcal{F}_b$ to the joints of $M[k_u]$, we get a subdivision of $M[k_u]$ into smaller bricks, which may have inessential joints. We denote this refined brick manifold by $M'[k_u]$. (Note that $M[k_u]$ and $M'[k_u]$ are the same as manifolds; only their brick structures differ.)

We shall show that $H(\cdot\,, 1)$ can be homotoped so that the restriction to $\mathcal{F}_b \cup \mathcal{F}$ is a uniform bi-Lipschitz embedding.

Lemma 9.5.20 *There exist a constant K'' depending only on $\xi(S)$ and a K''-Lipschitz homotopy $H' : (\mathcal{F}_b \cup \mathcal{F}) \times [0, 1] \to N[k_u]$, such that*

(i) $H'|(\mathcal{F}_b \cup \mathcal{F}) \times \{0\}$ *coincides with* $H(\mathcal{F}_b \cup \mathcal{F}, 1)$.
(ii) $H'|(\mathcal{F}_b \cup \mathcal{F}) \times [1/2, 1]$ *is a K''-bi-Lipschitz C^2-embedding.*

Proof Our argument is similar to the proof of Lemma 9.5.14. Let T be a component of $\partial M[k_u]$ intersecting B and T' its image in $N[k_u]$ under f_4. We first need to show that the $H(\cdot\,, 1)$ can be moved to a uniformly Lipschitz map which preserves the order of $T \cap (\mathcal{F}_b \cup \mathcal{F})$ except for the fronts of short bricks, by a homotopy whose displacement function is bounded from above by a uniform constant. Our situation is a little different from that of Lemma 9.5.13 since among our surfaces in $\mathcal{F}_b \cup \mathcal{F}$, there might be more than two components which are all homotopic to each other. Still, we can argue as in the proof of (vi) in Lemma 9.5.13, and see that the order of components F, F' of $\mathcal{F}_b \cup \mathcal{F}$ can be reversed only when they are homotopic in $M[k_u]$ and $H(F, 1) \cap H(F', 1) = \emptyset$. Now, by applying Lemma 9.5.19 to our refined brick manifold $M'[k_u]$, we see that the order between $F \cap T$ and $F' \cap T$ is preserved by $H(\cdot\,, 1)$ unless $F \cup F'$ bounds a brick B with $|h(B)| < C$ in $M'[k_u]$, which must be also a short brick of $M[k_u]$ since we did not introduce a new brick B with $|h(B)| < d_1$ in our subdivision of bricks. Thus we have shown that $H(\cdot\,, 1)$ can be homotoped with uniformly bounded displacement of points so that the order of $(\mathcal{F}_b \cup \mathcal{F}) \cap T$ is preserved for any component T of $\partial M[k_u]$ except for the order between the two fronts of the same short bricks. Let f_3'' be a uniform Lipschitz map from $M'[k_u]$ to $N[k_u]$, thus obtained.

Next we shall show that the same property as (vii) in Lemma 9.5.13 holds for $\mathcal{F}_b$ and f_3''; that is, for any δ, there is a number n_0 bounding the number of components of $f_3''(\mathcal{F}_b \cap T)$ which are within distance δ from $f_3''(F \cap T)$ for any component F of $\mathcal{F}_b$. Let $F_1, \ldots, F_n$ be distinct components of $\mathcal{F}_b$ such that $f_3''(F_1 \cap T), \ldots, f_3''(F_n \cap T)$ are within distance δ from $f_3''(F \cap T)$. Then $H(F_1, 1), \ldots, H(F_n, 1)$ are within distance $3\lambda_0 + \delta$ from $H(F, 1)$, where λ_0 is the constant which we defined in the proof of Lemma 9.5.13. Recall that for each slice τ of C_B, each component of $F_\tau \setminus \mathcal{V}$ is a thrice-punctured sphere. By Lemma 9.5.6, for distinct slices $F_{\tau_1}, \ldots, F_{\tau_n}$, there are at least ν non-homotopic tubes in $\mathcal{V}$ which at least one of $F_{\tau_1}, \ldots, F_{\tau_n}$ intersects with ν going to ∞ as $n \to \infty$. By Lemma 9.5.3, each tube has a core curve with length less than L. Since $H(\cdot\,, 1)$ is uniformly Lipschitz, the lengths of the images of the core curves are universally bounded. Suppose that there is no universal bound for n. Then by the usual argument using a geometric limit of model maps, we are lead to a contradiction since for any hyperbolic 3-manifold, a constant R and a base point x in the manifold, there are only finitely many homotopy classes which are represented by a closed curve of length less than L contained in the R-ball centred at x. Thus we have shown that $f_3''|\mathcal{F}_B$ has the same property as (vii) in Lemma 9.5.13. Combining this with Lemma 9.5.13, we see that for any δ, there exists n_0 bounding the number of components of $f_3''((\mathcal{F}_b \cup \mathcal{F}) \cap T)$ within distance δ from $f_3''(F \cap T)$ for any component of F of $\mathcal{F}_b \cup \mathcal{F}$.

By the same argument as Lemma 9.5.14, we see that $H(\cdot\,, 1)$ can be homotoped to a uniform Lipschitz map which embeds $\mathcal{F}_b \cup \mathcal{F}$ in such a way that two distinct components have disjoint δ-regular neighbourhoods. □

By this homotopy H', we can homotope f_3' to a map f_5 which is a uniform bi-Lipschitz map on each component of $\mathcal{F}_b \cup \mathcal{F}$ and whose restriction to a small regular neighbourhood of $\mathcal{F}_b \cup \mathcal{F}$ is an embedding. Recall that by the results in Sect. 9.5.6 and Lemmas 9.5.18 and 9.5.19, f_5 preserves the topological order of $\mathcal{F}_b \cup \mathcal{F}$ except for the fronts of short brick. If B is short, then B consists of less than $4^{\xi(S)}d_1$ blocks, hence the diameter of B, which can be bounded by a linear function of the number of blocks, is bounded by a constant depending only on $\xi(S)$. Therefore, we can isotope $f_5(\partial_- B)$ into a regular neighbourhood of $f_5(\partial_+ B)$ so that $f_5(\partial_- B) \prec_{\mathrm{top}} f_5(\partial_+ B)$ preserving the condition on the bi-Lipschitzness. We should note that short bricks of $M'[k_u]$ come from those of $M[k_u]$ and that by Assumption 9.4.3, two short bricks cannot be adjacent to each other. Therefore, we can perform this deformation for all short bricks so that $f_5(\partial_- B)$ and $f_5(\partial_+ B)$ have regular neighbourhoods with uniform width. Since the embedding of each cut by f_5 has a regular neighbourhood with uniform width, f_5 is bi-Lipschitz not only on each cut or joint but also with respect to the induced metric on the entire $\mathcal{F}_b \cup \mathcal{F}$.

To complete the proof of Theorem 9.5.1, it remains to deform f_5 in the complement of $\mathcal{F}_b \cup \mathcal{F}$ in $M[k_u]_{\mathrm{int}}$ to make it bi-Lipschitz without changing the map on $F \times [0, \infty)$ for every geometrically finite brick, which we parametrise as $F \times [-1, \infty)$ as before. This can be done by the same argument as §8.4 in [12] without any modification. Thus we have completed the proof of Theorem 9.5.1 by setting k to be k_u.

9.6 Proofs of Theorems

9.6.1 Geometric Limits of Geometrically Finite Bricks

Let G_n be a Kleinian surface group, and set N_n to be $\mathbb{H}^3/G_n$. Let $g_n : M_n \to (N_n)_0$ be a model map constructed in [12] which induces a bi-Lipschitz homeomorphism $g_n[k_u] : M_n[k_u] \to N_n[k_u]$. Suppose that M_n has a geometrically finite brick $B_n \cong F_n \times [-1, \infty)$ or $F_n \times (-\infty, 1]$. We note that for each geometrically finite end, Minsky [35] constructed a boundary block which is mapped into the augmented convex core and an exterior block for the exterior component of the augmented convex core containing the end. We have explained the details of this in

Sect. 9.5.1. Our geometrically finite brick is obtained as the union of these two. We shall consider only the case when $B_n \cong F_n \times [-1, \infty)$, for the other case can be dealt with in the same way just by changing the $+$-direction to the $-$-direction.

Lemma 9.6.1 *Let x_n be a point in B_n in the above situation. Suppose that with respect to the metric d_{B_n} on B_n defined in Sect. 9.4.5, we have $d_{B_n}(F_n \times \{-1\}, x_n) \to \infty$. Then the geometric limit of (a subsequence of) $\{(N_n, g_n(x_n))\}$ is elementary: i.e. isomorphic to $\mathbb{H}^3/\Gamma$ for an elementary Kleinian group Γ.*

Proof Let $C(N_n)$ be the convex core of $N_n = \mathbb{H}^3/G_n$. By the definition of our model maps, we see that $d_{N_n}(C(N_n), g_n(x_n)) \to \infty$. Let Γ be a Kleinian group such that $(\mathbb{H}^3/\Gamma, x_\infty)$ is the Gromov–Hausdorff limit of $\{(N_n, g_n(x_n))\}$ after passing to a subsequence. Suppose, seeking a contradiction, that there are non-commuting elements g, h in Γ. Then, there exist elements g_n, h_n in G_n such that $\lim g_n = g$ and $\lim h_n = h$. Consider the action of G_n on $\mathbb{H}^3$. Then g_n and h_n act on $\mathbb{H}^3$ as loxodromic or parabolic transformations. Let l_n be a geodesic in $\mathbb{H}^3$ which we set to be the common perpendicular of the axes of g_n and h_n if they are both loxodromic, or the geodesic ray perpendicular to the axis of the loxodromic one and tending to the fixed point at infinity of the parabolic one when only one of them is loxodromic, or the geodesic connecting the fixed points at infinity of the two elements if both are parabolic.

We claim that the function $t(g_n, h_n)(x) = \max\{d(x, g_n(x)), d(x, h_n(x))\}$ has a minimum at a point c_n on l_n. This can be seen by considering sets $V_{g_n}(r)$ and $V_{h_n}(r)$ consisting of points whose translation distances are less than or equal to r under g_n and h_n respectively. The smallest r for which $V_{g_n}(r) \cap V_{h_n}(r) \neq \emptyset$ realises the minimum of $t(g_n, h_n)$. (If $V_{g_n}(r)$ (resp. $V_{h_n}(r)$) reaches the axis of h_n (resp. g_n) while $V_{h_n}(r)$ (resp. $V_{g_n}(r)$) is empty, we take such r as the smallest.) By the convexity of these sets, we see that the intersection consists of one point c_n, and that it lies on l_n. Since both $\{g_n\}$ and $\{h_n\}$ converge, the smallest r is bounded from above independently of n. Since the configurations of $V_{g_n}(r)$, $V_{h_n}(r)$ up to isometries are compact, we see that $|t(g_n, h_n)(y) - 2d(y, c_n)|$ is bounded from above independently of n. (This follows from the fact that the displacement of a point can be approximated by twice the distance from the point to the axis if the translation length on the axis is bounded above.)

Obviously, l_n is contained in the Nielsen convex hull of G_n, i.e. the convex hull in $\mathbb{H}^3$ of the limit set Λ_{G_n}. Take a lift $\tilde{x}_n$ of $g_n(x_n)$ which converges to a lift $\tilde{x}_\infty$ of x_∞. Since $d_{N_n}(C(N_n), g_n(x_n)) \to \infty$, the distance of l_n from a lift $\tilde{x}_n$ of $g_n(x_n)$ in $\mathbb{H}^3$ goes to ∞ as $n \to \infty$; hence $d(\tilde{x}_n, c_n) \to \infty$. From the above observation, this implies that $t(g_n, h_n)(\tilde{x}_n) \to \infty$. This contradicts the facts that $g = \lim g_n$ and $h = \lim h_n$ translate $\tilde{x}_\infty$ by a finite distance. □

We now consider geometric limits of geometrically finite bricks B_m parametrised as $F_m \times [-1, \infty)$ as in Sect. 9.4.5. Since we are only interested in non-elementary geometric limits, by virtue of the previous lemma, we have only to consider the case when the basepoint lies on the real front along which the brick is pasted to other bricks. Let x_m be a point in B_m lying on $F_m \times \{-1\}$. Since each $\hat{F}_m = F_m \times$

$\{-1\}$ has the cylinder-$\Sigma_{0,3}$ metric υ_m, the sequence $\{(F_m \times \{-1\}, x_m)\}$ converges geometrically to a pointed surface $(\hat{F}_\infty, x_\infty)$ passing to a subsequence, and $\hat{F}_\infty$ also has a cylinder-$\Sigma_{0,3}$ metric υ_∞ and is homeomorphic to a subsurface F_∞ of S. On the other hand, $F_m \times \{0\}$ has a metric τ_m which is Euclidean in annular neighbourhoods of pants curves, and is hyperbolic outside. If we put a basepoint y_m on $F_m \times \{0\}$ whose distance from x_m is bounded as $m \to \infty$, the sequence $\{(F_m \times \{0\}, y_m)\}$ has a geometric limit $(\bar{F}_\infty, y_\infty)$ (passing to a subsequence) with $\bar{F}_\infty$ homeomorphic to F_∞, and having the same kind of metric as $F_m \times \{0\}$. Since $F_m \times [0, 1]$ has a metric which is bi-Lipschiz equivalent to the product of τ_m and the standard interval, $\{(F_m \times [0, 1], x_m)\}$ has a geometric limit homeomorphic to $F_\infty \times [0, 1]$ bounded by $\bar{F}_\infty$, regarded as $F_m \times \{0\}$, and $\hat{F}_\infty$, regarded as $F_\infty \times \{-1\}$. Moreover, since $F_m \times [0, \infty)$ has metric $e^{2r}\tau_m + dr^2$ $(r \in [0, \infty))$, this part converges to a manifold homeomorphic to $F_\infty \times [0, \infty)$ with the metric $e^{2r}\tau_\infty + dr^2$. Thus, we have seen that $\{B_m\}$ converges to a geometrically finite brick B_∞ homeomorphic to $F_\infty \times [-1, \infty)$ which has the metric defined in Sect. 9.4.5.

9.6.2 Proofs of Theorem A and Corollary B

Proof of Theorem A Let $\{G_n\}$ be a sequence of Kleinian surface groups which converges geometrically to a non-elementary Kleinian group G. Since $\{G_n\}$ converges geometrically to G, fixing a basepoint in $\mathbb{H}^3$, and projecting it to $\mathbb{H}^3/G_n$ and $\mathbb{H}^3/G$ as basepoints y_n and y_∞, we get a geometric convergence $(\mathbb{H}^3/G_n, y_n) \to (\mathbb{H}^3/G, y_\infty)$. By the original bi-Lipschitz model theorem [12], for each $n \in \mathbb{N}$, there exist a model manifold M_n and a model map $g_n : M_n \to (N_n)_0$ inducing a K-bi-Lipschitz homeomorphism $g_n : M_n[k_u] \to N_n[k_u]$, where $N_n = \mathbb{H}^3/G_n$. We let x_n be a point in M_n which is taken to y_n by g_n. The model manifold M_n consists of $M_n[0]$, which is decomposed into internal blocks and boundary blocks, and Margulis tubes. Since we assumed that G is non-elementary, x_n cannot go deeper and deeper into Margulis tubes as $n \to \infty$. Therefore, moving x_n and y_n within uniformly bounded distance without changing G up to conjugacy, we can assume that x_n lies in $M_n[0]$.

Since G_n is a Kleinian surface group, M_n is properly embedded in $S_0 \times (0, 1)$ for a compact core S_0 of S so that the boundary of a cusp neighbourhood which does not correspond to a boundary component of S_0 is a properly embedded open annulus both of whose ends go to the same end of $S_0 \times (0, 1)$, either to the $+$-direction or to the $-$-direction. We equip M_n with the structure of a brick manifold compatible with the block decomposition as follows. We first consider a proper embedding $\eta_n : M_n \to S \times (0, 1)$ with the following properties, obtained from the above embedding by isotoping blocks within $S \times (0, 1)$.

(1) The embedding η_n preserves the horizontal and the vertical leaves of each block. (Here for a block of the form $\Sigma \times J$, each $\Sigma \times \{t\}$ is a horizontal leaf and $\{x\} \times J$ is a vertical leaf.)

(2) Each Margulis tube in M_n is mapped to $A \times [t_1, t_2]$ for some essential annulus A on S and $t_1 < t_2$, and each torus boundary of M_n is mapped to the boundary of $A \times [t_1, t_2]$.
(3) Each open annulus boundary component of M_n except for those corresponding to cusps of S is mapped to the boundary of either $A \times [t, 1)$ or $A \times (0, t]$ for an essential annulus A on S and $t \in (0, 1)$.
(4) The geometrically finite ends of M_n are peripheral, i.e. lie on $S \times \{0, 1\}$.

This is exactly the situation as in the construction of a brick decomposition for $M^{(1)}_{\mathrm{int}}$ in Sect. 9.4.4. Therefore, we can endow M_n with a brick decomposition by defining each to be a maximal family of parallel leaves.

We now consider the geometric limit $(M[0], x_\infty)$ of $(M_n[0], x_n)$, possibly passing to a subsequence. Note that any internal block of $M_n[0]$ is isometric to either $\Sigma_{(0,4)} \times [0, 1]$ or $\Sigma_{(1,1)} \times [0, 1]$, or $\Sigma_{(0,3)} \times [0, 1]$, each with a standard metric. (We can consider a block decomposition in our sense or Minsky's. Either will do.) Therefore a geometric limit of internal blocks can also be regarded as blocks. On the other hand, as was seen in Sect. 9.6.1, any sequence of geometrically finite bricks in $M_n[0]$ converges geometrically to a geometrically finite brick in after taking a subsequence if we put a basepoint on the real front. Since G is non-elementary, by Lemma 9.6.1, if the x_n lie in geometrically finite bricks, we can assume that they lie on the real fronts of the bricks. These imply that the geometric limit $M[0]$ consists of geometrically finite bricks and the remaining part which is decomposed into blocks. (Here we are not considering yet the brick decomposition of $M[0]$.)

We denote by $M[0]_{\mathrm{int}}$ the part of $M[0]$ consisting of the limits of internal bricks. The complement of $M[0]_{\mathrm{int}}$ in $M[0]$ consists of geometrically finite bricks as was seen above. As before, we denote by $\mathcal{V}_n$ the union of tubes in the tight tube unions giving a block decomposition of M^0_n. (Recall that M^0_n is the complement of tubes in $\mathcal{V}_n$ intersecting M_n along annuli and is naturally identified with M_n.) For any k, we denote by $\mathcal{V}_n[k]$ the subset of $\mathcal{V}_n$ consisting of tubes V with $|\omega_{M_n}(V)| \geq k$. Recall that $M_n[k] = (M_n)^0 \setminus \mathcal{V}_n[k]$. We denote by $T_n[k]$ the union of Margulis tubes which is the image of $\mathcal{V}_n[k]$ by g_n. (Recall that we abuse the term "Margulis tube" to refer also to a tubular neighbourhood of a closed geodesic with uniformly bounded length even when it is greater than the Margulis constant.)

Each torus component T of $\partial M[0]$ is a geometric limit of torus components T_n of $\partial M_n[0]$. Since T_n converges geometrically, either $\{\omega_{M_n}(T_n)\}$ converges or goes to ∞. (Note that this excludes the case when $|\omega_{M_n}(T_n)|$ goes to ∞ but $\omega_{M_n}(T_n)$ does not.) If it converges, then T_n bounds a Margulis tube V_n converging geometrically to a Margulis tube V bounded by T. We denote by $\mathcal{V}_\infty$ the union of such tubes V. The gluing map of V_n to $M_n[0]$ converges to a gluing map of V to $M[0]$. We define the union of $M_n[0]$ and such tubes glued by the limit gluing maps to be M. Then it follows immediately that the geometric convergence of $(M_n[0], x_n)$ to $(M[0], x_\infty)$ extends that of (M_n, x_n) to (M, ∞). We denote by $M[k]$ the union of $M[0]$ and tubes in $\mathcal{V}_\infty$ for which $\lim_{n\to\infty} |\omega_{M_n}(T_n)| \leq k$. The argument above also implies in particular that g_n with base point x_n converges to a K-bi-Lipschitz homeomorphism $g : M[k_u] \to N[k_u]$. Since we put the metric on each $V_n \in \mathcal{V}_n[k_u]$ inherited from a Margulis tube determined by $\omega_{M_n}(V_n)$ and k_u was taken so that $g_n(\partial V_n)$ bounds an unknotted Margulis tube in N_n, each g_n is extended to a K-bi-Lipschitz map in each V_n. Therefore g is also extended to a K-bi-Lipschitz homeomorphism from M to N_0. We use the symbol M_{int} to denote the union of $M[0]_{\text{int}}$ and $\mathcal{V}_\infty$.

If $\lim_{n\to\infty} \omega_{M_n}(T_n) = \infty$, then $g(T)$ is the boundary of a torus cusp neighbourhood of N in the complement of N_0. If we put a basepoint on $\partial V_n = T_n$, then the geometric limit of V_n is also a $\mathbb{Z} \times \mathbb{Z}$-cusp which is K-bi-Lipschitz to the cusp neighbourhood bounded by $g(T)$ since by Lemmas 9.4.5 and 9.4.6, $\omega_{M_n}(T_n)$ controls the modulus of the Margulis tube bounded by $g_n(T_n)$. Note that by our definition of M, the limit cusp neighbourhood is not contained in our model manifold M. Therefore, M is not exactly a geometric limit of $\{M_n\}$ but is obtained from the geometric limit by removing cusp neighbourhoods; but for simplicity, we shall often refer below to M as a geometric limit of $\{M_n\}$.

The properties (ii) that M is acylindrical and (i) that ∂M consists of tori and annuli in the statement of Theorem A are derived from the same properties for N_0. We shall next show that M is a brick manifold. Recall that $M[0]_{\text{int}}$ admits a decomposition into blocks. Let $\rho_n : B_{r_n}(M_n, x_n) \to B_{K_n r_n}(M, x_\infty)$ be a (K_n, r_n)-approximate isometry associated to the geometric convergence of $\{(M_n, x_n)\}$ to (M, x_∞). We can arrange ρ_n so that for each block b of $M[0]_{\text{int}}$, its pull-back $\rho_n^{-1}(b)$ is also a block with respect to the block decomposition of the brick manifold M_n, and $\rho_n^{-1}|b$ preserves the vertical and horizontal leaves of b. Recall that the embedding η_n of M_n into $S \times (0, 1)$ preserves the vertical and the horizontal leaves of blocks. Therefore, at each point of M the (two-dimensional) horizontal directions and the vertical direction are well defined. The horizontal directions in M constitute a foliation whose leaves are incompressible in M and homeomorphic to an essential subsurface of S (including S itself) as we can see by considering their image under ρ_n^{-1} for large n. We define a leaf of this foliation to be a horizontal leaf of M. Horizontal leaves are transversely oriented, by defining the $+$-direction of the second factor of $S \times (0, 1)$ to be the positive direction.

Now, we define a brick in M to be a closed submanifold which is the closure of a maximal union of parallel horizontal leaves in M if it has non-empty interior. It is evident that the bricks defined in this way are pairwise disjoint. We can further show the following, which implies that M is a brick manifold.

Lemma 9.6.2 *Every point in M is contained in a brick. The bricks are locally finite.*

Proof Let x be a point in M, and F a horizontal leaf of M on which x lies. We say that a boundary component T of M touches F from above if $T \cap F \neq \emptyset$ and if any leaf near F and above F intersects T whereas any leaf below F is disjoint from T. Similarly, we define touching from below. Every component of ∂M is either a torus or an open annulus, and they may intersect a horizontal leaf along annuli. Recall that an annulus component of ∂M_n contains only one horizontal annulus, and a torus component contains only two horizontal annuli situated at different horizontal levels. This property is preserved by taking geometric limit, and hence if a component of ∂M intersects a horizontal leaf along annuli, the intersection consists of a single annulus. Moreover, since M is acylindrical, there are no two annuli on

$\partial M \cap F$ which are parallel on F and are contained in distinct components of ∂M. Therefore, the number of the components of $F \cap \partial M$ is uniformly bounded by a constant depending only on $\xi(S)$.

Now, recall that the height (with respect to the metric determined by blocks) of each component of ∂M_n is uniformly bounded from below by a positive constant ζ. We take a positive number $\theta < \zeta$, and let F' be a horizontal leaf of M above F at distance θ with respect to the metric determined by blocks. Then each component of ∂M within distance θ from F which does not lie below F must intersect either F or F'. Therefore, the number of such components is bounded by a constant depending only on $\xi(S)$. The same holds for components of ∂M within distance θ from F not lying above F.

Let h_1 be the minimum of the heights above F (with respect to the metric on M determined by blocks) of the components of ∂M intersecting F but not touching from below, which we allow to be ∞, including the case when there are no such components. Since there are finitely many components of ∂M intersecting F as was shown above, we have $h_1 > 0$. Next let h_1' be the minimal distance from F to the components of ∂M lying above F, which is defined to be ∞ if there are no such components. By the observation in the previous paragraph, there are only finitely many components of ∂M within a fixed distance from F, and hence we have $h_1' > 0$. We set $\bar{h}_1$ to be $\min\{h_1, h_1'\}$. Then, if we move F in the vertical direction on the positive side within distance $\bar{h}_1$ to another horizontal surface, then the new surface may lose (from a parallel copy of F) the interior of annuli which are intersection with components of ∂M touching from above, but the surface cannot change in other ways. Therefore all the horizontal leaves above F within distance $\bar{h}_1$ are parallel to each other.j It follows that if x lies outside the intersection with components of ∂M touching F from above, then x is contained in a brick which passes through F or is situated above F and touches F at the boundary from above. Similarly by defining h_2 and h_2' with changing the $+$-direction to the $-$-direction and setting $\bar{h}_2$ to be $\min\{h_2, h_2'\}$, we see that all the horizontal surface below F within distance $\bar{h}_2$ are parallel to each other. Also, if x lies outside the intersection with components of ∂M touching F from below, then x is contained in a brick which passes through F or is situated below F and touches F at its boundary from below. Since no components of ∂M can touch F from both above and below, this shows that x is always contained in a brick.

Furthermore, there are only finitely many bricks at distance less than $\min\{\bar{h}_1, \bar{h}_2\}$ since F is contained in the (non-empty) union of finitely many bricks whose heights are at least $\min\{\bar{h}_1, \bar{h}_2\}$ and one of which contains x. This shows the local finiteness of the bricks. □

By our definition of bricks in M and in M_n, for any brick B in M its pull-back $\rho_n^{-1}(B)$ is contained in one brick in M_n for large n. Now, we are in a position to use Lemma 9.4.1 to verify condition (iv) in Theorem A. For any $r \in \mathbb{N}$, let $M(r)$ be the submanifold of M consisting of bricks intersecting the r-ball centred at x_∞ with respect to the metric induced from those on blocks. Then $M(r)$ contains only finitely many bricks by Lemma 9.6.2. If we take a sufficiently large n, then we can

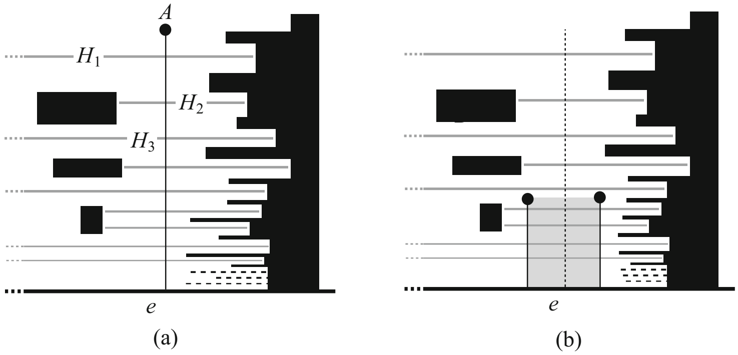


Fig. 9.7 *e* is a wild end

pull back $M(r)$ to M_n by ρ_n^{-1}. Since the pull-back of each brick is contained in a brick of M_n, we can embed $M(r)$ to $S \times (0, 1)$ by $\eta_n \circ \rho_n^{-1}$ preserving the vertical and the horizontal leaves. Since $M = \cup_{r=1}^{\infty} M(r)$, by Lemma 9.4.1, we can embed M into $S \times (0, 1)$ in such a way that every brick is mapped to a submanifold of the form $F \times J$. Since the geometrically finite ends of M_n are peripheral, we see that the same holds for M by Lemma 9.4.1. This completes the proof of (iv).

Finally, we shall show (iii), that there is no incompressible half-open annulus tending to a wild end e with core curve not homotopic to an annulus component of ∂M tending to e. Suppose, seeking a contradiction, that M has such an end e to which an incompressible half-open annulus A tends, and that the core curve of A is not homotopic into an annulus component of ∂M tending to e. Let $\{H_j\}$ be a sequence of properly embedded connected horizontal surfaces in M meeting A transversely and tending to e. (Since every point lies on some horizontal leaf, such a sequence of horizontal surfaces exist.) See Fig. 9.7a. For each j, the intersection $A \cap H_j$ is an essential simple closed curve, which we denote by l_j. By our assumption, l_j is not homotopic into an annulus component of ∂M tending to e. Therefore, $g(l_j)$ either represents a loxodromic element or is homotopic into a cusp which is disjoint from a small neighbourhood of e.

We first assume that $g(l_j)$ represents a loxodromic element. Let $h_j : H_j \to N_0$ be a pleated surface properly homotopic to $g|H_j$ realising l_j as a closed geodesic, which we denote by l^*. We should note that H_j is homeomorphic to an essential subsurface of S. For any $\delta > 0$, the pleated surfaces h_j have an upper bound depending only on $\chi(S)$ and δ for the diameters modulo their δ-thin parts. Since there are only finitely many ε_1-cusp neighbourhood within a bounded distance modulo the δ-thin part of N from l^* and the images of the h_j contain l^*, by taking a subsequence we can assume that the homotopy class of ∂H_j does not depend on j. By condition (ii) which we have already proved, this implies that the boundary components of M on which ∂H_j lies does not depend on j. It follows that there is an essential subsurface R of S such that all the H_j are vertically parallel to $R \times \{1/2\}$ in $S \times (0, 1)$. (Notice that they may not be parallel in M. To be more precise, we are claiming that the $\iota_M(H_j)$ are vertically parallel to $R \times \{1/2\}$ for the embedding ι_M of M into $S \times (0, 1)$ obtained above. We omit to write ι_M here.)

Let $i_j : R \to H_j$ be a homeomorphism compatible with a homotopy from $R \times \{1/2\}$ to H_j in $S \times (0, 1)$. Since the l_j are homotopic to each other in $S \times (0, 1)$, we can arrange the i_j so that there is a simple closed curve l on R such that $i_j(l) = l_j$ for all j. Recall that there are only finitely many ε_1-cusp neighbourhoods which pleated surfaces h_j can touch. We extend l to a pants decomposition P of R so that no curve is mapped to a curve freely homotopic into a cusp which some $h_j(H_j)$ touches. Since the condition that h_j is homotopic to $g|H_j$ and realises l guarantees the existence of such an extension, we can assume that h_j itself realises P. We now consider the sequence of pleated surfaces $\{h_j \circ i_j : R \to N_0\}$. Since there are only finitely many cusps which we must take into account, by applying the compactness of marked pleated surfaces without accidental parabolics (5.2.18 in Canary–Epstein–Green [15]), we see that passing to a subsequence, $\{h_j \circ i_j\}$ converges to a pleated surface from a component R' of $R \setminus \alpha$ containing l, where α is a possibly empty union of disjoint non-parallel essential annuli in R, uniformly on every compact subset of R'. It follows that there exists $j_0 \in \mathbb{N}$, such that all $h_j \circ i_j|R'$ $(j \geq j_0)$ are properly homotopic in N_0. Pulling back this to M, we see that there is no component of $S \times (0, 1) \setminus M$ which obstructs homotopies between the $i_j|R'$. Hence, the subsurfaces $i_j(R')$ of H_j are vertically parallel in M for all large j. Therefore, there exists a submanifold $R' \times [0, 1)$ embedded in M preserving the horizontal and vertical leaves, which contains a neighbourhood of the end of A such that $R' \times \{t\}$ tends to e as $t \to 1$. See Fig. 9.7b.

We shall next show that we have the same kind of product region even when $g(l_j)$ represents a parabolic class. Let c be a cusp of N homotopic to $g(l_j)$. Then, we consider a pleated surface $h_j : H_j \setminus \mathrm{Int}N(l_j) \to N_0$ taking $\partial N(l_j)$ to the boundary of the cusp neighbourhood of c instead of the one realising l_j as a closed geodesic, where $N(l_j)$ denotes an annular neighbourhood of l_j. Even in this case, we have the finiteness of pleated surfaces which can be reached from the δ-cusp neighbourhood U_c of c within a bounded distance modulo the thin part. Therefore, as before, we can show that the H_j are parallel in $S \times (0, 1)$ after taking a subsequence.

As before, we can consider a homeomorphism $i_j : R \to H_j$ compatible with the inclusion of R into S, and can assume that $h_j \circ i_j$ realises a pants decomposition P containing l none of whose curves except for l is mapped to a cusp which can be reached by $h_j(H_j)$. Then as in the previous case, there is a possibly empty union α of non-parallel disjoint essential annuli on R, and for components R_1, R_2 of $R \setminus (N(l) \cup \alpha)$ adjacent to $N(l)$, which may coincide if l is non-separating, the pleated surfaces $h_j \circ i_j|R_1 \cup R_2$ converge uniformly on every compact set of $R_1 \cup R_2$. Let R' be $R_1 \cup R_2 \cup N(l)$. Since the $h_j \circ i_j|R'$ are homotopic to each other for large j, we see that the subsurfaces $i_j(R')$ on H_j are vertically parallel to each other. This shows that there is a region $R' \times [0, 1)$ embedded in M preserving the horizontal and vertical leaves which contains a neighbourhood of the end of A such that $R' \times \{t\}$ tends to e as $t \to 1$.

In both cases, i.e. whether $g(l_j)$ represents a loxodromic class or a parabolic class, if every sequence of properly embedded connected horizontal surfaces tending to e is eventually contained in $R' \times [0, 1)$ defined above, then $R' \times [0, 1)$ constitutes a neighbourhood of e, contradicting the assumption that e is wild. Suppose that this is not the case. Then some component c of $\mathrm{Fr}R'$ is not homotopic to a core curve of an annulus component of ∂M tending to e. Therefore, we can repeat the above argument replacing A with $c \times [0, 1) \subset R' \times [0, 1)$ and get a larger subsurface R'' properly containing R' and a leaf-preserving embedding $R'' \times [0, 1)$ such that $R'' \times \{t\}$ tends to e as $t \to 1$. Since the topological type of S is fixed, this process terminates in finitely many steps, and we get a neighbourhood of e in the form $R_0 \times [0, 1)$ for some essential subsurface R_0 of S (which might be S itself) such that $\mathrm{Fr}R_0 \times [0, 1)$ lies on ∂M. By our definition of brick decomposition of M, this $R_0 \times [0, 1)$ is contained in one brick and e must be simply degenerate. This contradicts the assumption that e is wild. □

Proof of Corollary B By Theorem A, there is a brick manifold M having the properties listed in the theorem with a bi-Lipschitz homeomorphism $g : M \to N_0$. By Lemma 9.4.1, M has at most countably many ends; hence so does N_0. □

9.6.3 Proof of Theorem C

Proof of Theorem C Let M be a labelled brick manifold satisfying conditions (i)–(iv) in Theorem A with end invariants given so that condition (EL) in Sect. 9.4.4 is satisfied. Let $\mathcal{K}$ be a brick complex with $\bigvee \mathcal{K} = M$. By Sects. 9.4.2 and 9.4.4, M admits a decomposition into blocks. We use the symbols $\mathcal{V}$ and $\mathcal{V}[k]$ etc. to denote the unions of tubes inducing the decomposition into blocks as before. This implies that condition (BB) in Sect. 9.4.4 also holds. Since M is assumed to be embedded in $S \times (0, 1)$, we often identify M and its image in $S \times (0, 1)$.

For a simply degenerate brick $B = F \times [s, t)$ in $\mathcal{K}$, we consider a monotone increasing sequence $\{p_n\}$ of positive numbers tending to t such that, for any $n \in \mathbb{N}$, every component of $F \times \{p_n\} \setminus \mathrm{Int}\mathcal{V}$ is homeomorphic to $\Sigma_{0,3}$ and $B(p_n) = F \times [s, p_n]$ contains at least n components of $\mathcal{V}[0]$. We construct $B(p_n)$ in the same way when $B = F \times (t, s]$, just turning everything upside down. Let $\{\mathcal{K}_n\}$ be an ascending sequence of finite brick complexes with $\bigcup_n \mathcal{K}_n = \mathcal{K}$. We may choose such $\mathcal{K}_n$ so that $M_n = \bigvee \mathcal{K}_n$ is connected for any $n \in \mathbb{N}$. Since all geometrically finite bricks in $\mathcal{K}$ are peripheral in $S \times (0, 1)$, the their number is at most $-2\chi(S)$. Hence we can choose $\{\mathcal{K}_n\}$ so that $\mathcal{K}_1$ contains $\mathcal{K}_{\mathrm{gf}}$.

Consider a brick complex $\mathcal{K}_n^-$ obtained from $\mathcal{K}_n$ by replacing every simply degenerate bricks B in $\mathcal{K}_n$ with $B(p_n)$, and set $M_n^- = \bigvee \mathcal{K}_n^-$. For a simply degenerate brick B in $\mathcal{K}_n$ and for all $i \geq n$, the brick B is contained in $\mathcal{K}_i$ since $\{\mathcal{K}_i\}$ is ascending. Since $B = \bigcup_{i \geq n} B(p_i)$ by our definition of $B(p_n)$, we have $B \subset \bigcup_i M_i^-$. Therefore we see that $M = \bigcup_n M_n^-$.

We fix a base point x_0 in $M_1^- \cap M[0]$. Let $W_n[0]$ be the component of $M_n^- \cap M[0]$ containing x_0, and W_n the union of $W_n[0]$ and the components of $\mathcal{V}[0]$ whose boundaries lie on $\partial W_n[0]$. By the definition of W_n, we have $W_n \subset M_n^- \cap M^0$. For

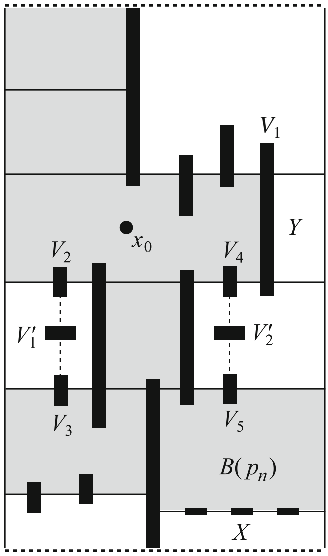


Fig. 9.8 This figure illustrates Z_n. The shaded region represents W_n and the union of black rectangles is $\mathcal{V}_n'$. $B(p_n) \in \mathcal{K}_n^-$ is contained in a simply degenerate brick B in $\mathcal{K}_n$ with $B = B(p_n) \cup X$. V_1 splits M_n^- into W_n and $Y = M_n^- \setminus W_n$. V_2 and V_3 (resp. V_4 and V_5) are components of $\mathcal{V}_n^{\mathrm{ext}}$ parallel to each other in $S \times (0, 1)$. V_1' (resp. V_2') obstructs an annulus between V_2 and V_3 (resp. V_4 and V_5)

any $n \in \mathbb{N}$, there exists $m \geq n$ such that every component of $\mathcal{V}[0]$ intersecting M_n^- is contained in the component of M_m^- containing x_0 since there are only finitely many components of $\mathcal{V}[0]$ intersecting M_n^-. This means in particular that $M_n^- \cap M^0$ is contained in W_m, and hence that $\bigcup_m W_m = \bigcup_n (M_n^- \cap M^0) = M^0$. Taking a subsequence if necessary, we may assume that W_1 contains all the bricks in $\mathcal{K}_{\mathrm{gf}}$.

Let $\mathcal{V}_n^{\mathrm{ext}}$ be the union of all the components of $\mathcal{V} \setminus \mathrm{Int} W_n$ intersecting ∂W_n. It should be noted that $\mathcal{V}_n^{\mathrm{ext}}$ might contain a component of $\mathcal{V} \setminus \mathcal{V}[0]$. By the definition of W_n, each component of $W_n \cap \mathcal{V}_n^{\mathrm{ext}}$ is an annulus. Since $M[0]$ is acylindrical, there is no essential annulus A in W_n with $\partial A \subset W_n \cap \mathcal{V}_n^{\mathrm{ext}}$. Still there might be an annulus A in $S \times (0, 1)$ with $\partial A \subset \mathcal{V}_n^{\mathrm{ext}}$. Figure 9.8 illustrates such a situation. By the acylindricity of $M[0]$, for such an annulus A, either there is a tube V_A in $\mathcal{V}$ obstructing A, or A goes out of M (i.e. A cannot be homotoped into M). In the latter case, A must go out from a simply degenerate end B by condition (iii) of Theorem A. Since the core curves of $\mathcal{V}$ converges to an ending lamination, which is filling, we see that also in this case there is a tube V_A in $\mathcal{V}$ obstructing A. Since there are only finitely many homotopy classes of such annuli, we can choose finitely many pairwise disjoint tubes $V_1', \ldots, V_m'$ of $\mathcal{V}$ in $S \times (0, 1) \setminus W_n$ which obstruct all such annuli. Then by setting $\mathcal{V}_n' = \mathcal{V}_n^{\mathrm{ext}} \cup V_1' \cup \cdots \cup V_m'$ and $Z_n = S \times (0, 1) \setminus \mathcal{V}_n'$ in $S \times (0, 1)$, and defining its brick decomposition using maximal families of parallel horizontal leaves as usual, we see that Z_n is an acylindrical finite brick manifold with a brick decomposition $\mathcal{L}_n$ which is an extension of $\mathcal{K}_n|_{W_n}$. (Figure 9.8 is an example of Z_n.) Note that Z_n is not necessarily a subset of Z_{n+1} although $W_n \subset W_{n+1}$.

Since W_n contains all the bricks of $\mathcal{K}_{\mathrm{gf}}$ and since they are peripheral, we have $\partial_\infty M \subset \partial_\infty Z_n$. Using the conformal structure given on $\partial_\infty M$, we regard Z_n as a labelled brick manifold. We can take tight tube unions so that their restrictions to W_n coincide with $\mathcal{V} \cap W_n$. As was shown in Sect. 9.4.4, these tubes induce a decomposition of $Z_n[0]$ into blocks. By condition (BB) in Sect. 9.4.4, the closure of each component of $\partial W_n \setminus \mathcal{V}_n^{\text{ext}}$ is homeomorphic to $\Sigma_{0,3}$. It follows that for any B in $\mathcal{L}_n$ with $\partial_\pm B \cap \mathcal{V}_n^{\text{ext}} \neq \emptyset$, each component of $\partial_\pm B \setminus \text{Int}\mathcal{V}_n^{\text{ext}}$ is homeomorphic to $\Sigma_{0,3}$. Therefore, this block decomposition of $Z_n[0]$ can be taken so that its restriction to W_n is equal to the original block decomposition on $W_n[0]$. As in Sect. 9.4.5, we define a model metric on $Z_n[0]$ using the blocks and the conformal structure on $\partial_\infty Z_n$, and the model metric on $Z_n[0]$ is extended to the one on Z_n as before so that each component of $Z_n \setminus Z_n[0]$ is a Margulis tube having a metric inherited from its realisation in a hyperbolic 3-manifold. Since $d_{Z_n}(x_0, Z_n \setminus W_n)$ goes to ∞ as $n \to \infty$ with respect to the model metric d_{Z_n} on Z_n, the geometric limit of $\{Z_n\}$ is equal to the geometric limit M^0 of $\{W_n\}$. It is easy to check that Z_n is irreducible and atoroidal.

By Thurston's uniformisation theorem for atoroidal Haken manifolds [48] (see Morgan [37] and Kapovich [23] for the proof), there exists a geometrically finite hyperbolic 3-manifold N_n with a homeomorphism $f_n : Z_n \to (N_n)_0$ which can be extended to the conformal map from $\partial_\infty Z_n$ to $\partial_\infty N_n$. By Theorem 9.5.1 (or the original bi-Lipschitz theorem by Brock–Canary–Minsky), we may assume that f_n is a K-bi-Lipschitz map. Since the geometric limit of Z_n based at x_0 is M^0, by the Ascoli–Arzelà theorem, $\{f_n\}$ converges uniformly on any compact set of M^0 to a K-bi-Lipschitz map $f : M^0 \to N_0$, where N is a geometric limit of N_n. By our definition of M_n^- and W_n, each simply degenerate brick $F \times J$ has a sequence of tubes in $\mathcal{V}_n^{\text{ext}}$, taken one for each n, whose longitudes l_n regarded as simple closed curves on F converge to the ending lamination $\nu(e)$ given on the end e contained in $F \times J$. By our definition of the metric on M^0, the lengths of the l_n with respect to the model metric on M^0 are uniformly bounded from above. Since f is bi-Lipschitz, the closed geodesics l_n^* in N freely homotopic to $f(l_n)$ have also uniformly bounded lengths. This shows that l_n^* must tend to the end $f(e)$ by the argument of §§6.3–6.4 of Bonahon [6]. Therefore, the end $f(e)$ of N_0 has the ending lamination $f_*(\nu(e))$.

Let G_n be a Kleinian group with $\mathbb{H}^3/G_n = N_n$. By the main theorem of [39], there is a sequence of geometrically finite hyperbolic 3-manifolds $N_n^k = \mathbb{H}^3/G_n^k$ without $\mathbb{Z}$-cusps such that G_n^k converges algebraically to G_n. We can choose N_n^k so that the domain of discontinuity of G_n^k converges to that of G_n in the sense of Carathéodory by defining G_n^k to be obtained by pinching the conformal structure at infinity along curves corresponding to the $\mathbb{Z}$-cusps of N_n and using Lemma 3 of Abikoff [1]. By Proposition 4.2 of Jørgensen–Marden [22], this implies that G_n^k converges strongly to G_n as $k \to \infty$. By performing hyperbolic Dehn surgeries along the torus cusps of N_n^k of type $(1, u_n)$ with sufficiently large $u_n \in \mathbb{N}$, we have quasi-Fuchsian manifolds $N_n^{'k}$ geometrically approximating N_n^k closer and closer as $k \to \infty$ as was shown in Bonahon–Otal [7] and Ohshika [38]. This gives rise to a sequence of quasi-Fuchsian manifolds $N_n^{'k}$ converging geometrically to N_n as $k \to \infty$. By the diagonal argument, we have a sequence of quasi-Fuchsian manifolds N_n' converging geometrically to N. This completes the proof of Theorem C. □

9.6.4 Proof of Theorem D

Proof of Theorem D Let G_1 and G_2 be non-elementary geometric limits of Kleinian surface groups isomorphic to $\pi_1(S)$ preserving parabolicity, and $f : N_1 = \mathbb{H}^3/G_1 \to N_2 = \mathbb{H}^3/G_2$ a homeomorphism preserving their end invariants. We may assume that $f((N_1)_0) = (N_2)_0$. By Theorem A, there exists a brick manifold M and a K-bi-Lipschitz homeomorphism $\eta_1 : M \to (N_1)_0$ preserving the end invariants. Then the composition $\eta_2 = f \circ \eta_1 : M \to (N_2)_0$ is also a homeomorphism preserving the end invariants. By Theorem 9.5.1, we can properly homotope η_2 to a K-bi-Lipschitz homeomorphism, which we denote by the same symbol η_2. Therefore $\eta_2 \circ \eta_1^{-1} : (N_1)_0 \to (N_2)_0$ is a bi-Lipschitz homeomorphism preserving the end invariants, which can be extended to a bi-Lipschitz map $\Phi : N_1 \to N_2$. This Φ can be lifted to a bi-Lipschitz homeomorphism $\widetilde{\Phi} : \mathbb{H}^3 \to \mathbb{H}^3$ between the universal coverings, which is equivariant with respect to the covering translations. Furthermore $\widetilde{\Phi}$ is extended to a quasi-conformal homeomorphism $\widetilde{\Phi}_\partial$ on the Riemann sphere $\widehat{\mathbf{C}}$ such that $\widetilde{\Phi}_\partial|_{\Omega_{G_1}}$ is a conformal homeomorphism from Ω_{G1} to Ω_{G_2}. On the other hand, the injectivity radii in the convex cores of our manifolds N_1 and N_2 are bounded above by the existence of uniform models for Kleinian surface groups as was proved in [12]. Theorem 2.9 in McMullen [32], which generalises Sullivan's rigidity theorem, says that any quasi-conformal deformation of a (torsion-free) Kleinian group whose convex core has injectivity radii bounded above has support in the region of discontinuity. Applying this theorem to our G_1 and G_2 we see that $\eta_2 \circ \eta_1^{-1}$, which is properly homotopic to f, is properly homotopic to an isometry. □

References

1. W. Abikoff, On boundaries of Teichmüller spaces and on Kleinian groups. III. Acta Math. **134**, 211–237 (1975)
2. I. Agol, Tameness of Hyperbolic 3-manifolds, preprint. https://arxiv.org/abs/math/0405568
3. J. Anderson, R. Canary, Algebraic limits of Kleinian groups which rearrange the pages of a book. Invent. Math. **126**, 205–214 (1996)
4. R. Benedetti, C. Petronio, *Lectures on Hyperbolic Geometry* (Universitext, Springer, Berlin, 1992)
5. S. Bleiler, A. Casson, Automorphisms of surfaces after Nielsen and Thurston, London Mathematical Society Student Texts, vol. 9 (Cambridge University, Cambridge, 1988), iv+105 pp.
6. F. Bonahon, Bouts des variétés hyperboliques de dimension 3. Ann. Math. **124**, 71–158 (1986)
7. F. Bonahon, J.-P. Otal, Variétés hyperboliques avec géodésiques arbitrairement courtes. Bull. London Math. Soc. **20**, 255–261 (1988)
8. B. Bowditch, Intersection numbers and the hyperbolicity of the curve complex. J. Reine Angew. Math. **598**, 105–129 (2006)
9. B. Bowditch, Length bounds on curves arising from tight geodesics. Geom. Funct. Anal. **17**, 1001–1042 (2007)
10. B. Bowditch, The ending lamination theorem, preprint (2011)

11. J. Brock, Iteration of mapping classes and limits of hyperbolic 3-manifolds. Invent. Math. **143**, 523–570 (2001)
12. J. Brock, R. Canary, Y. Minsky, The classification of Kleinian surface groups, II: The Ending Lamination Conjecture. Ann. of Math. **176**(2), 1–149 (2012)
13. D. Calegari, D. Gabai, Shrinkwrapping and the taming of hyperbolic 3-manifolds. J. Am. Math. Soc. **19**, 385–446 (2006)
14. R. Canary, Ends of hyperbolic 3-manifolds. J. Am. Math. Soc. **6**, 1–35 (1993)
15. R. Canary, D. Epstein, P. Green, Notes on notes of Thurston, in *Fundamentals of Hyperbolic Geometry: Selected Expositions*. London Mathematical Society Lecture Note Series, vol. 328 (Cambridge University, Cambridge, 2006), pp. 1–115
16. D.B.A. Epstein, A. Marden, Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces, in *Analytical and geometric aspects of hyperbolic space (Coventry/Durham, 1984)*. London Mathematical Society, Lecture Note Series, vol. 111 (Cambridge University, Cambridge, 1987)
17. A. Fathi, F. Laudenbach, V. Poénaru, Travaux de Thurston sur les surfaces, in *Astérisque* (Socity of Mathematical, France, 1979), pp. 66–67
18. M. Freedman, J. Hass, P. Scott, Least area incompressible surfaces in 3-manifolds. Invent. Math. **71**, 609–642 (1983)
19. U. Hamenstädt, Train tracks and the Gromov boundary of the complex of curves, in *Spaces of Kleinian groups*. London Mathematical of Society, Lecture Note Series, vol, 329 (Cambridge University, Cambridge, 2006), pp. 187–207
20. W. Harvey, Boundary structure of the modular group, in *Riemann Surfaces and Related Topics: Proceedings of the 1978 Stony Brook Conference* (State University New York, Stony Brook, 1978), pp. 245–251. Annual of Mathematical Studies No. 97 (Princeton University, Princeton, 1981)
21. J. Hempel, *3-Manifolds*. Annals of Mathematics Studies, No. 86 (Princeton University, Princeton, 1976)
22. T. Jørgensen, A. Marden, Geometric and algebraic convergence of Kleinian groups. Math. Scand. **66**, 47–72 (1990)
23. M. Kapovich, Hyperbolic manifolds and discrete groups, in *Progress in Mathematics Studies*, vol. 183 (Birkhäuser, New York, 2000)
24. S. Kerckhoff, W. Thurston, Non-continuity of the action of the modular group at Bers' boundary of Teichmüller space. Invent. Math. **100**, 25–47 (1990)
25. E. Klarreich, The boundary at infinity of the curve complex and the relative Teichmüller space, preprint (1999)
26. C. Lecuire, Modèles et laminations terminales (d'après Minsky et Brock-Canary-Minsky), Astérisque, vol. 361 (2014), pp. 299–323. Exp. No. 1068, ix
27. A. Marden, *Outer Circles: An Introduction to Hyperbolic* 3*-manifolds* (Cambridge University, Cambridge, 2007)
28. A. Marden, *Hyperbolic Manifolds: An Introduction in 2 and 3 Dimensions* (Cambridge University, Cambridge, 2016), xviii+515 pp.
29. H. Masur, Y. Minsky, Geometry of the complex of curves, I: Hyperbolicity. Invent. Math. **138**, 103–149 (1999)
30. H. Masur, Y. Minsky, Geometry of the complex of curves, II: Hierarchical structure. Geom. Funct. Anal. **10**, 902–974 (2000)
31. K. Matsuzaki, M. Taniguchi, *Hyperbolic Manifolds and Kleinian Groups* (Oxford University, Oxford, 1998)
32. C. McMullen, *Renormalization and 3-manifolds Which Fiber over the Circle*. Annals of Mathematics Studies, vol. 142 (Princeton University, Princeton, 1996)
33. C. McMullen, Complex earthquakes and Teichmüller theory. J. Am. Math. Soc. **11**, 283–320 (1998)
34. Y. Minsky, On rigidity, limit sets and end invariants of hyperbolic 3-manifolds. J. Am. Math. Soc. **7**, 539–588 (1994)

35. Y. Minsky, The classification of Kleinian surface groups I: models and bounds. Ann. of Math. **171**, 1–107 (2010)
36. Mahan Mj, K. Ohshika, Discontinuous motions of limit sets, preprint. https://arxiv.org/abs/1704.00269
37. J. Morgan, On Thurston's uniformisation theorem for three-dimensional manifolds, in *The Smith Conjecture* (Academic Press, New York, 1984), pp. 37–125
38. K. Ohshika, Geometric behaviour of Kleinian groups on boundaries for deformation spaces. Quart. J. Math. Oxford **43**(2), 97–111 (1992)
39. K. Ohshika, Geometrically finite Kleinian groups and parabolic elements. Proc. Edinburgh Math. Soc. **41**(2), 141–159 (1998)
40. K. Ohshika, Reduced Bers boundaries of Teichmüller spaces. Ann. Inst. Fourier **64**, 145–176 (2014)
41. K. Ohshika, Geometric limits and their applications, in *Handbook of Group Actions II* (International Press/Higher Education Press, 2015), pp. 245–270
42. K. Ohshika, *Divergence, Exotic Convergence and Self-Bumping in Quasi-Fuchsian Spaces* (2018). https://arxiv.org/abs/1010.0070, to appear in Ann. Fac. Sci. Toulouse
43. K. Ohshika, T. Soma, Geometry and topology of geometric limits II, to appear in *In the tradition of Thurston : Geometry and groups* ed. K. Ohshika, A. Papadopoulos, Springer Cham.
44. J.-P. Otal, Les géodésiques fermées d'une variété hyperbolique en tant que nœuds, in *Kleinian Groups and Hyperbolic 3-manifolds (Warwick 2001)*. London Mathematical Society Lecture Note Series, vol. 299 (Cambridge University Press, Cambridge, 2003), pp. 95–104
45. J.-P. Otal, William P. Thurston: "Three-dimensional manifolds, Kleinian groups and hyperbolic geometry". Jahresber. Dtsch. Math.-Ver. **116**, 3–20 (2014)
46. T. Soma, Geometric approach to ending lamination conjecture, preprint. https://arxiv.org/abs/0801.4236
47. W. Thurston, *The Geometry and Topology of* 3*-Manifolds*. Lecture Notes (Princeton university, Princeton, 1978). http://www.msri.org/publications/books/gt3m/
48. W. Thurston, Three dimensional manifolds, Kleinian groups and hyperbolic geometry. Bull. Am. Math. Soc. **6**, 357–381 (1982)

Chapter 10
Laminar Groups and 3-Manifolds

Hyungryul Baik and KyeongRo Kim

Abstract Thurston showed that the fundamental group of a closed atoroidal 3-manifold admitting a co-oriented taut foliation acts faithfully on the circle by orientation-preserving homeomorphisms. This action on the circle is called a universal circle action, due to the rich information it carries. In this chapter, we first review Thurston's theory of universal circles and follow-up work of other authors. We note that the universal circle action of a 3-manifold group always admits an invariant lamination. A group acting on the circle with an invariant lamination is called a laminar group. In the second half of the chapter, we discuss the theory of laminar groups and prove some interesting properties of laminar groups under various conditions.

Keywords Tits alternative · Laminations · Circle homeomorphisms · Fuchsian groups · Fibered 3-manifolds · Pseudo-Anosov surface diffeomorphism

MSC Classes 20F65, 20H10, 37C85, 37E10, 57M60

10.1 Introduction

A few years before Perelman came up with his proof of the Poincaré conjecture using the theory of Ricci flow [28, 29] (built upon the work of Hamilton [23]), Thurston showed his vision to finish the geometrization program using foliations in 3-manifolds in [31]. Although Thurston left the manuscript unfinished after Perelman's resolution of the geometrization conjecture, [31] contains abundant beautiful ideas which are closely related to many interesting results by a number of authors including Ghys [22], Calegari-Dunfield [11], Calegari [9], Fenley [16, 17], Barbot-Fenley [5], Gabai-Kazez [20, 21], Mosher [26], and Frankel [19].

H. Baik (✉) · K. Kim
Department of Mathematical Sciences, KAIST, Daejeon, South Korea
e-mail: hrbaik@kaist.ac.kr; cantor14@kaist.ac.kr

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_10

One of the main themes of [31] is to combine a few approaches to 3-manifolds which are proven to be successful and fruitful. In particular, a deep connection between codimension-1 objects in 3-manifold and 3-manifold group actions on the low-dimensional spaces has been investigated. One of the main theorems in this chapter is the following.

Theorem 10.1.1 (Thurston's Universal Circle for Co-orientable Taut Foliations [31]) *Let M be a closed atoroidal 3-manifold admitting a co-orientable taut foliation $\mathscr{F}$. Then there exists a faithful homomorphism* $\rho_{univ} : \pi_1(M) \to \mathrm{Homeo}^+(S^1)$.

A codimension-one foliation $\mathscr{F}$ of a manifold M is called *taut* if there exists a closed embedded loop γ in M which is transverse to the leaves of $\mathscr{F}$ and intersects every leaf of $\mathscr{F}$ at least once. In fact, the action in Thurston's theorem is not just any group action on the circle. Thurston called the circle obtained in the above theorem a *universal circle* for the taut foliation $\mathscr{F}$. Let us denote it by S^1_{univ}. The name suggests that ρ_{univ} is not just an action but it "sees" the structure of the foliation. In fact, a universal circle consists of following data:

(1) Let $\widetilde{\mathscr{F}}$ be the covering foliation on $\mathscr{F}$ in the universal cover $\widetilde{M}$ of M. For each leaf λ of $\widetilde{\mathscr{F}}$, there exists a circle $S^1_\infty(\lambda)$ so that the action of $\pi_1(M)$ on the leaves extends continuously to the set of such circles.
(2) For each leaf λ of $\widetilde{\mathscr{F}}$, there exists a monotone map $\phi_\lambda : S^1_{univ} \to S^1_\infty(\lambda)$, i.e., a continuous surjection so that the preimage of each point in the range is connected.
(3) For each $\alpha \in \pi_1(M)$ and for each leaf λ, the following diagram commutes:

$$\begin{array}{ccc} S^1_{univ} & \xrightarrow{\rho_{univ}(\alpha)} & S^1_{univ} \\ \phi_\lambda \downarrow & & \downarrow \phi_{\alpha(\lambda)} \\ S^1_\infty(\lambda) & \xrightarrow{\alpha} & S^1_\infty(\alpha(\lambda)) \end{array}$$

(4) (comparability condition) For each leaf λ of $\widetilde{\mathscr{F}}$, the maximal connected intervals in S^1_{univ} which are mapped to points by ϕ_λ are called the *gaps* associated to λ and the complement of the gaps is called the *core* associated to λ. For any two incomparable leaves μ, λ, the core associated to μ is contained in a single gap associated to λ and vice versa.

For the construction of the universal circle, we borrow material largely from [11], so for the interested readers, please consult [11] for details. Here we recall the main ingredients and rough ideas to see the big picture. As we will see in the construction, there are some choices involved and as a result, a universal circle is not unique. Perhaps coming up with a canonical way of obtaining a universal circle via some universal property would be desirable.

Many results analogous to Theorem 10.1.1 have been obtained in the literature under the presence of other codimension-1 objects or flows in the 3-manifold. For instance, Calegari obtained the result for 3-manifolds with quasi-geodesic flows [9], and Calegari and Dunfield showed this result in the case of essential laminations with solid torus guts [11]. Later Hamm in his PhD thesis [24] generalized Calegari-Dunfield's work to a more general class of essential laminations.

In Sects. 10.2–10.4, we briefly review these works. In Sect. 10.5, we observe that in all those cases, the action on the circle comes with an invariant lamination. This motivates the study of groups acting on the circle with invariant laminations (and such groups are called laminar groups). In Sects. 10.6–10.11, we discuss some recent and on-going work in the theory of laminar groups. We emphasize that by no means the review of the material in the literature in Sects. 10.2–10.5 can serve as a thorough survey for all related works.

10.2 S^1-Bundle over the Leaf Space

Let $\mathscr{F}$ be a co-oriented taut foliation in a 3-manifold M. Let $\widetilde{\mathscr{F}}$ be the foliation on the universal cover $\widetilde{M}$ of M which covers $\mathscr{F}$, and let $L = L(\widetilde{\mathscr{F}})$ be the leaf space of this covering foliation. As a set, each point of L corresponds to a leaf of $\widetilde{\mathscr{F}}$. To give a topology, we say that a sequence of leaves μ_i *converges* to a leaf μ_∞ if for every compact subset K of $\widetilde{M}$, $\mu_i \cap K$ converges to $\mu_\infty \cap K$ in the Hausdorff topology.

The leaf space L is a one-dimensional manifold in the sense that each point has a neighborhood homeomorphic to $\mathbb{R}$ but L does not have to be Hausdorff. In fact, the leaf space L is Hausdorff if and only if it is homeomorphic to $\mathbb{R}$. In that case, we say that $\mathscr{F}$ is an $\mathbb{R}$-*covered foliation*.

In all other cases, L is not Hausdorff. The co-orientation of $\mathscr{F}$ gives an orientation on each embedded line segment of the leaf space. Therefore, it induces a partial order on the leaf space L. For two leaves α, β of $\widetilde{\mathscr{F}}$, we say that $\alpha < \beta$ if there exist an embedded closed interval in L whose endpoints are α, β, and it is an oriented path from α to β with respect to the induced orientation. One caveat is that we need to know that there exists no closed transversal to $\widetilde{\mathscr{F}}$. This will be shown later in this section. $\mathscr{F}$ is $\mathbb{R}$-covered if and only if the induced partial order on L is a total order. In general, if $\mathscr{F}$ is not $\mathbb{R}$-covered, there are incomparable leaves.

We say that $\mathscr{F}$ is *branched in the forward direction* (a.k.a. has one sided branching in the positive direction [10]) if it is not $\mathbb{R}$-covered and if for any two leaves α, β of $\widetilde{\mathscr{F}}$ there exists a leaf γ of $\widetilde{\mathscr{F}}$ such that $\alpha > \gamma$ and $\beta > \gamma$. Similarly, one can define a *branching in the backward direction*. In this chapter, when we think of a non $\mathbb{R}$-covered foliation, we only consider the case where $\mathscr{F}$ has two-sided branching i.e., it is branched in both forward and backward direction for simplicity. For what we discuss in this section, this assumption is not so relevant.

We would like to construct what can be called an S^1-bundle over $L(\widetilde{\mathscr{F}})$ in some sense. In other words, we would like to assign one copy of the circle to each leaf, but where does it come from? To begin with, we recall the result of Candel.

In general, for a manifold M with dimension $n \geq 3$, a two-dimensional lamination is called a Riemann surface lamination if each leaf is a Riemann surface.

More precisely, suppose that M admits an atlas with product charts $\phi_p : U_p \to B_p \times K_p$ where B_p is a domain in $\mathbb{C}$, K_p a closed subspace of $\mathbb{R}^{n-2}$, U_p an open subset of M, and ϕ_p a homeomorphism. We further assume that each coordinate change has the form $\phi_p \circ \phi_q^{-1}(b, k) = (\psi(b, k), \rho(k))$ where ψ, ρ are continuous functions and ψ is holomorphic in b. Such an atlas Λ is called a Riemann surface lamination. We will focus on the case where M is a 3-manifold, and Λ is a surface lamination in M. In fact, we assume M to be a closed hyperbolic 3-manifold throughout the rest of the chapter, since the theory of universal circles was built upon the assumption that the 3-manifold is closed and atoroidal (we know now it is hyperbolic due to the geometrization theorem). Also, B_p is always taken to be the unit disk D. Hence we consider the product charts $U_i = D \times K_i$.

Candel obtained a significant generalization of the classical uniformization theorem for Riemann surfaces in the setting of Riemann surface laminations. In particular, this provides a sufficient condition for (M, Λ) to admit a Riemannian metric so that its restriction to Λ is a leaf-wise hyperbolic metric. We only recall the main ideas. For details on Candel's work, we refer the readers to [12] or [10].

The classical uniformization theorem says that if a closed Riemann surface has negative Euler characteristic, then it admits a hyperbolic metric. To state a similar result for laminations, we need to develop a notion which plays a role similar to the Euler characteristic. To do this, we first need to discuss invariant transverse measures on laminations. An invariant transverse measure μ for a lamination Λ is a collection of nonnegative Borel measure on the leaf space of Λ in each product chart which is compatible on the overlap of distinct charts.

Now when Λ is a Riemann surface lamination, the leafwise metric determines a leafwise closed 2-form, say Ω. The product measure $\mu \times \Omega$ is a signed Borel measure on the total space Λ. We call the total mass of this measure the Euler characteristic $\chi(\mu)$ of μ. As in the case of the classical uniformization theorem, the sign of the Euler characteristic is important.

Note that if $U = D \times K$ is a product chart for Λ, then $(\mu \times \Omega)(U) = \int_K (\int_{D \times k} \Omega) d\mu(k)$. When Λ admits a leafwise hyperbolic metric, then $\int_{D \times k} \Omega$ is negative and μ is a nonnegative measure by definition, hence $(\mu \times \Omega)(U)$ is negative for each product chart U. As a consequence, we have $\chi(\mu) < 0$. Candel proved that the converse is also true.

Theorem 10.2.1 (Candel's Uniformization Theorem [12]) *Let Λ be a Riemann surface lamination. Then Λ admits a leafwise hyperbolic metric if and only if the Euler characteristic $\chi(\mu)$ is negative for all nontrivial invariant transverse measure μ.*

While Candel's theorem is not specifically about hyperbolic manifolds, we go back to our own setting where M is a closed hyperbolic 3-manifold and $\mathscr{F}$ is a co-orientable taut foliation. First, we observe that no leaf of $\mathscr{F}$ is the 2-sphere S^2. This follows from the Reeb stability theorem.

Theorem 10.2.2 (Reeb Stability Theorem) *Let $\mathscr{F}$ be a cooriented taut foliation in a closed 3-manifold M. Suppose that $\mathscr{F}$ has a leaf homeomorphic to S^2. Then M is homeomorphic to $S^2 \times S^1$ and $\mathscr{F}$ is the product foliation by spheres.*

Sketch of the Proof Since $\pi_1(S^2)$ is trivial, the holonomy along any path on the spherical leaf is trivial. Therefore, the spherical leaf has a neighborhood which is foliated as a product. This shows that the set of spherical leaves form an open subset of M.

Since M is compact, We know that if we have a sequence of closed leaves λ_i which converge to a leaf λ, then λ is also closed. If all λ_i are spheres, then in a small neighborhood of λ, the projection along the vertical direction in each product chart defines a covering map from λ_i to λ for large enough i. Since $\mathscr{F}$ is co-oriented, λ is also necessarily a sphere. Therefore, the set of spherical leaves forms a closed subset of M. Since the set is both open and closed, it should be M itself. □

Since M is assumed to be hyperbolic in our case, we do not have any spherical leaf.

Also no leaf is a torus. Indeed, since M is atoroidal, if any leaf is a torus, then it would bound a solid torus. One can foliate the solid torus where the boundary is also a leaf, and it is called a Reeb component. First, one can foliate $H = \{(x, y, z) \in \mathbb{R}^3 : z \geq 0\}$ by the horizontal planes $\{(x, y, z) : z = c\}$. Quotient $H \setminus \{(0, 0, 0)\}$ by the equivalence relation $(x, y, z) \sim (2x, 2y, 2x)$. In this case, one can easily see that if a transversal in M travels from the complement of the Reeb component into the Reeb component by passing through the boundary leaf of the Reeb component (the torus leaf), it cannot escape from the Reeb component again. Hence, if $\mathscr{F}$ is a taut foliation, it cannot have a Reeb component.

From this, one can conclude that each leaf of $\mathscr{F}$ is of hyperbolic type. Therefore, the condition of Candel's theorem is satisfied, and M admits a leafwise hyperbolic metric.

By a work of Rosenberg [30] which is an important improvement of the seminal work of Novikov [27], we know the followings about M and $\mathscr{F}$:

(i) M is irreducible.
(ii) For each leaf α of $\mathscr{F}$, the inclusion map $\alpha \hookrightarrow M$ induces an injective homomorphism $\pi_1(\alpha) \to \pi_1(M)$.
(iii) Every closed transversal to $\mathscr{F}$ is nontrivial in $\pi_1(M)$.

Here we immediately see that the leaf space L is a tree in the sense that there is no cycle embedded in L. If there exists such a cycle, it corresponds to a closed transversal to $\widetilde{\mathscr{F}}$, so it projects down to a closed transversal to $\mathscr{F}$. Then it must be nontrivial in $\pi_1(M)$ while it lifts to a loop in the universal cover of M, a contradiction. We remark that the word tree is not used to mean a simplicial tree. It is a simply connected (not necessarily Hausdorff) 1-manifold.

From this result, one can immediately deduce the following theorem.

Theorem 10.2.3 *Let $M, \mathscr{F}$ be as above. Then every leaf of $\widetilde{\mathscr{F}}$ is a properly embedded plane in $\widetilde{M}$.*

Sketch of a Proof First of all, every leaf of $\widetilde{\mathscr{F}}$ is simply connected. Let $\widetilde{\lambda}$ be a leaf of $\widetilde{\mathscr{F}}$, and γ a loop on $\widetilde{\lambda}$. Note that $\widetilde{\lambda}$ is a covering of some leaf λ of $\mathscr{F}$.

Since $\widetilde{M}$ is simply connected, γ is homotopically trivial in $\widetilde{M}$, so it must be homotopically trivial in M. On the other hand, by the theorem of Novikov–Rosenberg above, λ is π_1-injectively embedded in M. Thus, γ must be trivial in λ. By the homotopy lifting property, this implies that the original loop γ is homotopically trivial in $\widetilde{\lambda}$. Since γ is arbitrary, this implies that $\widetilde{\lambda}$ is simply connected.

Now by the Reeb stability theorem, no leaf is a sphere. Hence all leaves of $\widetilde{\mathscr{F}}$ must be planes. For a leaf $\widetilde{\lambda}$ of $\widetilde{\mathscr{F}}$, if it is covered by product charts so that in each chart, the intersection with $\widetilde{\lambda}$ is connected (each connected component is called a plaque), then it must be properly embedded. Therefore, if $\widetilde{\lambda}$ is not properly embedded, there exists a product chart where $\widetilde{\lambda}$ intersects in at least two plaques. In that case, one can make a closed loop in $\widetilde{M}$ such that first, one can use the transversal in that product chart to connect two points in different plaques of $\widetilde{\lambda}$, and close it up by a path contained in $\widetilde{\lambda}$. Now this path in $\widetilde{\lambda}$ is covered by finitely many product charts, so one can tilt it to get a transversal which is very close to the original path (see Fig. 10.1. In our case, the charts U_1 and U_n could coincide). Using this technique, one gets a closed transversal $\widetilde{\gamma}$ to $\widetilde{\mathscr{F}}$ which intersects $\widetilde{\lambda}$. The transversal gets mapped to a closed transversal γ in M and by Part (iii) of the Novikov–Rosenberg theorem stated above, γ must be homotopically nontrivial. On the other hand, since $\widetilde{M}$ is simply connected, $\widetilde{\gamma}$ is homotopically trivial, a contradiction. We conclude that every leaf is properly embedded. □

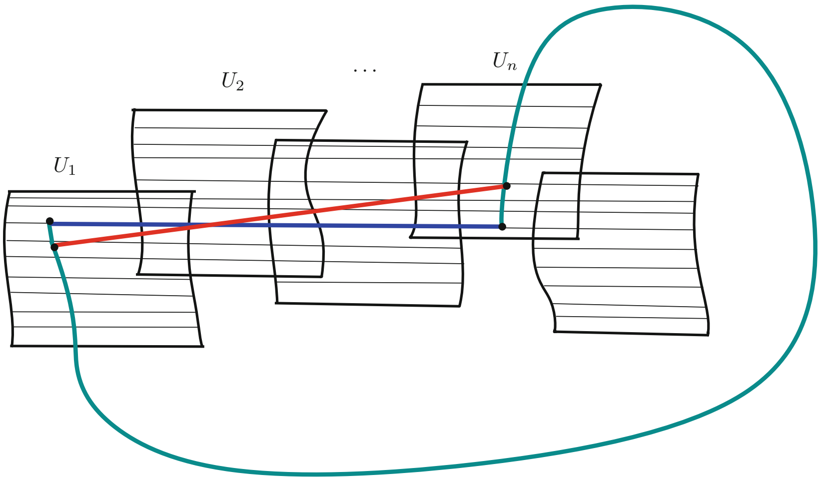


Fig. 10.1 Consider the loop obtained by concatenating the blue arc which is contained in a leaf of the foliation with the green arc which is assumed to be transverse to the foliation. The blue arc is tilted to the red arc to make the whole loop transverse to the foliation. Note that one cannot draw the green arc so that it intersects the chart U_n from below, since this would contradict the fact that the foliation is co-oriented

Combining this result with Candel's theorem, we find a metric on M so that each leaf of $\widetilde{\mathscr{F}}$ equipped with the induced path metric is isometric to the hyperbolic plane $\mathbb{H}^2$. For each leaf λ of $\widetilde{\mathscr{F}}$, since λ can be identified with $\mathbb{H}^2$ and the ideal boundary of $\mathbb{H}^2$ is homeomorphic to the circle (called the circle at infinity), we get the circle at infinity $S^1_\infty(\lambda)$ for λ. Now we define the *circle bundle at infinity* E_∞ as the set of all circles at infinity for the leaves of $\widetilde{\mathscr{F}}$. In other words, $E_\infty = \cup_{\lambda \in L} S^1_\infty(\lambda)$. E_∞ can be obtained from the "cylinders" over each transverse arc to $\widetilde{\mathscr{F}}$ by patching them together appropriately. We explain what this means in the next section.

10.3 Leaf Pocket Theorem and the Special Sections

Now we have circles, one for each leaf of $\widetilde{\mathscr{F}}$. We need to combine them to make one big mother circle which we will call a universal circle. This is done as follows: in the previous section, we defined E_∞ as a set, so we first give a description of its topology. Second, we note that there are some special sections for the bundle E_∞ which are preserved under the deck group action on $\widetilde{M}$. Third, we observe that they can be circularly ordered so that the deck group action is order-preserving. Finally, taking an order completion of the set of special sections, we get a circle.

To do this, we need to understand both "tangential geometry" and "transverse geometry" of $\mathscr{F}$. For the tangential geometry, here is one useful lemma.

Lemma 10.3.1 *There exists $\epsilon > 0$ such that every leaf of $\widetilde{\mathscr{F}}$ is quasi-isometrically embedded in its ϵ-neighborhood.*

Proof For each point p in M, consider a product chart U_p which is evenly covered by the universal covering map so that one connected component of the preimage of U_p is a product chart where each leaf of $\widetilde{\mathscr{F}}$ intersects at most once. The last condition can be satisfied for the reason explained in the proof of Theorem 10.2.3.

By compactness of M, there exist finitely many points $p_1, \ldots, p_n$ so that M is covered by $U_{p_1}, \ldots, U_{p_n}$. For simplicity of the notation, we write U_{p_i} as U_i. Again since M is compact, we can apply the Lebesgue number lemma to conclude that there exists $\epsilon > 0$ such that every ball of radius 2ϵ is contained in one of the product charts U_i.

Now let λ be any leaf of $\widetilde{\mathscr{F}}$, and let N be the ϵ-neighborhood of λ. By our choice of ϵ, lifts of the product charts U_i cover the entire N. Since these are lifts of finitely many product charts, they have uniformly bounded geometry. This shows that λ is quasi-isometrically embedded in N. □

A positive number ϵ as in the above lemma is called a *separation constant* of $\mathscr{F}$.

The transverse geometry of $\mathscr{F}$ is described in the so-called leaf pocket theorem. To state the theorem, we first need to define the *endpoint map*. Let λ be a leaf of $\widetilde{\mathscr{F}}$, and p a point in it. Then for any vector u in the unit tangent space $UT_p\lambda$ at p, let $e(u)$ be the endpoint in $S^1_\infty(\lambda)$ of the geodesic ray in λ determined by u. This defines a map, which again we call e, from the unit tangent bundle $UT\widetilde{\mathscr{F}}$ of $\widetilde{\mathscr{F}}$ to E_∞. Now we give E_∞ the finest topology so that the map $e : UT\widetilde{\mathscr{F}} \to E_\infty$ is continuous.

Now we explain what we meant by "patching cylinders" in the last section. Let τ be any transverse arc to $\widetilde{\mathscr{F}}$. Then $UT\widetilde{\mathscr{F}}_\tau$ is literally a cylinder. If $e(v_1) = e(v_2)$ for $v_1 \in UT\widetilde{\mathscr{F}}_{\tau_1}$ and $v_2 \in UT\widetilde{\mathscr{F}}_{\tau_2}$ for two transverse arcs τ_1, τ_2, then we identify v_1 and v_2. Hence E_∞ is obtained from the disjoint union of cylinders of the form $UT\widetilde{\mathscr{F}}_\tau$ under these identifications.

Going back to the transverse geometry of the foliation, we call a map $m : I \times \mathbb{R}_{\geq 0} \to \widetilde{M}$ a *marker* if $m(\{k\} \times \mathbb{R}_{\geq 0})$ is a geodesic ray in a single leaf of $\widetilde{\mathscr{F}}$ for each $k \in I$ and $m(I \times \{t\})$ is a transverse arc with length no greater than $\epsilon/3$ for all $t \in \mathbb{R}_{\geq 0}$ where ϵ is a separation constant for $\mathscr{F}$.

Let $p \in \widetilde{M}$ and λ be a leaf of $\widetilde{\mathscr{F}}$ containing p. Suppose there exists a marker m such that $p = m(k, 0)$ for some $k \in I$. This means that there exists a transversal $m(I \times \{0\})$ at p, the holonomy along the geodesic ray $m(\{k\} \times \mathbb{R}_{\geq 0})$ emanating from p being defined for the whole time. Said differently, along this ray, nearby leaves are not pulled away from the leaf λ too fast. The following theorem of Thurston shows that for arbitrary $p \in \widetilde{M}$, there exist abundant directions with this property. This describes the transverse geometry of $\mathscr{F}$.

The original proof of the leaf pocket theorem given by Thurston in [31] uses the existence of harmonic measures for foliations. An alternative, purely topological proof is given by Calegari-Dunfield [11]. We omit the discussion of the proof here and only briefly explain how this theorem is applied to get a set of cyclically ordered set of sections.

Theorem 10.3.2 (Leaf Pocket Theorem [11, 31]) *For every leaf λ of $\widetilde{\mathscr{F}}$, the set of endpoints of markers is dense in $S^1_\infty(\lambda)$.*

Abusing the notation, we also call the set of endpoints of a marker a *marker*. Let C be a cylinder in E_∞ i.e, $C = \cup_{\lambda \in I} S^1_\infty(\lambda)$ where I is an interval in L. C is foliated by circles at infinities for the leaves corresponding to points in I. The first thing to observe is that any two markers contained in C are either disjoint or their union is an interval transverse to the circle fibers in C.

This is actually a consequence of the tangential geometry of $\mathscr{F}$ (more precisely the existence of a separation constant ϵ). Suppose that two markers m_1, m_2 intersect at a point in $S^1_\infty(\lambda)$ but have distinct endpoints on $S^1_\infty(\mu)$ for some leaves $\lambda, \mu \in I$. On λ, the geodesic rays of m_i's become arbitrarily close to each other, since they have the same endpoints on the ideal boundary. Hence, by shortening the markers horizontally (i.e., by restricting markers to $[t, \infty)$ for sufficiently large t), we may assume that they are within $\epsilon/3$-distance from each other on λ with respect to the metric on $\widetilde{M}$. Since each marker is $\epsilon/3$-thin, the geodesic rays of m_i's on μ are within ϵ-distance from each other again with respect to the metric on $\widetilde{M}$. However, those rays diverge on μ, hence with respect to the hyperbolic metric on μ, the rays get arbitrarily far away from each other. This contradicts the fact that μ is quasi-isometrically embedded in its ϵ-neighborhood.

From this fact together with the leaf pocket theorem, we can start constructing sections on C. First, pick a set T of finitely many markers on C so that each non-boundary circle fiber of C intersects at least one marker at an interior point of the marker, and the boundary circle fibers meet at least one marker at the endpoint.

To make our description simple, let us parametrize I (recall that C is a circle bundle over an interval I in L) to be the closed interval [0, 1], and the leaf corresponding to point $t \in [0, 1]$ is denoted by λ_t. Let $p \in S^1_\infty(\lambda_t) \subset C$ for some t. We can choose a "left-most" path through p with respect to T in the following way: On $S^1_\infty(\lambda_t)$, we start from p and move anti-clockwise until we hit a marker. At the marker, follow the marker upward (increasing the parameter t) until the end of the marker. At the end, move anti-clockwise as much as you can until you hits another marker. Follow the marker upward. In this way we construct a path from p to $S^1_\infty(\lambda_1)$.

Let us call this path $\gamma_{p,T}$. Now make the set T bigger by adding more markers on C to get a new set T' of markers. If new markers do not intersect the path $\gamma_{p,T}$, there is nothing to do in the sense that $\gamma_{p,T} = \gamma_{p,T'}$. Hence let us assume that a new marker m intersects the path $\gamma_{p,T}$. This means that at some t, $\gamma_{p,T}$ moves horizontally on $S^1_\infty(\lambda_t)$ but the marker m crosses it vertically. Hence, when we construct the path with respect to the set $T \cup \{m\}$ of markers, our path should stop at $m \cap S^1_\infty(\lambda_t)$ and follow m upward, and then move horizontally anti-clockwise again until hitting other markers in the set. Then one can observe that the path $\gamma_{p,T'}$ is slightly perturbed to the right compared to $\gamma_{p,T}$. To make this more precise, one can unwrap the circle fibered of C to the real line $\mathbb{R}$ to get a simply-connected cover of C which is now foliated by horizontal lines. Here we see this cover so that on each line fiber, moving to the left corresponds to moving anti-clockwise on a circle fiber on C. Then clearly the new path $\gamma_{p,T'}$ is on the right compared to $\gamma_{p,T}$ (here one should fix a lift $\widetilde{p}$ of p and consider the lifts of the paths passing through $\widetilde{p}$). An important point is that they cannot cross each other, although they are likely to coalesce.

Now for any two paths γ and δ on C, we say $\gamma \leq \delta$ if the lift of δ through $\widetilde{p}$ is on the right side to the lift of γ through $\widetilde{p}$ in the universal cover of C. This gives a partial ordering on the set of paths on C. For any two sets of markers $T \subset T'$, we get $\gamma_{p,T} \leq \gamma_{p,T'}$. Now we define a section $\tau_p : I \to C$ by $\tau_p(\nu) = \sup\{\min \gamma_{p,T} \cap S^1_\infty(\nu) : T \text{ is a set of markers}\}$. Here the minimum means the projection of the left-most point in the universal cover of C, and the supremum exists because the lifts of paths $\gamma_{p,T}$ to the cover of C through $\widetilde{p}$ are bounded from above by the vertical line through $\widetilde{p}$. This new path τ_p is continuous since the set of markers meets each circle fiber at a dense subset. Consequently, we get a continuous section τ_p of the circle-bundle C over I and call it a *left-most section* starting from p.

Starting from p, one can also move downward in the leaf space L. In this case, instead of using the left-most paths, we take right-most paths by moving clockwise on each circle fiber and following markers downward. This is called a *right-most section* starting from p. Hence, for each embedded line A in L, one can get a section τ_p of the bundle $E_\infty|_A$ over A by taking a left-most section when we move upward from p along A, and taking a right-most section when we move downward from p along A. But we would like to extend τ_p as a section for the bundle $E_\infty \to L$.

Before we proceed, we need one definition. Consider a sequence (μ_i) of leaves of $\widetilde{\mathscr{F}}$ which are contained in a single totally ordered segment of L and increasing with respect to that order. We call such a sequence *monotone ordered*. Suppose there exists a collection of leaves $\{\lambda_j\}$ of $\widetilde{\mathscr{F}}$ such that μ_i converges on compact subsets of $\widetilde{M}$ in the Hausdorff topology to the union of leaves λ, then we call the collection $\{\lambda_j\}$ together with the monotone ordered sequence (μ_i) a *cataclysm*. Here the convergence means that for any compact subset K of M, $\mu_i \cap K$ converges to $(\cup_j \lambda_j) \cap K$ in the Hausdorff topology. In fact, it is more appropriate to consider the cataclysms up to some natural equivalence relation because the sequence (μ_i) is not an essential part of the data. So as long as we have two monotone ordered sequences contained in a single totally ordered segment of L which converge to the same collection of leaves $\{\lambda_j\}$, we say that those two cataclysms are *equivalent*. Abusing the notation, we will just call the collection $\{\lambda_j\}$ a *cataclysm*.

For the purpose of visualization, see Fig. 10.2 for one lower-dimensional example. On the left-hand side of the figure, what is drawn is a foliation of $\mathbb{R}^2$ which is divided into three parts by two red vertical lines. The left-most part and right-most part, they are foliated by black vertical lines, and in the middle region, it is foliated by blue curves where each end of each blue curve is asymptotic to one of the red lines. On the right-hand side of the figure, its leaf space is drawn. By construction, the leaf space is not Hausdorff. Here the red leaves form a cataclysm, and a monotone ordered sequence can be taken to be a sequence of blue leaves.

Let λ be a leaf so that $p \in S^1_\infty(\lambda)$ and let μ be any other leaf in L. There exists a unique broken path from λ in μ which is obtained in the following way: first collapse each cataclysm to a point in L to get an actual tree Y, take the unique path from λ to μ in Y, and pull it back to L. This broken path is a union of embedded intervals in L with occasional jumps between two leaves in the same cataclysm. Say, in this broken path, τ_p comes down to λ_1 and it jumps to λ_2 which is in the same cataclysm

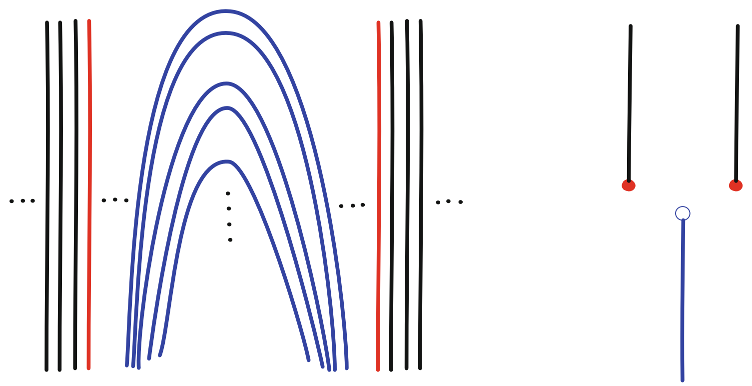


Fig. 10.2 A two-dimensional example of a foliation and its leaf space

with λ_1 and then move upward from there. Say μ_i is a monotone ordered sequence converging to λ_1 and λ_2.

Suppose I_1, I_2 are two intervals in L such that they coincide in an half-open interval I and differ at only one vertex, μ_i are in I, and $I_j = I \cup \{\lambda_j\}$ for $j = 1, 2$. For each $j \in \{1, 2\}$, let m_j, m'_j be any two markers so that they have one endpoint on $S^1_\infty(\lambda_j)$ and the rest lie in the circle-bundle C over I. For later use, let us call the circle-bundle over I_j, say C_j, for each j. First note that m_1 and m_2 are disjoint on C. Otherwise, since they are $\epsilon/3$-thin, again we get a contradiction to the fact that ϵ is a separation constant for $\mathscr{F}$. Also, for each μ_i which intersects all the markers m_1, m'_1, m_2, m'_2, the pairs (m_1, m'_1) and (m_2, m'_2) are unlinked in the circle $S^1_\infty(\mu_i)$. Here we say that two pairs (a, b) and (c, d) of points of the circle are *unlinked* if both a, b are contained in the closure of a single connected component of $S^1 \setminus \{c, d\}$. If they are linked, since λ_1 gets close to μ_i near the pair (m_1, m'_1) and λ_2 gets close to μ_i near the pair (m_2, m'_2), either λ_1 and λ_2 are comparable in L or they must intersect. We know that λ_1 and λ_2 are incomparable, so this is impossible. Consequently, one can take disjoint arcs J_1, J_2 of $S^1_\infty(\mu_i)$ so that the set of endpoints of the markers in C_j on $S^1_\infty(\mu_i)$ are completely contained in J_j.

Let $S^1_{\lambda_1\lambda_2}$ be the circle obtained from $S^1_\infty(\mu_i)$ by collapsing each connected component of the complement of the closure of the set of intersections with the markers through either λ_1 or λ_2. Naturally for each j, there exists a monotone map $\phi_j : S^1_{\lambda_1\lambda_2} \to S^1_\infty(\lambda_j)$. For instance, ϕ_1 collapses the arc obtained from the image of the closure of the set of intersections with the markers through λ_2 under the monotone map $S^1_\infty(\mu_i) \to S^1_{\lambda_1\lambda_2}$, and similarly for ϕ_2. Then the preimage of $\tau_p(\lambda_1)$ under ϕ_1 gets mapped to a single point in $S^1_\infty(\lambda_2)$ via ϕ_2. Let this point be $\tau_p(\lambda_2)$. We continue by constructing a left-most section starting at $\tau_p(\lambda_2)$. This procedure allows us to construct τ_p along the broken path from λ to μ, therefore we get a well-defined value $\tau_p(\mu)$. We call a section for $E_\infty \to L$ a *special section* if it is τ_p for some $p \in E_\infty$ and constructed as above.

Let $\mathscr{S}$ be the set of all special sections. First of all, they are built upon the set of markers which is preserved under the $\pi_1(M)$-action, since the markers are constructed using the geometry of the foliation. One can also check easily that $\mathscr{S}$ admits a natural cyclic order. For a triple $(\tau_{p_1}, \tau_{p_2}, \tau_{p_3})$, there must exist $\mu \in L$ so that $\tau_{p_1}(\mu), \tau_{p_2}(\mu), \tau_{p_3}(\mu)$ are distinct. Hence they inherit a cyclic order from the orientation on $S^1_\infty(\mu)$. Clearly this cyclic order is preserved by the $\pi_1(M)$-action, since the cyclic order on each cataclysm is determined by the geometry of the foliation as well. Of course we put many details under the rug. For the full detail of the proof, see Section 6 of [11].

By taking the completion of the set of special sections of E_∞ as an ordered set, one gets a circle S^1_{univ} where $\pi_1(M)$ acts by order-preserving homeomorphisms.

Recall the definition of a universal circle given as a set of data in the introduction. We also need a monotone map $\phi_\lambda : S^1_{univ} \to S^1_\infty(\lambda)$ for each leaf λ of $\widetilde{\mathscr{F}}$. For a point p on S^1_{univ} corresponding to a special section, $\phi_\lambda(p)$ is just the evaluation of the section at λ. From the construction, it is clear that ϕ_λ is monotone, and commutativity of the diagram in the definition of the universal circle holds. Also, for incomparable leaves λ_1, λ_2, since the core of ϕ_{λ_1} is the closure of the points in S^1_{univ} corresponding to the special sections through a point on $S^1_\infty(\lambda_1)$ and since they are entirely collapsed to a single point in $S^1_\infty(\lambda_2)$ (recall the part where we constructed the circle $S^1_{\lambda_1\lambda_2}$ above), the core of ϕ_{λ_1} must be contained in a gap of ϕ_{λ_2}. This is actually contained in a single gap because the fact that the markers through $S^1_\infty(\lambda_1)$ are unlinked with the markers through $S^1_\infty(\lambda_2)$ implies that the same fact holds for special sections.

One last thing to check is that the action on S^1_{univ} is faithful. In the case of $\mathbb{R}$-covered foliations, one can find a transverse pseudo-Anosov flow, and in that case the faithfulness can be verified using the ideas in [7]. See also Sect. 10.4 to see the detail of the pseudo-Anosov flow case. Hence we consider only the case that $\mathscr{F}$ has branching. Let H be the kernel of the action $\rho_{univ} : \pi_1(M) \to \text{Homeo}\, S^1_{univ}$.

Suppose H is nontrivial. Let h be any nontrivial element of H and let λ be any leaf of $\widetilde{\mathscr{F}}$. We have the following commutative diagram:

$$\begin{array}{ccc} S^1_{univ} & \xrightarrow{\rho_{univ}(h)} & S^1_{univ} \\ \phi_\lambda \big\downarrow & & \big\downarrow \phi_{h(\lambda)} \\ S^1_\infty(\lambda) & \xrightarrow{h} & S^1_\infty(h(\lambda)) \end{array}$$

Since h acts trivially on the universal circle, the top map $\rho_{univ}(h)$ is the identity map.

If $h(\lambda) = \lambda$, then by the above diagram, we know that h acts trivially on $S^1_\infty(\lambda)$. This implies that h acts on λ as the identity but this is impossible since h is a nontrivial element of $\pi_1(M)$. Hence $h(\lambda)$ is different from λ.

Second, we observe that λ and $h(\lambda)$ are comparable. Indeed, suppose they are incomparable. By the commutativity of the above diagram, any gap associated with λ is contained in a gap associated with $h(\lambda)$, but also the core associated with λ is contained in a single gap associated with $h(\lambda)$, a contradiction. Therefore, the leaves λ and $h(\lambda)$ are comparable.

Let λ, μ be two distinct leaves contained in the same cataclysm in L. From above discussion, $H\lambda$ is an infinite set contained in a line X of L, and similarly, $H\mu$ is an infinite set contained in a line Y of L. For each $h \in H$, $h(\lambda)$ and $h(\mu)$ are two distinct leaves contained in the same cataclysm. This shows that there exists infinitely many pairs of points $(x, y) \in X \times Y$ such that x and y are contained in the same cataclysm. But this is impossible for two lines in L, since there cannot be a nontrivial cycle in L. This is a contradiction, so we conclude that H must be trivial, i.e., the $\pi_1(M)$-action on S^1_{univ} is faithful.

10.4 The Case of Quasi-Geodesic and Pseudo-Anosov Flows

Let $\mathfrak{F}$ be a flow on the closed hyperbolic 3-manifold M. As we lifted a taut foliation in the 3-manifold to the covering foliation of the universal cover, we can consider the lifted flow $\widetilde{\mathfrak{F}}$ in the universal cover of M. We say that $\mathfrak{F}$ is a quasi-geodesic flow if each flow line of $\widetilde{\mathfrak{F}}$ is a quasi-geodesic in $\widetilde{M}$ which can be identified with the hyperbolic 3-space $\mathbb{H}^3$.

Pseudo-Anosov flows form another important class of flows in 3-manifolds. A flow $\mathscr{F}$ is pseudo-Anosov if it locally looks like a branched covering of an Anosov flow.

These two notions are closely related. First, Steven Frankel [18] announced the resolution of Calegari's flow conjecture which says that any quasi-geodesic flow on a closed hyperbolic 3-manifold can be deformed to a flow that is both quasi-geodesic and pseudo-Anosov. On the other hand, not every pseudo-Anosov flow is quasi-geodesic. Fenley [14] constructed a large class of Anosov flows in hyperbolic 3-manifolds which are not quasi-geodesic. Later he gave a necessary and sufficient condition for a pseudo-Anosov flow to be quasi-geodesic in [17]. These are optimal results.

Calegari [9] showed that if M admits a quasi-geodesic flow, then $\pi_1(M)$ acts faithfully on the circle where the circle is the boundary of the group-equivariant compactification of the space of flow lines of the covering flow $\widetilde{\mathfrak{F}}$. In some sense, the work of Ghys in [22] is a prototype of the result of Calegari. Roughly speaking, Ghys proved that the leaf space of the weak stable foliation of an Anosov flow on a circle bundle is a line, and established a map from the leaf space into the circle.

On the other hand, Calegari-Dunfield [11] showed the same result in the case M admits a pseudo-Anosov flow. Hence, modulo Frankel's upcoming paper, the construction of the action on the circle for quasi-geodesic flows can be reduced to the one for pseudo-Anosov flows. In this section, we will review the work of Calegari-Dunfield for the 3-manifolds admitting a pseudo-Anosov flow.

As shown in the seminal paper of Cannon and Thurston [13], the suspension flow of hyperbolic mapping tori can be chosen to be both quasi-geodesic and pseudo-Anosov. They used this to show that lifts of surface fibers of a fibered hyperbolic 3-manifold extend continuously to the ideal boundary of $\widetilde{M}$ (therefore their boundaries give group-equivariant surjections from S^1 to S^2, which are commonly called Cannon–Thurston maps). This was later generalized by Fenley [16]. Hence, it might be instructive to consider the suspension flows when we think of a pseudo-Anosov flow. In the case of a suspension flow for a hyperbolic mapping torus M, one can consider the suspension of stable and unstable singular measured foliations on the surface for the monodromy to obtain two-dimensional stable and unstable singular foliations in M. Analogously, in the case of a general pseudo-Anosov flow, M has two-dimensional stable and unstable singular foliations.

Let $\mathscr{F}^u$ be the unstable foliation in M for a pseudo-Anosov flow $\mathfrak{F}$. $\mathscr{F}^u$ can be split open to a lamination Λ. Λ can be obtained from $\mathscr{F}^u$ by first removing the singular leaves, and for each singular leaf removed, we insert a finite-sided ideal polygon bundle over the circle so that the leaves of Λ are precisely the nonsingular leaves of $\mathscr{F}^u$ together with one leaf for every face of a singular leaf of $\mathscr{F}^u$. Just like in the case of the taut foliations, one can consider the lifted lamination $\widetilde{\Lambda}$ in $\widetilde{M}$ and the leaf space L of $\widetilde{\Lambda}$. One caution here is that a vertex in L is either a non-boundary leaf or a closed complementary region of $\widetilde{\Lambda}$. Since a complementary region comes from a singular leaf, it is natural to identify the whole thing as a single point in the space of leaves.

At a point in L, it does not locally look like an open interval of the real line, but instead each point of L has a neighborhood which is totally orderable, and between any two points, there exists a unique path which is a concatenation of such orderable segments. This structure is called an *order tree*.

One of the key statements in [11] is the following:

Theorem 10.4.1 (Calegari-Dunfield [11]) *Let M be a closed hyperbolic 3-manifold. If M admits a very full lamination with orderable cataclysms, then $\pi_1(M)$ acts faithfully on the circle by orientation-preserving homeomorphisms.*

Sketch of the Proof We remark that Calegari-Dunfield showed a stronger result by weakening the assumption that the lamination is very full. They allowed the complementary regions of the lamination to be so-called solid torus guts, and in that case, it is shown that one can fill in the lamination with additional leaves to get a very full lamination while preserving many nice properties.

As we explained above, the laminations we obtain from pseudo-Anosov flows (including the stable and unstable laminations in the hyperbolic mapping tori) are very full which means that each complementary region is a finite-sided ideal polygon bundle over the circle. To see how this condition is used, we first fix orientations on the core curves of the complementary regions of Λ. This determines a natural cyclic order on the faces of each cataclysm, hence gives a natural cyclic order on the set of segments sharing exactly one vertex. This order is $\pi_1(M)$-invariant by construction.

The second condition of having orderable cataclysms means that there exists an ordering on each cataclysm which is invariant under the action of the stabilizer of the cataclysm in $\pi_1(M)$. A set of segments of L which differ only by a single vertex correspond to a cataclysm, so they also have natural ordering which is $\pi_1(M)$-invariant by definition of orderable cataclysms.

In summary, a set of segments of L which share exactly one vertex are cyclically ordered and a set of segments of L which differ only at a vertex are linearly ordered. Furthermore, these orderings are $\pi_1(M)$-invariant. From this data, one can realize L as a "planar order tree". There are three types of points in L: first a cataclysm point, i.e., a point corresponding to a leaf in a cataclysm, second a singular point which corresponds to a closed complementary region, and finally an ordinary point which belongs to none of the previous two cases. Let p be an arbitrary point in L. To be concrete, let us assume p is an ordinary point. Draw the point p as an arbitrary point in $\mathbb{R}^2$, maybe the origin, and let I be the orderable segment containing p where endpoints are either cataclysm points or singular points but any other points are ordinary points. If an endpoint is singular, one can draw the incident segments so that the cyclic order on them matches with the cyclic order on their realization inherited from the plane. If an endpoint is a cataclysm point, again one can draw the other segments "incident" at the cataclysm with respect to the linear order on them. Continuing this process, we can realize L as an order tree on the plane.

Let e_1, e_2, e_3 be three distinct ends of L. Pick a point p in L and let r_i be the ray from p to e_i for $i = 1, 2, 3$. Since e_1, e_2, e_3 are all distinct, the r_i's must get separated at some point, and form a subtree of L. Based on our realization of L on $\mathbb{R}^2$, the rays r_i are naturally cyclically ordered, which gives a cyclic ordering on the triple (e_1, e_2, e_3). Note that the ordering on the triple (e_1, e_2, e_3) does not depend on the choice of p.

This defines a cyclic ordering on the set E of ends of L, and by construction, it is $\pi_1(M)$-invariant. Hence we obtained a cyclically ordered set E where $\pi_1(M)$ acts by order-preserving maps. E is equipped with the topology determined by its order: for $e \in E$, the sets $\{x \in E \setminus \{a, b\} | (b, x, a) \text{ is positively oriented}\}$ for some $a, b \in E$ where (a, e, b) is positively oriented form a basis for the topology on E. Then there exists a unique continuous order-preserving embedding of E into S^1 up to homeomorphisms. By collapsing each connected component of the complement of the closure of the image of E, we get a circle where $\pi_1(M)$ acts by orientation-preserving homeomorphisms. Here the circle is obtained as the order-completion of E, and we will denote it as $\overline{E}$.

Suppose a nontrivial element α of $\pi_1(M)$ acts trivially on this circle. For each complementary region of $\widetilde{\Lambda}$, let p be the vertex of L corresponding to the complementary region. Consider all infinite rays in L starting at p; this defines a subset of E. The fact that α fixes this set implies that α actually fixes p. In other words, when we consider the action of α on $\widetilde{M}$, it preserves the given complementary region. Hence, all complementary regions are preserved by α. Each complementary region of $\widetilde{\Lambda}$ is a $\mathbb{Z}$-cover of a complementary region of Λ. Hence if α preserves a complementary region of $\widetilde{\Lambda}$, then it admits an invariant quasi-geodesic. If α preserves another complementary region, α would admit another quasi-geodesic axis who endpoints are disjoint from the one we already had, a contradiction. We have shown that the $\pi_1(M)$-action on the circle constructed above is faithful. □

To apply the above theorem to our case, it remains to see that our lamination Λ has orderable cataclysms. This observation is due to Fenley [15]. Note that each leaf of $\widetilde{\Lambda}$ is foliated by the flow lines of $\widetilde{\mathfrak{F}^u}$ contained in that leaf. Whenever we talk about the foliation on a leaf, we refer to this foliation coming from $\widetilde{\mathfrak{F}^u}$. Let $\{\lambda_j\}$ be (an equivalence class of) a cataclysm and let (μ_i) be a monotone ordered sequence of nonsingular leaves of $\widetilde{\Lambda}$ converging to $\{\lambda_j\}$ on compact subsets of $\widetilde{M}$. For each j, choose a sequence of points $p_{ij} \in \mu_i$ so that p_{ij} converges to a point q_j in λ_j as i tends to ∞.

Candel's theorem again applies here: M admits a metric so that each μ_i is isometric to $\mathbb{H}^2$. Then the foliation on μ_i from $\widetilde{\Lambda}$ is a foliation by bi-infinite geodesics which all share one endpoint (this is an unstable lamination so the flow lines are oriented so that they flow from this common endpoint). Hence the leaf space of the foliation on each μ_i is $\mathbb{R}$, hence naturally totally ordered. The set $\{p_{ij}\}$

of points on μ_i has a natural order on the indices j with respect to this order. For each j, we can take a small product chart U_j around q_j. For all large enough i, the plaque P_j obtained as the intersection $U_j \cap \mu_i$ contains p_{ij} and P_j converges to $U_j \cap \lambda_j$ as foliated disks. Hence, the order relation between p_{ij} and $p_{ij'}$ remains the same for all sufficiently large i. Hence, this gives an ordering on the set $\{\widetilde{q_j}\}$ which can be used as an ordering on the cataclysm $\{\lambda_j\}$. Since the flow lines of $\widetilde{\mathscr{F}^u}$ are preserved under the $\pi_1(M)$-action, our ordering on the cataclysm is invariant under the action of its stabilizer in $\pi_1(M)$. Hence, the unstable lamination for a pseudo-Anosov flow has orderable cataclysms so the above theorem applies. We finally obtain

Theorem 10.4.2 (Calegari-Dunfield [11]) *Let M be a closed hyperbolic 3-manifold which admits a pseudo-Anosov flow. Then $\pi_1(M)$ acts faithfully on the circle by orientation-preserving homeomorphisms.*

10.5 Invariant Laminations for the Universal Circles and Laminar Groups

A lamination Λ on S^1 is defined to be a closed subset of the set of all unordered pairs of two distinct points of S^1 so that any two elements are unlinked. Recall that two pairs (a, b) and (c, d) of points of the circle are unlinked if both a, b are contained in the closure of a single connected component of $S^1 \setminus \{c, d\}$. Note that if $a = c$ and $b \neq d$, the pairs (a, b) and (c, d) are still unlinked according to our definition.

One can visualize Λ by identifying the circle with the ideal boundary of $\mathbb{H}^2$ and then realizing each element as the endpoints of a bi-infinite geodesic. We call this geodesic lamination a *geometric realization* of Λ. Since the geometric realization is unique up to isotopy, we will freely go back and forth between a lamination on the circle and its geometric realization to discuss its properties.

We first consider the case that M is a closed hyperbolic 3-manifold and $\mathscr{F}$ is a co-orientable taut foliation with a branching. In Sect. 10.3, we saw that there exists a set of special sections which has a $\pi_1(M)$-invariant cyclic order and it can be completed to get a universal circle S^1_{univ} where $\pi_1(M)$ acts faithfully by orientation-preserving homeomorphisms.

Now we see that this action preserves laminations. We will construct a lamination Λ^+ assuming the leaf space L is branched in the forward direction. In the case L has a branching in the backward direction, one can construct another lamination Λ^- in a completely analogous way. For each leaf λ in L, let $L^+(\lambda)$ denote the connected component of $L \setminus \{\lambda\}$ containing at least one leaf μ with $\mu > \lambda$. For a subset X of L, we say $core(X)$ is the union of the cores associated with the leaves in X. Let $\Lambda(core(X))$ be the boundary of the convex hull of the closure of $core(X)$ in $\mathbb{H}^2$. Finally, define $\Lambda^+(\lambda)$ to be $\Lambda(core(L^+(\lambda)))$, and Λ^+ to be the closure of the union $\cup_{\lambda \in L} \Lambda^+(\lambda)$. Note that Λ^+ is completely determined by the structure of L.

To see that this is indeed a lamination, we need to show that for $\lambda, \mu \in L$, no leaf of $\Lambda^+(\lambda)$ is linked with a leaf of $\Lambda^+(\mu)$. This is easy to see when λ, μ are comparable, since one of $\Lambda^+(\lambda)$ and $\Lambda^+(\mu)$ is contained in the other. When they are incomparable, there are two cases. One case is that $\lambda \notin \Lambda^+(\mu)$ and $\mu \notin \Lambda^+(\lambda)$. In this case, $\Lambda^+(\mu)$ and $\Lambda^+(\lambda)$ are disjoint, so this is again straightforward. Finally, let us assume that $\lambda \in \Lambda^+(\mu)$ and $\mu \in \Lambda^+(\lambda)$. In this case, $\Lambda^+(\lambda) \cup \Lambda^+(\mu) = L$. Hence $core(L) = core(\Lambda^+(\lambda)) \cup core(\Lambda^+(\mu))$, so the boundaries of the convex hulls do not cross in $\mathbb{H}^2$.

Up to here, we did not really need to assume that L is branched in the forward direction. To see Λ^+ is nonempty, we need this assumption. From the assumption that L has a branching in the forward direction, there exist leaves μ, λ so that $\lambda \notin \Lambda^+(\mu)$ and $\mu \notin \Lambda^+(\lambda)$. As we noted above, $\Lambda^+(\mu)$ and $\Lambda^+(\lambda)$ are disjoint, so their cores are unlinked. In particular, $core(\Lambda^+(\lambda))$ is not dense in S^1_{univ}, which is sufficient to conclude that Λ^+ is nonempty.

Now we get an invariant lamination for the universal circle action for the pseudo-Anosov flow. Let us consider the setup of Sect. 10.4. Let $p_1, \ldots, p_k$ be points in L corresponding to a set of representative of orbits of cataclysm points under the $\pi_1(M)$-action. Say each p_i corresponds to a complementary region which is an ideal n_i-gon bundle over the circle. Then $L \setminus \{p_i\}$ consists of n_i subtrees of L. Choose $q_1, \ldots, q_{n_i}$ on $\overline{E}$ which separate the ends of distinct subtrees of $L \setminus \{p_i\}$. We may assume that the n_i-tuple $(q_1, \ldots, q_{n_i})$ is positively oriented with respect to the cyclic order on $\overline{E}$. Then we consider the pairs (q_j, q_{j+1}) for each $j = 1, \ldots, n_i$ mod $n_i + 1$. We do this for each p_i and take the union of $\pi_1(M)$-orbits of all those pairs, and call it Λ. This process can be done so that elements of Λ are pairwise unlinked. By taking a closure of Λ in the space of unordered pairs of points of $\overline{E}$, we get a $\pi_1(M)$-invariant lamination.

In summary,

Theorem 10.5.1 *Let M be a closed hyperbolic 3-manifold with either a taut foliation, a quasi-geodesic flow or a pseudo-Anosov flow. Then $\pi_1(M)$ acts faithfully on the circle by orientation-preserving homeomorphisms with an invariant lamination.*

From this result, it is natural to ask if a group acting faithfully on the circle by orientation-preserving homeomorphisms with invariant laminations has any interesting property. We call such a group a *laminar group*.

One might first wonder whether there are some natural examples of laminar groups other than the 3-manifold groups we have seen. In fact, all surface groups are laminar groups. Let S_g be a closed connected orientable surface of genus $g \geq 2$, and fix a hyperbolic metric on S_g. The deck group action of $\pi_1(S_g)$ on $\mathbb{H}^2$ extends to an action on $\partial_\infty \mathbb{H}^2$ by homeomorphisms. In this case, any geodesic lamination on S_g defines a lamination on $\partial_\infty \mathbb{H}^2$ which is $\pi_1(S_g)$-invariant. In this case, one can easily construct infinitely many invariant laminations with a lot of structures.

A lamination Λ on S^1 is called *very full* if when it is realized as a geodesic lamination on $\mathbb{H}^2$ via an arbitrary identification of S^1 with $\partial_\infty \mathbb{H}^2$, all the complementary regions are finite-sided ideal polygons. For later use, let us call this geodesic lamination on $\mathbb{H}^2$ a geometric realization of Λ. In the case of $\pi_1(S_g)$, there are infinitely many very full invariant laminations on $\partial_\infty \mathbb{H}^2$. One way to get a very full lamination is to start with a pants-decomposition by simple closed geodesics and then decompose each pair of pants into two ideal triangles by three bi-infinite geodesics which spiral toward boundary components. Then all complementary regions of the resulting lamination are ideal triangles. Since there are infinitely many different pants-decompositions, we get infinitely many different very full invariant laminations. In fact, this argument can be easily generalized to any (complete) hyperbolic surface except the three-punctured sphere, even the ones with infinite area.

In [3], the first author showed that this is actually the characterizing property for hyperbolic surface groups. In fact, we only need three invariant laminations instead of infinitely many invariant laminations. Roughly speaking, a group acting faithfully on the circle acts like a hyperbolic surface group if and only if it admits three different very full invariant laminations. Via an arbitrary identification of S^1 with $\partial_\infty \mathbb{H}^2$, we always identify $\mathrm{PSL}_2(\mathbb{R})$ with a subgroup of $\mathrm{Homeo}^+ S^1$. A precise version of this theorem is the following:

Theorem 10.5.2 (Baik [3]) *Let $G < \mathrm{Homeo}^+ S^1$ be a torsion-free discrete subgroup. Then the followings are equivalent:*

- *G is conjugated into $\mathrm{PSL}_2(\mathbb{R})$ by an element of $\mathrm{Homeo}^+ S^1$ so that $\mathbb{H}^2/G$ is not a three-punctured sphere.*
- *G admits three very full invariant laminations $\Lambda_1, \Lambda_2, \Lambda_3$ where a point p of S^1 is a common endpoint of leaves from Λ_i and Λ_j for $i \neq j$ if and only if it is a cusp point of G (i.e., a fixed point of a parabolic element).*

One can deduce the following simplified version immediately from the above theorem.

Corollary 10.5.3 (Characterization of Cusp-Free Hyperbolic Surface Groups) *Let $G < \mathrm{Homeo}^+ S^1$ be a torsion-free discrete subgroup. Then G is conjugate into $\mathrm{PSL}_2(\mathbb{R})$ by an element of $\mathrm{Homeo}^+ S^1$ so that $\mathbb{H}^2/G$ has no cusps if and only if G admits three very full invariant laminations $\Lambda_1, \Lambda_2, \Lambda_3$ so that leaves from Λ_i and Λ_j with $i \neq j$ do not share an endpoint.*

The proof is pretty long so we do not try to recall it here, but we would like to talk about some key ingredients. One very important observation on the very full lamination Λ is that each point p in S^1 which is not an endpoint of any leaf of Λ has a nested sequence of neighborhoods (I_j) so that I_j shrinks to p and for each j there exists a leaf ℓ_j of Λ whose endpoints are precisely the endpoints of I_j. Such a sequence of leaves (ℓ_j) is called a *rainbow* at p. In short,

Lemma 10.5.4 (Baik [3]) *Let Λ be a very full lamination on S^1. For each $p \in S^1$, either there exists a leaf of Λ which has p as an endpoint, or there exists a rainbow at p.*

Another key ingredient is actually a big hammer called convergence group theorem. Let G be a group acting on a compactum X. We say that the G-action is a *convergence group action* if the induced diagonal action of G on $X \times X \times X - \Delta$ where Δ is the big diagonal is properly discontinuous.

Theorem 10.5.5 (Convergence Group Theorem (Gabai), (Casson–Jungreis), (Tukia), (Hinkkanen), ...) *Suppose a group G acts on S^1 as a convergence group. Then G is conjugate into* $\mathrm{PSL}_2(\mathbb{R})$.

Due to this remarkable theorem, one only needs to check that if G admits three very full laminations, then G acts on S^1 as a convergence group. Suppose not. By definition, this means that there exist a sequence $((x_i, y_i, z_i))$ of three distinct points in S^1 and a sequence (g_i) of elements of G such that $(x_i, y_i, z_i) \to (x_\infty, y_\infty, z_\infty)$ and $(g_i x_i, g_i y_i, g_i z_i) \to (x'_\infty, y'_\infty, z'_\infty)$ where $(x_\infty, y_\infty, z_\infty)$ and $(x'_\infty, y'_\infty, z'_\infty)$ are triples of distinct points in S^1. One can then check that for various possibilities for $x_\infty, y_\infty, z_\infty, x'_\infty, y'_\infty, z'_\infty$ which are either an endpoint of leaves or having rainbows in each Λ_i, each case cannot happen by finding a leaf which is forced to be mapped to a pair which is linked to the given leaf. For details, consult [3].

10.6 Basic Notions and Notation to Study the Group Action on the Circle

So far we have provided a brief review of previously known results. Starting from this section, we now move toward some recent results on this topic. First we need to set up the notation.

Let S^1 be the multiplicative topological subgroup of $\mathbb{C}$ defined as

$$S^1 = \{z \in \mathbb{C} : |z| = 1\}.$$

So far we have used the term cyclic order, but from now on, we will call it a circular order, since it is more suitable for the context. To give more precise definitions, let us consider the stereographic projection $p : S^1 \setminus \{1\} \to \mathbb{R}$ defined as

$$p(z) = \frac{\mathrm{Im}(z)}{\mathrm{Re}(z) - 1}.$$

Obviously, p is a homeomorphism with respect to the standard topologies. For our convenience, we define the degenerate set $\Delta_n(G)$ of a set G to be the set

$$\Delta_n(G) = \{(g_1, \cdots, g_n) \in G^n : g_i = g_j \; for\ some\ i \neq j\}$$

of all n-tuples with some repeated elements.

Definition 10.6.1 For $n \geq 3$, an element $(x_1, \cdots, x_n)$ in $(S^1)^n - \Delta_n(S^1)$ is a *positively oriented n-tuple* on S^1 if for each $i \in \{2, \cdots, n-1\}$, $p(x_1^{-1} x_i) < p(x_1^{-1} x_{i+1})$. An element $(x_1, \cdots, x_n)$ in $(S^1)^n - \Delta_n(S^1)$ is a *negatively oriented n-tuple* on S^1 if for each $i \in \{2, \cdots, n-1\}$, $p(x_1^{-1} x_{i+1}) < p(x_1^{-1} x_i)$.

We use the definition of circular order in the following form.

Definition 10.6.2 *A circular order* on a set G is a map $\varphi : G^3 \to \{-1, 0, 1\} \subseteq \mathbb{Z}$ with the following properties:

(DV) φ kills precisely the degenerate set, i.e.

$$\varphi^{-1}(0) = \Delta_3(G).$$

(C) φ is a 2-cocycle, i.e.

$$\varphi(g_1, g_2, g_3) - \varphi(g_0, g_2, g_3) + \varphi(g_0, g_1, g_3) - \varphi(g_0, g_1, g_2) = 0$$

for all $g_0, g_1, g_2, g_3 \in G$.

Furthermore, if G is a group, then a *left-invariant circular order* on G is a circular order on G as set that also satisfies the homogeneity property:

(H) φ is homogeneous, i.e.

$$\varphi(g_0, g_1, g_2) = \varphi(hg_0, hg_1, hg_2)$$

for all $h \in G$ and $(g_0, g_1, g_2) \in G^3$.

By abuse of language, we will refer to a “left-invariant circular order of a group” simply as a “circular order.” To learn about invariant circular orders of groups, see [4].

Let us define a circular order φ on the multiplicative group S^1 in the following way. For $p \in \Delta_3(S^1)$, $\varphi(p) = 0$. For $p \in (S^1)^3 - \Delta_3(S^1)$, we assign $\varphi(p) = 1$ if p is positively oriented in S^1, and $\varphi(p) = -1$ if p is negatively oriented in S^1. We can easily check that φ is a circular order of the group S^1.

We also set up terminologies for intervals in S^1. The reason is that in the rest of the chapter, we will reformulate laminations on S^1 as sets of intervals with certain conditions. Using this new perspective, we will give a detailed discussion of laminar groups. First, we call a nonempty proper connected open subset of S^1 an *open interval* in S^1. Technically, we distinguish the following two cases.

Definition 10.6.3 Let u, v be two elements of S^1.

(1) If $u \neq v$, $(u, v)_{S^1}$ is the set

$$(u, v)_{S^1} = \{p \in S^1 : \varphi(u, p, v) = 1\}.$$

We call it a *nondegenerate open interval* in S^1.

(2) If $u = v$, $(u, v)_{S^1}$ is the set

$$(u, v)_{S^1} = S^1 - \{u\}.$$

We call it a *degenerate open interval* in S^1.

If $(u, v)_{S^1}$ is a nondegenerate open interval, then we denote $(v, u)_{S^1}$ by $(u, v)^*_{S^1}$, and call it the *dual interval* of $(u, v)_{S^1}$.

We can check that the set of nondegenerate open intervals of S^1 is a base for the topology of S^1 which is induced from the standard topology of $\mathbb{C}$. For convenience, we also use the following notation. Let $(u, v)_{S^1}$ be a nondegenerate open interval. Then, we denote

(1) for $z \in S^1$, $z(u, v)_{S^1} = (zu, zv)_{S^1}$,
(2) $[u, v)_{S_1} = \{u\} \cup (u, v)_{S^1}$,
(3) $(u, v]_{S^1} = (u, v)_{S^1} \cup \{v\}$,
(4) $[u, v]_{S^1} = (u, v)_{S^1} \cup \{u, v\}$.

We can derive the following list of properties about dual intervals. Let I and J be two nondegenerate open intervals.

(1) $(I^*)^* = I$,
(2) $I^c = \overline{I^*}$,
(3) If $I \subseteq J$, then $J^* \subseteq I^*$,
(4) If $I \cap J = \emptyset$, then $I \subseteq J^*$,
(5) If $I \cap J = \emptyset$, then $|\bar{I} \cap \bar{J}| \leq 2$,

where $\bar{I}$ is the closure of I. Recall that every open subset of $\mathbb{R}$ can be obtained by a countable disjoint union of open intervals of $\mathbb{R}$. The same is true for S^1.

Proposition 10.6.4 *Every proper open set of S^1 is an at most countable union of disjoint open intervals.*

10.7 Lamination Systems on S^1 and Laminar Groups

Using the notation and terminology defined in the previous section, we introduce the notion of lamination systems on S^1. This is a set of intervals in S^1 with certain conditions which corresponds to leaves of our usual notion of a lamination on S^1. Before defining lamination systems, we need to introduce a condition which is analogous to the unlinkedness in laminations on S^1.

Definition 10.7.1 Let I and J be two nondegenerate open intervals. If $I \subseteq J$ or $I^* \subseteq J$, then we say that the two points set $\{I, I^*\}$ *lies on* J (see Fig. 10.3). If $\bar{I} \subseteq J$ or $\overline{I^*} \subseteq J$, then we say that the two points set $\{I, I^*\}$ *properly lies on* J.

Let us define lamination systems.

Definition 10.7.2 Let $\mathscr{L}$ be a nonempty family of nondegenerate open intervals of S^1. $\mathscr{L}$ is called a *lamination system* on S^1 if it satisfies the following three properties:

(1) If $I \in \mathscr{L}$, then $I^* \in \mathscr{L}$.
(2) For any $I, J \in \mathscr{L}$, $\{I, I^*\}$ lies on J or J^*.

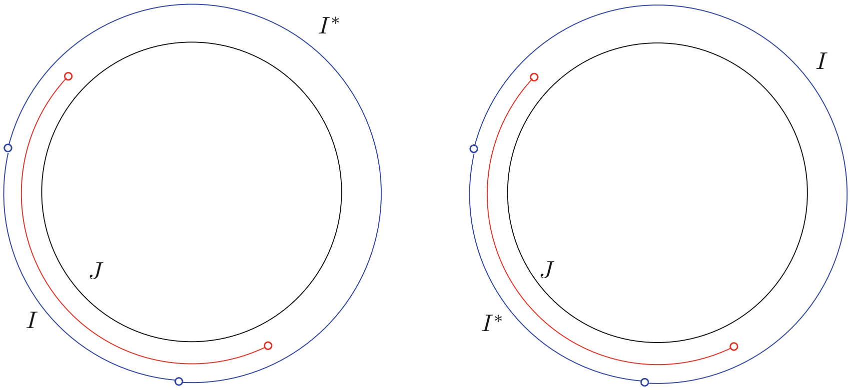


Fig. 10.3 The red segment represents the nondegenerate open interval J and the blue parts represent I and I^*. Two figures show all possible cases where $\{I, I^*\}$ lies on J

(3) If there is a sequence $\{I_n\}_{n=1}^{\infty}$ on $\mathscr{L}$ such that for $n \in \mathbb{N}$, $I_n \subseteq I_{n+1}$, and $\bigcup_{n=1}^{\infty} I_n$ is a nondegenerate open interval in S^1, then $\bigcup_{n=1}^{\infty} I_n \in \mathscr{L}$.

The original definition of laminations on S^1 is a closed subset of the set of all pairs in S^1 with unlinkedness condition. In a lamination system, each two points set corresponds to the set of two connected components of the complement of the two points.

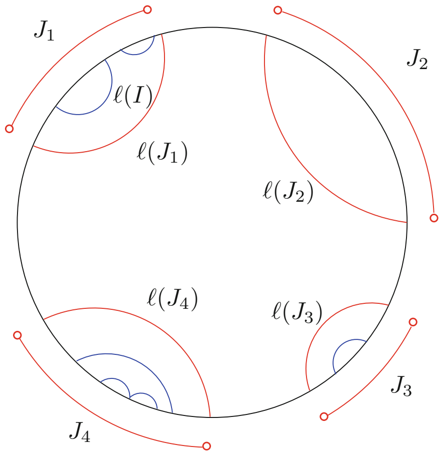
In this sense, we define leaves and gaps on a lamination system $\mathscr{L}$ as following. A subset $\mathscr{G}$ of $\mathscr{L}$ is a *leaf* of $\mathscr{L}$ if $\mathscr{G} = \{I, I^*\}$ for some $I \in \mathscr{L}$. We denote such a leaf $\mathscr{G}$ by $\ell(I)$. With this definition of leaf, we can see that the second condition of lamination system implies unlinkedness of leaves of laminations of S^1. Likewise, a subset $\mathscr{G}$ of $\mathscr{L}$ is a *gap* of $\mathscr{L}$ if $\mathscr{G}$ satisfies the following two conditions:

(1) Elements of $\mathscr{G}$ are disjoint.
(2) For each $I \in \mathscr{L}$, there is a J in $\mathscr{G}$ on which $\ell(I)$ lies (see Fig. 10.4).

By the second condition on gaps, every gap is nonempty. Obviously, a leaf is also a gap with two elements. So we say that a gap is a *non-leaf gap* if it is not a leaf. Then, we denote $S^1 - \bigcup_{I \in \mathscr{G}} I$ as $v(\mathscr{G})$ and call it a *vertex set* of $\mathscr{G}$ or an *end points set* of $\mathscr{G}$. Each element of a vertex set is called a *vertex* or an *end point*. Note that in general, a vertex set need not be a discrete subset of S^1. Geometrically, the convex hull $conv(v(\mathscr{G}))$ of $v(\mathscr{G})$ in $\overline{\mathbb{H}^2}$ is the geometric realization of a gap $\mathscr{G}$.

The third condition on lamination systems is analogous to the closedness of laminations on S^1. From now on, to describe the limit of a sequence of leaves, we define the notion of convergence of a sequence of leaves.

Fig. 10.4 The red chords on the disk are the geometric realization of $\ell(J_i)$. In this figure, the geodesic lamination is the union of red and blue chords and a gap is $\{J_1, J_2, J_3, J_4\}$. Note that any blue chord $\ell(I)$ lies on J_i



Definition 10.7.3 Let $\mathscr{L}$ be a lamination system, and $\{\ell_n\}_{n=1}^{\infty}$ be a sequence of leaves on $\mathscr{L}$. Let J be a nondegenerate open interval. We say that $\{\ell_n\}_{n=1}^{\infty}$ *converges to* J if there is a sequence $\{I_n\}_{n=1}^{\infty}$ on $\mathscr{L}$ such that for each $n \in \mathbb{N}$, $\ell_n = \ell(I_n)$, and

$$J \subseteq \liminf I_n \subseteq \limsup I_n \subseteq \overline{J}.$$

We denote this by $\ell_n \to J$.

This definition is symmetric in the following sense.

Proposition 10.7.4 *Let $\mathscr{L}$ be a lamination system and $\{\ell_n\}_{n=1}$ be a sequence of leaves on $\mathscr{L}$. Let J be a nondegenerate open interval. Suppose that there is a sequence $\{I_n\}_{n=1}^{\infty}$ on $\mathscr{L}$ such that for each $n \in \mathbb{N}$, $\ell_n = \ell(I_n)$ and*

$$J \subseteq \liminf I_n \subseteq \limsup I_n \subseteq \overline{J}.$$

Then

$$J^* \subseteq \liminf I_n^* \subseteq \limsup I_n^* \subseteq \overline{J^*}.$$

Proof Since $J \subseteq \liminf I_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} I_n \subseteq \overline{J}$, so $J^* = \overline{J}^c \subseteq \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} I_n^c \subseteq J^c = \overline{J^*}$. So, $\limsup I_n^* = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} I_n^* \subseteq \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \overline{I_n^*} = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} I_n^c \subseteq J^c = \overline{J^*}$. It remains to show that $J^* \subseteq \liminf I_n^*$. Since $J \subseteq \limsup I_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} I_n \subseteq \overline{J}$, so

$J^* = \overline{J}^c \subseteq \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} I_n^c = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \overline{I_n^*} \subseteq J^c = \overline{J^*}$. Denote $J^* = (u, v)_{S^1}$ and choose $w \in J$. For each $n \in \mathbb{N}$, define $(u_n, v_n)_{S^1}$ as follows:

$$(u_n, v_n)_{S^1} = wp^{-1}((p(w^{-1}u) + \frac{L}{3n}, p(w^{-1}v) - \frac{L}{3n}))$$

where p is the stereographic projection and $L = p(w^{-1}v) - p(w^{-1}u)$. Then for all $n \in \mathbb{N}$, $[u_n, v_n]_{S^1} \subseteq (u_{n+1}, v_{n+1})_{S^1} \subseteq J^*$.

Fix $m \in \mathbb{N}$. Since $[u_m, v_m]_{S^1} \subseteq J^* \subseteq \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \overline{I_n^*}$, there is a natural number N_m such that $\{u_m, v_m\} \subseteq \bigcap_{n=N_m}^{\infty} \overline{I_n^*}$. Note that since $\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \overline{I_n^*} \subseteq \overline{J^*}$, there is a natural number N such that $w \notin \bigcup_{n=N}^{\infty} \overline{I_n^*}$. Let $M_m = \max\{N, N_m\}$. Then for all $k \geq M_m$, $\{u_m, v_m\} \subseteq \overline{I_k^*}$ and $w \notin \overline{I_k^*}$.

From now on, we show that for all $k \geq M_m$, $(u_m, v_m)_{S^1} \subseteq I_k^*$. Fix $k \geq M_m$ and denote $I_k^* = (a_k, b_k)_{S^1}$. Note that since $[u_m, v_m]_{S^1} \subseteq J^*$ and $w \in J$, $w \notin [u_m, v_m]_{S^1}$. If $\{a_k, b_k\} = \{u_m, v_m\}$, then $(a_k, b_k)_{S^1} = (u_m, v_m)_{S^1}$ since $w \notin [u_m, v_m]_{S^1}$. If not, there is an element $v \in \{u_m, v_m\} - \{a_k, b_k\}$. First, consider the case $v = u_m$. Since $\{u_m, v_m\} \subseteq \overline{I_k^*}$, it is $v_m \in [a_k, u_m)_{S^1}$ or $v_m \in (u_m, b_k]_{S^1}$. If $v_m \in [a_k, u_m)_{S^1}$, then $(v_m, u_m)_{S^1} \subseteq [a_k, u_m)_{S^1} \subseteq [a_k, b_k]_{S^1}$. However, this is a contradiction since $w \in (v_m, u_m)_{S^1}$ and $w \notin [a_k, b_k]_{S^1}$. Therefore, $v_m \in (u_m, b_k]_{S^1}$ and so $(u_m, v_m)_{S^1} \subseteq (u_m, b_k]_{S^1} \subseteq [a_k, b_k]_{S^1}$. Thus, $(u_m, v_m)_{S^1} \subseteq (a_k, b_k)_{S^1} = I_k^*$. Similarly, we can prove the case $v = v_m$.

Therefore, for all $m \in \mathbb{N}$, $(u_m, v_m)_{S^1} \subseteq \bigcap_{n=M_m}^{\infty} I_n^* \subseteq \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} I_n^* = \liminf I_n^*$.

Thus, $J^* = \bigcup_{m=1}^{\infty} (u_m, v_m)_{S^1} \subseteq \liminf I_n^*$. □

Since the third condition on lamination systems guarantees that the limit of an ascending sequence on a lamination system is in the lamination system, we need to consider descending sequences to say about limits of arbitrary sequences on lamination systems. The following lemma implies closedness of descending sequences in a lamination system $\mathcal{L}$.

Lemma 10.7.5 *Let* $\{I_n\}_{n=1}^{\infty}$ *be a sequence on a lamination system* $\mathcal{L}$ *such that* $I_{n+1} \subseteq I_n$ *for all* $n \in \mathbb{N}$, *and* $\bigcup_{n=1}^{\infty} I_n^* = J \in \mathcal{L}$. *Then* $\operatorname{Int}\left(\bigcap_{n=1}^{\infty} I_n\right) = J^* \in \mathcal{L}$.

Proof Since $\bigcup_{n=1}^{\infty} I_n^* = J$, so

$$\overline{J^*} = J^c = \bigcap_{n=1}^{\infty}(I_n^*)^c = \bigcap_{n=1}^{\infty}\overline{I_n}.$$

So, $\bigcap_{n=1}^{\infty} I_n \subseteq \bigcap_{n=1}^{\infty}\overline{I_n} = \overline{J^*}$. Since $\bigcup_{n=1}^{\infty} I_n^* = J$, so for all $n \in \mathbb{N}$, $I_n^* \subseteq J$ and so $J^* \subseteq I_n$. Therefore, $J^* \subseteq \bigcap_{n=1}^{\infty} I_n$. Thus, $J^* \subseteq \bigcap_{n=1}^{\infty} I_n = \overline{J^*}$ and so $\mathrm{Int}\left(\bigcap_{n=1}^{\infty} I_n\right) = J^*$. □

With this lemma, the following proposition shows the closedness of lamination systems.

Proposition 10.7.6 *If a sequence $\{\ell_n\}_{n=1}^{\infty}$ of leaves of a lamination system $\mathscr{L}$ converges to a nondegenerate open interval J, then $J \in \mathscr{L}$.*

Proof Since the sequence $\{\ell_n\}$ converges to J, there is a sequence $\{I_n\}_{n=1}^{\infty}$ on the lamination system $\mathscr{L}$ such that for all $n \in \mathbb{N}$, $\ell_n = \ell(I_n)$ and

$$J \subseteq \liminf I_n \subseteq \limsup I_n \subseteq \overline{J}.$$

Since $\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} I_n \subseteq \overline{J}$, and so $J^* = \overline{J}^c \subseteq \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} I_n^c$, there is $N \in \mathbb{N}$ such that $\bigcap_{n=N}^{\infty} I_n^c \neq \emptyset$. Then we can get $\left(\bigcup_{n=k}^{\infty} I_n\right)^c \neq \emptyset$ for all $k \geq N$ since $\bigcap_{n=k}^{\infty} I_n^c \subseteq \bigcap_{n=k+1}^{\infty} I_n^c$ for all $k \geq N$. On the other hand, since $J \subseteq \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} I_n$, there is $N' \in \mathbb{N}$ such that $\bigcap_{n=N'}^{\infty} I_n \neq \emptyset$. Choose $p \in \bigcap_{n=N'}^{\infty} I_k$, and set $M = \max\{N, N'\}$. Since for $n \geq M$, I_n is a connected open subset of S^1, and contains p, so for $k \geq M$, $\bigcup_{n=k}^{\infty} I_n$ is nonempty open connected subset of S^1. Since for all $k \geq M$, $\left(\bigcup_{n=k}^{\infty} I_n\right)^c \neq \emptyset$ and so $\bigcup_{n=k}^{\infty} I_n$ is a proper subset of S^1, then for each $k \geq M$, $\bigcup_{n=k}^{\infty} I_n$ is an open interval. If for all $k \geq M$, $\bigcup_{n=k}^{\infty} I_n$ is a degenerate open interval, then $\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} I_n = \bigcap_{k=M}^{\infty}\bigcup_{n=k}^{\infty} I_n$ is a degenerate open interval, but this contradicts $\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} I_n \subseteq \overline{J}$. Therefore, there is $K \in \mathbb{N}$ such that $K \geq M$, and for all $k \geq K$, $\bigcup_{n=k}^{\infty} I_n$ is nondegenerate. Let $L = \bigcup_{n=K}^{\infty} I_n$. Then, for all $n \geq K$, $I_n \subseteq L$ and so $L^* \subseteq I_n^*$. Therefore, we get that for all $n \geq K$, $p \in I_n \subseteq L$.

From now on, we show that for any $n, m \geq K$, $I_n \subseteq I_m$ or $I_m \subseteq I_n$. Choose $n, m \geq K$. If $I_n \subseteq I_m^*$, then $p \in I_m^*$, and so this contradicts $p \in I_m$. If $I_m^* \subseteq I_n$, then $L^* \subseteq I_m^* \subseteq I_n$, and so this contradicts $I_n \subseteq L$. So, we can get what we want.

Then, we show that $\bigcup_{n=k}^{\infty} I_n \in \mathscr{L}$ for all $k \geq K$. Fix $k \geq K$. Let $J_i = \bigcup_{n=k}^{k+i} I_n$ for $i \in \mathbb{N}$. Then $J_1 = I_k \cup I_{k+1}$. Since $I_k \subseteq I_{k+1}$ or $I_{k+1} \subseteq I_k$, so $J_1 = I_k$ or $J_1 = I_{k+1}$, and so $J_1 \in \mathscr{L}$. Assume that $J_m \in \mathscr{L}$ for some $m \in \mathbb{N}$. If $I_{k+m+1} \subseteq I_j$ for some $j \in \{k, k+1, \cdots, k+m\}$, then $J_{m+1} = J_m \in \mathscr{L}$. If not, $I_j \subseteq I_{k+m+1}$ for all $j \in \{k, k+1, \cdots, k+m\}$, and so $J_{m+1} = I_{k+m+1} \in \mathscr{L}$. Therefore, by mathematical induction, $J_i \in \mathscr{L}$ for all $i \in \mathbb{N}$. Moreover, $J_i \subseteq J_{i+1}$ for all $i \in \mathbb{N}$. Since by the condition of K, $\bigcup_{i=1}^{\infty} J_i = \bigcup_{n=k}^{\infty} I_n$ is a nondegenerate open interval and so $\bigcup_{i=1}^{\infty} J_i \in \mathscr{L}$, so $\bigcup_{n=k}^{\infty} I_n \in \mathscr{L}$.

From now on, let $L_i = \bigcup_{n=K+i}^{\infty} I_n$ for $i \in \mathbb{N}$. Then, for all $i \in \mathbb{N}$, $L_i \in \mathscr{L}$ by the above, and $L_{i+1} \subseteq L_i$. Since $J \subseteq \bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} I_n \subseteq \bar{J}$, and $\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} I_n = \bigcap_{k=K+1}^{\infty}\bigcup_{n=k}^{\infty} I_n = \bigcap_{i=1}^{\infty} L_i$, so $J \subseteq \bigcap_{i=1}^{\infty} L_i \subseteq \bar{J}$. So, $J^* \subseteq \bigcup_{i=1}^{\infty} \overline{L_i^*} \subseteq \overline{J^*}$, and so $\bigcup_{i=1}^{\infty} L_i^* \subseteq \bigcup_{i=1}^{\infty} \overline{L_i^*} \subseteq \overline{J^*}$. Therefore, $\bigcup_{i=1}^{\infty} L_i^*$ is nondegenerate. Since for all $i \in \mathbb{N}$, $L_i^* \in \mathscr{L}$, and $L_i^* \subseteq L_{i+1}^*$, $\bigcup_{i=1}^{\infty} L_i^* \in \mathscr{L}$. Thus by Lemma 10.7.5, $\operatorname{Int}\left(\bigcap_{i=1}^{\infty} L_i\right) = \left(\bigcup_{i=1}^{\infty} L_i^*\right)^* \in \mathscr{L}$. Since $J \subseteq \bigcap_{i=1}^{\infty} L_i \subseteq \bar{J}$, $J = \operatorname{Int}\left(\bigcap_{i=1}^{\infty} L_i\right) \in \mathscr{L}$. □

Moreover, by Proposition 10.7.4 we can prove that if $\ell_n \to J$, then $\ell_n \to J^*$, and so $J^* \in \mathscr{L}$. So, we can make the following definition.

Definition 10.7.7 Let $\mathscr{L}$ be a lamination system, and $\{\ell_n\}_{n=1}^{\infty}$ be a sequence of leaves on $\mathscr{L}$. Let ℓ be a leaf of $\mathscr{L}$. Then, we say that $\{\ell_n\}_{n=1}^{\infty}$ *converges to* ℓ if $\ell_n \to I$ for some $I \in \ell$.

So far, we have studied the definition of a lamination system. From now on, we talk about the shape of lamination systems. First, as we can see in the proof of Proposition 10.7.6, the following structure is useful to deal with configuration of leaves.

Definition 10.7.8 Let $\mathscr{L}$ be a lamination system on S^1 and I be a nondegenerate open interval. Then, for $p \in I$, we define C_p^I as the set $C_p^I = \{J \in \mathscr{L} \colon p \in J \subseteq I\}$.

As we observed in the proof of Proposition 10.7.6, C_p^I is totally ordered by the inclusion.

Proposition 10.7.9 *C_p^I is totally ordered by the set inclusion $\subseteq$.*

Proof If C_p^I has at most one element, it is true. Assume that C_p^I has at least two elements. Let J and K be two distinct elements of $\mathscr{L}$. If $J \subseteq K^*$, then $p \in L \subseteq K^*$ and so it is a contradiction since $p \in K$. If $K^* \subseteq J$, then $K^* \subseteq J \subseteq I$, and so $I^* \subseteq K$. But it contradicts $K \subseteq I$. Thus, $J \subseteq K$ or $K \subseteq J$. □

The following lemma tells about the maximal and minimal elements of C_p^I.

Lemma 10.7.10 *Let $\mathscr{L}$ be a lamination system on S^1, and I be a nondegenerate open interval. Let x be an element of I. Assume that C_x^I is nonempty. Then, there is a sequence $\{J_n\}_{n=1}^{\infty}$ on C_x^I such that for all $n \in \mathbb{N}$, $J_n \subseteq J_{n+1}$, and $\bigcup_{n=1}^{\infty} J_n = \bigcup_{K \in C_x^I} K$. Also, there is a sequence $\{K_n\}_{n=1}^{\infty}$ on C_x^I such that for all $n \in \mathbb{N}$, $K_{n+1} \subseteq K_n$, and $\bigcap_{n=1}^{\infty} K_n = \bigcap_{K \in C_x^I} K$.*

Proof First, we show the first statement. Since $x \in K$ for all $K \in C_x^I$, $\bigcup_{K \in C_x^I} K$ is a connected open set. Because $\bigcup_{K \in C_x^I} K \subseteq I$, or equivalently $I^c \subseteq \left(\bigcup_{K \in C_x^I} K\right)^c$, $\bigcup_{K \in C_x^I} K$ is a nondegenerate open interval. So, we can write $\bigcup_{K \in C_x^I} K = (u, v)_{S^1}$ for some $u, v \in S^1$ with $u \neq v$. Choose $z \in (v, u)_{S^1}$. We define a sequence $\{I_n\}_{n=1}^{\infty}$ of nondegenerate intervals as

$$I_n = zp^{-1}((p(z^{-1}u) + \frac{L}{3n}, p(z^{-1}v) - \frac{L}{3n}))$$

where p is the stereographic projection map used in the definition of the orientation, $L = p(z^{-1}v) - p(z^{-1}u)$. Then, for all $n \in \mathbb{N}$, $I_n \subseteq I_{n+1}$ and $\bigcup_{n=1}^{\infty} I_n = (u, v)_{S^1}$.

From now on, we construct a sequence $\{K_n\}_{n=1}^{\infty}$ in C_x^I such that $I_n \subseteq K_n$. For $n \in \mathbb{N}$, we denote $I_n = (p_n, q_n)_{S^1}$. Then since $\partial I_n \subseteq (u, v)_{S^1}$ and $\bigcup_{K \in C_x^I} K = (u, v)_{S^1}$, there are K_{p_n} and K_{q_n} in C_x^I such that $p_n \in K_{p_n}$, and $q_n \in K_{q_n}$. By Proposition 10.7.9, $K_{p_n} \subseteq K_{q_n}$ or $K_{q_n} \subseteq K_{p_n}$. If $K_{q_n} \subseteq K_{p_n}$, then $\partial I_n \subset K_{p_n}$. So $(p_n, q_n)_{S^1} \subseteq K_{p_n}$ or $(q_n, p_n)_{S^1} \subseteq K_{p_n}$. Since $z \notin K_{p_n}$, so $I_n = (p_n, q_n)_{S^1} \subseteq K_{p_n}$. In this case, we set $K_n = K_{p_n}$ Likewise, if $K_{p_n} \subseteq K_{q_n}$, then $I_n \subseteq K_{q_n}$, and so we set $K_n = K_{q_n}$.

For $n \in \mathbb{N}$, we define J_n as $J_n = \bigcup_{m=1}^{n} K_m$. As in the argument in Proposition 10.7.6, $J_n \in C_x^I$ for all $n \in \mathbb{N}$. Then $\{J_n\}_{n=1}^{\infty}$ is a sequence on C_x^I such that for all $n \in \mathbb{N}$, $I_n \subseteq J_n \subseteq J_{n+1} \subseteq (u, v)_{S^1}$. Therefore,

$$\bigcup_{n=1}^{\infty} J_n = \bigcup_{n=1}^{\infty} I_n = (u, v)_{S^1} = \bigcup_{K \in C_x^I} K.$$

The second statement can be also proved in a similar way. Let $A = \bigcap_{K \in C_x^I} K$. Then,

$$A = \bigcap_{K \in C_x^I} K \subseteq \bigcap_{K \in C_x^I} \overline{K}$$

and so

$$\bigcup_{K \in C_x^I} K^* = \bigcup_{K \in C_x^I} \overline{K}^c \subseteq A^c = \bigcup_{K \in C_x^I} K^c = \bigcup_{K \in C_x^I} \overline{K^*} \subseteq \overline{\bigcup_{K \in C_x^I} K^*}.$$

So we get

$$\bigcup_{K \in C_x^I} K^* \subseteq A^c \subseteq \overline{\bigcup_{K \in C_x^I} K^*}.$$

Since for all $K \in C_x^I$, $x \in K \subseteq I$ and so $x \notin K^*$ and $I^* \subseteq K^*$, $\bigcup_{K \in C_x^I} K^*$ is a nonempty proper connected open set and so it is an open interval in S^1. Then, we can write $\bigcup_{K \in C_x^I} K^* = (v, u)_{S^1}$ for some $u, v \in S^1$. Then

$$(v, u)_{S^1} \subseteq A^c \subseteq \overline{(v, u)_{S^1}}$$

and so

$$\overline{(v, u)}^c_{S^1} \subseteq A \subseteq (v, u)^c_{S^1}.$$

Since $(v, u)_{S^1}$ is homeomorphic to $\mathbb{R}$, it is Lindelöff. So, there is a sequence $\{K_n\}_{n=1}^{\infty}$ on C_x^I such that $\bigcup_{n=1}^{\infty} K_n^* = \bigcup_{K \in C_x^I} K^* = (v, u)_{S^1}$ since $\bigcup_{K \in C_x^I} K^*$ is an open cover of $(v, u)_{S^1}$.

If $u = v$, then $\emptyset = \overline{(v, u)}^c_{S^1} \subseteq A \subseteq (v, u)^c_{S^1} = \{u\}$. Since

$$\{x\} \subseteq \bigcap_{n=1}^{\infty} K_n \subseteq \bigcap_{n=1}^{\infty} \overline{K_n} = \bigcap_{n=1}^{\infty} (K_n^*)^c = \left(\bigcup_{n=1}^{\infty} K_n^* \right)^c = (v, u)^c_{S^1} = \{u\},$$

so $\{x\} \subseteq \bigcap_{n=1}^{\infty} K_n \subseteq \{u\}$. Therefore, $\{x\} = \bigcap_{n=1}^{\infty} K_n = \{u\}$. Thus, since $x \in A$ and $A \subseteq \{u\}$, and so $\{x\} = A = \{u\}$, $\bigcap_{n=1}^{\infty} K_n = A = \{x\}$. For each $n \in \mathbb{N}$, define $J_n = \bigcap_{m=1}^{n} K_m$. Then, for all $n \in \mathbb{N}$, $J_n \in C_x^I$ and $J_{n+1} \subseteq J_n$. Thus, since $\bigcap_{n=1}^{\infty} K_n = \bigcap_{n=1}^{\infty} J_n$, the sequence $\{J_n\}_{n=1}^{\infty}$ is the sequence that we want.

If $u \neq v$, then

$$(u, v)_{S^1} = \overline{(v, u)}^c_{S^1} \subseteq A \subseteq (v, u)^c_{S^1} = [u, v]_{S^1}.$$

There are four cases: $A = (u, v)_{S^1}$, $A = [u, v]_{S^1}$, $A = [u, v)_{S^1}$ and $A = (u, v]_{S^1}$.

First, if $A = [u, v]_{S^1}$, then

$$[u, v]_{S^1} = A = \bigcap_{K \in C_x^I} K \subseteq \bigcap_{n=1}^{\infty} K_n \subseteq \bigcap_{n=1}^{\infty} \overline{K_n} = \bigcap_{n=1}^{\infty} (K_n^*)^c = \left(\bigcup_{n=1}^{\infty} K_n^* \right)^c = [u, v]_{S^1}.$$

Therefore,

$$[u, v]_{S^1} = A = \bigcap_{n=1}^{\infty} K_n.$$

Then, for each $n \in \mathbb{N}$, define $J_n = \bigcap_{m=1}^{n} K_m$. By the construction of $\{J_n\}_{n=1}^{\infty}$, for all $n \in \mathbb{N}$, $J_n \in C_x^I$ and $J_{n+1} \subseteq J_n$. Thus, since $\bigcap_{n=1}^{\infty} K_n = \bigcap_{n=1}^{\infty} J_n$, the sequence $\{J_n\}_{n=1}^{\infty}$ is a sequence that we want.

Next, if $A = (u, v)_{S^1}$, then there are K_u and K_v in C_x^I such that $u \notin K_u$ and $v \notin K_v$. Since C_x^I is totally ordered, so $K_u \subseteq K_v$ or $K_v \subseteq K_u$. Therefore, one of K_u and K_v, say K', does not intersect $\{u, v\}$. Since

$$(u, v)_{S^1} = A \subseteq K' \subseteq S^1 - \{u, v\} = (u, v)_{S^1} \cup (v, u)_{S^1}$$

and K' is connected, $K' = (u, v)_{S^1}$. Therefore, $(u, v)_{S^1} \in C_x^I$ and so, for each $n \in \mathbb{N}$, define $J_n = (u, v)_{S^1}$. Then the sequence $\{J_n\}_{n=1}^{\infty}$ is the sequence that we want.

If $A = (u, v]_{S^1}$, then there is an element L in C_x^I such that $u \notin L$. Define $C = \{K \in C_x^I : K \subseteq L\}$. Since $C \subseteq C_x^I$, C is also totally ordered by the inclusion. Note that $A = \bigcap_{K \in C_x^I} K = \bigcap_{K \in C} K$ and $\bigcup_{K \in C_x^I} K^* = \bigcup_{K \in C} K^*$. So, since $\bigcup_{K \in C} K^*$ is a nondegenerate open interval and so is Lindelöff, there is a sequence $\{L_n\}_{n=1}^{\infty}$ of C such that $\bigcup_{K \in C} K^* = \bigcup_{n=1}^{\infty} L_n^*$. Then

$$(u, v]_{S^1} = A = \bigcap_{K \in C} K \subseteq \bigcap_{n=1}^{\infty} L_n \subseteq \bigcap_{n=1}^{\infty} \overline{L_n} = \bigcap_{n=1}^{\infty} (L_n^*)^c = \left(\bigcup_{n=1}^{\infty} L_n^* \right)^c = [u, v]_{S^1}.$$

So we get

$$(u, v]_{S^1} = A \subseteq \bigcap_{n=1}^{\infty} L_n \subseteq [u, v]_{S^1}.$$

Since for all $n \in \mathbb{N}$, $u \notin L_n$, $(u, v]_{S^1} = A = \bigcap_{n=1}^{\infty} L_n$. Then, for each $n \in \mathbb{N}$, define $J_n = \bigcap_{m=1}^{n} L_m$. By the construction of $\{J_n\}_{n=1}^{\infty}$, for all $n \in \mathbb{N}$, $J_n \in C_x^I$ and

$J_{n+1} \subseteq J_n$. Thus, the sequence $\{J_n\}_{n=1}^{\infty}$ is the one that we want. The proof of the case $A = [u, v)_{S^1}$ is similar to the case $A = (u, v]_{S^1}$. □

Note that if $C_p^I \neq \emptyset$, then by Lemma 10.7.10, $\bigcup_{J \in C_p^I} J \in \mathscr{L}$. On a lamination system $\mathscr{L}$, when a sequence of leaves converges to a leaf ℓ, it approaches ℓ in two different sides of ℓ. Geometrically, if there is no converging sequence of leaves on one side, then there is a non-leaf gap on that side. To describe this situation, we use the following definition.

Definition 10.7.11 Let $\mathscr{L}$ be a lamination system on S^1, and $I \in \mathscr{L}$. Let $\{\ell_n\}_{n=1}^{\infty}$ be a sequence of leaves of $\mathscr{L}$. Then we call $\{\ell_n\}_{n=1}^{\infty}$ an *I-side sequence* if for all $n \in \mathbb{N}$, $I \notin \ell_n$, and ℓ_n lies on I, and $\ell_n \to I$. And we say that I is *isolated* if there is no I-side sequence on $\mathscr{L}$. Moreover, a leaf ℓ is *isolated* if each element of ℓ is isolated.

The following Lemma shows that the previous statement is true.

Lemma 10.7.12 *Let $\mathscr{L}$ be a lamination system on S^1 and $I \in \mathscr{L}$. Suppose that I^* is isolated. Then, there is a non-leaf gap $\mathscr{G}$ such that $I \in \mathscr{G}$.*

Proof For any point p in I^* at which $2 \leq |C_p^{I^*}|$, we define $J_p = \bigcup_{J \in C_p^{I^*} - \{I^*\}} J$. Then J_p is a nondegenerate open interval since J_p is a nonempty connected open subset of S^1 and $I \subseteq {J_p}^c$. By applying Lemma 10.7.10 to $C_p^{J_p}$, $J_p \in \mathscr{L}$ and since I^* is isolated, $J_p \subsetneq I^*$.

Note that J_p and J_q are disjoint or coincide whenever $p \neq q \in I^*$.

Define $\mathscr{G} = \{I\} \cup \{J_p : p \in I^* \textit{ and } 2 \leq |C_p^{I^*}|\}$. Then if there is $K \in \mathscr{L}$ with $\ell(K) \neq \ell(I)$, then $\ell(K)$ lies on I or I^*. If $\ell(K)$ lies on I^*, then there is L in $\ell(K)$ which is contained on I^*. Since for any x in L, $L \in C_x^{I^*}$, $\ell(K)$ lies on J_x for any $x \in L$. Thus, $\mathscr{G}$ is the non-leaf gap in which I is. □

The following lemma is about the configuration of two gaps. It is a kind of generalization of unlinkedness condition of two leaves to unlinkedness condition of two gaps.

Lemma 10.7.13 *Let $\mathscr{L}$ be a lamination system on S^1, and $\mathscr{G}$, $\mathscr{G}'$ be two gaps with $|\mathscr{G}|, |\mathscr{G}'| \geq 2$. Then, $\mathscr{G} = \mathscr{G}'$ or there are I in $\mathscr{G}$, and I' in $\mathscr{G}'$ such that $I^* \subseteq I'$, and for all $J \in \mathscr{G}$, $\ell(J)$ lies on I', and for all $J' \in \mathscr{G}'$, $\ell(J')$ lies on I.*

Proof Assume that $\mathscr{G} \neq \mathscr{G}'$. If $\mathscr{G} \subsetneq \mathscr{G}'$, then there is I in $\mathscr{G}' - \mathscr{G}$. Then, since $\mathscr{G}$ is a gap, there is J in $\mathscr{G}$ on which $\ell(I)$ lies. Since $\mathscr{G}'$ is a gap, then for all $K \in \mathscr{G}' - \{I\}$, $I \cap K = \emptyset$, and so $I \cap J = \emptyset$. Since $I \cap J = \emptyset$, $I \subseteq J^*$ and since $\ell(I)$ lies on J and $I \cap J = \emptyset$, then $I^* \subseteq J$ and $J^* \subseteq I$. Therefore, $I = J^*$, and so $\{I, I^*\} \subseteq \mathscr{G}'$. Thus, $\mathscr{G}'$ is a leaf $\ell(I)$, and so $\mathscr{G}$ is a one point subset of $\mathscr{G}'$. However, this contradicts $|\mathscr{G}'| \geq 2$. Similarly, we can get that $\mathscr{G}' \subsetneq \mathscr{G}$ is not possible.

So, there are J in $\mathscr{G} - \mathscr{G}'$ and J' in $\mathscr{G}' - \mathscr{G}$. Since $\mathscr{G}$ and $\mathscr{G}'$ are gaps, there are K in $\mathscr{G}$ and K' in $\mathscr{G}'$ such that $\ell(J')$ lies on K and $\ell(J)$ lies on K'. Since $\mathscr{G}'$ is a gap,

$\ell(K)$ lies on L' for some $L' \in \mathscr{G}'$. Likewise, since $\mathscr{G}$ is a gap, $\ell(K')$ lies on L for some $L \in \mathscr{G}$.

First, consider the case $L' \neq J'$. Since $\ell(K)$ lies on L', $K \subseteq L'$ or $K^* \subseteq L'$. If $K \subseteq L'$, then $J' \subseteq K$ cannot occur, and so $(J')^* \subseteq K \subseteq L'$ since $\ell(J')$ lies on K. Since by assumption, $L' \cap J' = \emptyset$, so $L' \subseteq (J')^*$. Therefore $(J')^* = K = L'$, and so $\mathscr{G}'$ should be a leaf $\ell(J')$. Then since $K = (J')^*$, and $I \subseteq K^*$ for all $I \in \mathscr{G} - \{K\}$, so $I \subseteq J'$ for all $I \in \mathscr{G} - \{K\}$. Therefore, for all $I \in \mathscr{G}$, $\ell(I)$ lies on J'. Since $K^* = J' \subseteq J'$ and, trivially, $\ell(J')$ lies on K, K and J' are the elements that we want to find. On the other hand, if $K^* \subseteq L'$, then for all $I \in \mathscr{G} - \{K\}$, $I \subseteq K^* \subseteq L'$ since $I \subseteq K^*$. Therefore, for all $I \in \mathscr{G}$, $\ell(I)$ lies on L'. And by the assumption, $(L')^* \subseteq K$. Likewise, for all $I' \in \mathscr{G}'$, $\ell(I')$ lies on K. So, in this case, K and L' are the elements that we want.

Second, consider the case $L' = J'$. Then $\ell(J')$ lies on K and $\ell(K)$ lies on J'. If $J' \subseteq K$, then $K^* \subseteq (J')^*$. Since $\ell(K)$ lies on J', there are two possibility. One is $K \subseteq J'$, and so $K = J'$. However, it contradicts $J' \in \mathscr{G}' - \mathscr{G}$. The other is $K^* \subseteq J'$, but it also contradicts $K^* \subseteq (J')^*$. Therefore, $(J')^* \subseteq K$. Since $\mathscr{G}'$ is a gap, for all $I' \in \mathscr{G}' - \{J'\}$, $I' \subseteq (J')^*$, and so for all $I' \in \mathscr{G}'$, $\ell(I')$ lies on K. And by the assumption, $K^* \subseteq J'$. Likewise, for all $I \in \mathscr{G}$, $\ell(I)$ lies on J'. Thus, K and J' are the elements that we want. □

On a lamination system $\mathscr{L}$ on S^1, a gap $\mathscr{G}$ with $|v(\mathscr{G})| < \infty$ is called an *ideal polygon*. In particular, an ideal polygon is called a *non-leaf ideal polygon* if it is not a leaf. For an ideal polygon $\mathscr{G}$, since $v(\mathscr{G})$ is a finite set, we can write $v(\mathscr{G}) = \{x_1, x_2, \cdots, x_n\}$ where $|v(\mathscr{G})| = n$, and $(x_1, x_2, \cdots, x_n)$ is a positively oriented n-tuple. Moreover, we can represent $\mathscr{G} = \{(x_1, x_2)_{S^1}, (x_2, x_3)_{S^1}, \cdots, (x_{n-1}, x_n)_{S^1}, (x_n, x_1)_{S^1}\}$. Then we say that a lamination system $\mathscr{L}$ is *very full* if every gap of $\mathscr{L}$ is an ideal polygon. Let $E(\mathscr{L}) = \bigcup_{I \in \mathscr{L}} v(\ell(I))$ and call it the *end points set* of $\mathscr{L}$. A lamination system $\mathscr{L}$ is called *dense* if $E(\mathscr{L})$ is a dense subset of S^1. Let $p \in S^1$ and $\mathscr{L}$ be a dense lamination system. Suppose that there is a sequence $\{I_n\}_{n=1}^{\infty}$ on $\mathscr{L}$ such that for all $n \in \mathbb{N}$, $I_{n+1} \subseteq I_n$, and $\bigcap_{n \in \mathbb{N}} I_n = \{p\}$. We call such a sequence a *rainbow* at p. In [3], it is observed that very full laminations have abundant rainbows (see Theorem 10.7.14 for a precise statement). In the rest of the section, we recall some results from [3] and [1] about invariant laminations and give alternative proofs in the language of lamination systems.

Theorem 10.7.14 ([3]) *Let $\mathscr{L}$ be a very full lamination system. For $p \in S^1$, either p is in $E(\mathscr{L})$ or p has a rainbow. These two possibilities are mutually exclusive.*

Proof Let p be a point of S^1. First we show that if there is no $I \in \mathscr{L}$ such that $p \in v(\ell(I))$, then p has a rainbow. Assume that there is no I in $\mathscr{L}$ such that $p \in v(\ell(I))$. Since $\mathscr{L}$ is nonempty, there is an element I in $\mathscr{L}$. Then, by assumption, $p \notin v(\ell(I)) = \partial I$. Since S^1 has a partition $\{I, \partial I, I^*\}$, p belongs to either I or I^*. Say that $p \in I$. Then, C_p^I is nonempty. By Lemma 10.7.10, there is a sequence

$\{K_n\}_{n=1}^{\infty}$ on C_p^I such that for all $n \in \mathbb{N}$, $K_{n+1} \subseteq K_n$, and $\bigcap_{n=1}^{\infty} K_n = \bigcap_{K \in C_p^I} K$. If $\bigcap_{n=1}^{\infty} K_n = \{p\}$, we are done. If not, $\bigcup_{n=1}^{\infty} K_n^*$ is a nondegenerate open interval J with $p \in J^c = \overline{J^*}$ and $J \in \mathscr{L}$, and by Lemma 10.7.5, $\text{Int}\left(\bigcap_{n=1}^{\infty} K_n\right) = J^* \in \mathscr{L}$. If $p \in \partial J^*$,then $p \in E(\mathscr{L})$ and so it contradicts the assumption. So $p \in J^*$. Then $p \in J^* \subseteq I$, and so J^* is the minimal element of C_p^I. Now, we want to show that J^* is isolated. Suppose that there is a J^*-side sequence $\{\ell_n\}_{n=1}^{\infty}$. Then there is a sequence $\{I_n\}_{n=1}^{\infty}$ on $\mathscr{L}$ such that for all $n \in \mathbb{N}$, $\ell_n = \ell(I_n)$, and

$$J^* \subseteq \liminf I_n \subseteq \limsup I_n \subseteq \overline{J^*}.$$

So, since $p \in J^* \subseteq \liminf I_n$, there is $m \in \mathbb{N}$ such that $p \in \bigcap_{n=m}^{\infty} I_n$. Therefore, for all $k \geq m$, $p \in I_k$ and so $I_k \nsubseteq J$. Then choose $q \in J$. Since $\limsup I_n \subseteq \overline{J^*}$, there is $m' \in \mathbb{N}$ such that $q \notin \bigcup_{n=m'}^{\infty} I_n$. Therefore, for all $k \geq m'$, $J \nsubseteq I_k$. So, for $k \geq \max\{m, m'\}$, $J^* \subseteq I_k$ or $I_k \subseteq J^*$. Since for all $n \in \mathbb{N}$, $\ell_n \neq \ell(J^*)$, then for $k \geq \max\{m, m'\}$, $J^* \subsetneq I_k$ or $I_k \subsetneq J^*$. Since for all $n \in \mathbb{N}$, ℓ_n lies on J^*, so for $k \geq \max\{m, m'\}$, $I_k \subsetneq J^*$. Moreover, for $k \geq \max\{m, m'\}$, $p \in I_k \subsetneq J^* \subseteq I$, and so $I_k \in C_p^I$. This contradicts the minimality of J^* on C_p^I. Therefore, J^* is isolated. Then by Lemma 10.7.12, there is the non-leaf gap $\mathscr{G}$ such that $J \in \mathscr{G}$. Since $\mathscr{L}$ is very full, $v(\mathscr{G})$ is finite, and so $\bigcup_{I \in \mathscr{G}} \partial I = v(\mathscr{G})$. Note that S^1 has a partition $\mathscr{G} \cup \{v(\mathscr{G})\}$. By assumption, $p \notin v(\mathscr{G})$, so there is $K \in \mathscr{G} - \{J\}$ such that $p \in K$. Since $K \subsetneq J^* \subseteq I$, $K \in C_p^I$. but this contradicts the minimality of J^* on C_p^I. Thus, $\bigcap_{n=1}^{\infty} K_n = \{p\}$.

Finally, we want to show that if there is a leaf ℓ such that $p \in v(\ell)$, p has no rainbow. Suppose that there are a rainbow $\{I_n\}_{n=1}^{\infty}$, and a leaf ℓ such that $p \in v(\ell)$. Since for all $n \in \mathbb{N}$, $p \in I_n$, ℓ lies on I_n. Choose $n \in \mathbb{N}$. Then there is an element I in ℓ such that $I \subsetneq I_n$. If $I^* \subsetneq I_{n+1}$, then $I^* \subsetneq I_{n+1} \subseteq I_n$, but this is a contradiction. So, $I \subsetneq I_{n+1}$. Therefore, $I \subseteq \bigcap_{n=1}^{\infty} I_n$, but this is not possible since $\bigcap_{n=1}^{\infty} I_n = \{p\}$. □

Corollary 10.7.15 ([3]) *Let $\mathscr{L}$ be a very full lamination system of S^1. Then, $E(\mathscr{L})$ is dense in S^1.*

Proof Suppose that $E(\mathscr{L})$ is not dense. Then, there is a point p in S^1 which has an open neighborhood K which is a nondegenerate open interval with $E(\mathscr{L}) \cap \overline{K} = \emptyset$.

And by Theorem 10.7.14, there is a rainbow $\{I_n\}_{n=1}^{\infty}$ at p. Fix $n \in \mathbb{N}$ and denote $I_n = (u_n, v_n)_{S^1}$ and $K = (s,t)_{S^1}$. Let φ be the circular order of S^1. Note that $\varphi(u_n, p, v_n) = 1$. Since $E(\mathscr{L}) \cap \overline{K} = \emptyset$, $\varphi(s, u_n, t) = \varphi(s, v_n, t) = -1$ and so $\varphi(t, u_n, s) = \varphi(t, v_n, s) = 1$. Since $\varphi(s, p, t) = 1$ and $\varphi(t, s, p) = 1$, so $\varphi(t, u_n, p) = \varphi(t, v_n, p) = 1$. Since by the cocycle condition on the four points (t, u_n, v_n, p),

$$\varphi(u_n, v_n, p) - \varphi(t, v_n, p) + \varphi(t, u_n, p) - \varphi(t, u_n, v_n) = 0,$$

then

$$-1 - 1 + 1 - \varphi(t, u_n, v_n) = 0,$$

Hence $\varphi(t, u_n, v_n) = -1$. Therefore, $\varphi(u_n, t, v_n) = 1$. Likewise, $\varphi(s, u_n, t) = \varphi(s, v_n, t) = -1$ and so $\varphi(s, t, u_n) = \varphi(s, t, v_n) = 1$. Since $\varphi(s, p, t) = 1$, then $\varphi(s, p, u_n) = \varphi(s, p, v_n) = 1$. Since by the cocycle condition on the four points (s, p, u_n, v_n),

$$\varphi(p, u_n, v_n) - \varphi(s, u_n, v_n) + \varphi(s, p, v_n) - \varphi(s, p, u_n) = 0,$$

then

$$-1 - \varphi(s, u_n, v_n) + 1 - 1 = 0.$$

Hence $\varphi(s, u_n, v_n) = -1$. Therefore, $\varphi(u_n, s, v_n) = 1$.

We have shown that $\varphi(u_n, s, v_n) = \varphi(u_n, t, v_n) = 1$ and we have $\varphi(s, u_n, t) = \varphi(s, v_n, t) = -1$ since $E(\mathscr{L}) \cap \overline{K} = \emptyset$. From now on, we show that $K \subseteq I_n$. Let q be a point in K. Then $\varphi(s, q, t) = 1$ and since $\varphi(s, u_n, t) = \varphi(s, v_n, t) = -1$ and so $\varphi(s, t, u_n) = \varphi(s, t, v_n) = 1$, we get that $\varphi(s, q, u_n) = \varphi(s, q, v_n) = 1$. Then by applying the cocycle condition to four points (u_n, s, q, v_n),

$$\varphi(s, q, v_n) - \varphi(u_n, q, v_n) + \varphi(u_n, s, v_n) - \varphi(u_n, s, q) = 0.$$

Since $\varphi(u_n, s, q) = \varphi(s, q, u_n)$,

$$\varphi(s, q, v_n) - \varphi(u_n, q, v_n) + \varphi(u_n, s, v_n) - \varphi(u_n, s, q) = 1 - \varphi(u_n, q, v_n) + 1 - 1 = 0.$$

Therefore, $\varphi(u_n, q, v_n) = 1$ and so $q \in I_n$. We are done. This implies that for all $n \in \mathbb{N}$, $K \subseteq I_n$, so $K \subseteq \bigcap_{n=1}^{\infty} I_n$, but this contradicts the definition of a rainbow. Thus, $E(\mathscr{L})$ is dense. □

Indeed, very fullness does not guarantee the existence of non-leaf gaps. More precisely, a lamination system, of which the geometric realization is a geodesic lamination which foliates the whole hyperbolic plane, is very full, but there is no non-leaf gap. So, we need some notions to rule out this situation and to guarantee the existence of a gap on a lamination system. So, the following definitions on a lamination system describe the situation which is analogous to that in $\overline{\mathbb{H}^2}$, there is no open disk foliated by leaves on a given geodesic lamination which is a geometric realization of a lamination on S^1.

Definition 10.7.16 Let $\mathscr{L}$ be a lamination system and $\{I, J\}$ be a subset of $\mathscr{L}$. Then, $\{I, J\}$ is called a *distinct pair* if $I \cap J = \emptyset$, and $\{I, J\}$ is not a leaf. A distinct pair $\{I, J\}$ is *separated* if there is a non-leaf gap $\mathscr{G}$ such that $I \subseteq K$ and $J \subseteq L$ for some $K, L \in \mathscr{G}$, not necessarily $K \neq L$. And $\mathscr{L}$ is *totally disconnected* if every distinct pair is separated.

Two lamination systems $\mathscr{L}_1$ and $\mathscr{L}_2$ have *distinct endpoints* if $E(\mathscr{L}_1) \cap E(\mathscr{L}_2) = \emptyset$. When we study two lamination systems, the distinct endpoints condition enforces totally disconnectedness on lamination systems.

Lemma 10.7.17 ([1]) *If two dense lamination systems have distinct endpoints, then each of the lamination systems is totally disconnected.*

Proof Let $\mathscr{L}_1$ and $\mathscr{L}_2$ be two dense lamination systems with distinct endpoints. First, we show that $\mathscr{L}_1$ is totally disconnected. Suppose that a subset $\{I, J\}$ of $\mathscr{L}_1$ is a distinct pair. $I^* \cap J^*$ is a non-empty open set. Since $E(\mathscr{L}_2)$ is dense in S^1, we can choose $p \in I^* \cap J^* \cap E(\mathscr{L}_2)$. If there is $K \in \mathscr{L}_1$ such that $p \in K \subseteq I^* \cap J^*$, then $K \in C_p^{I^*} \cap C_p^{J^*}$ and so $C_p^{I^*} \cap C_p^{J^*}$ is nonempty where we consider $C_p^{I^*}$ and $C_p^{J^*}$ on $\mathscr{L}_1$. Note that $C_p^{I^*} \cap C_p^{J^*}$ is totally ordered by $\subseteq$. Let M be the union of elements of $C_p^{I^*} \cap C_p^{J^*}$. Then M is a nondegenerate open interval with $p \in M \subseteq I^* \cap J^*$. So, C_p^M is equal to $C_p^{I^*} \cap C_p^{J^*} = \{K \in \mathscr{L}_1 : p \in K \subseteq I^* \cap J^*\}$. Moreover, by Lemma 10.7.10, $M \in \mathscr{L}_1$.

We want to show that M^* is isolated on $\mathscr{L}_1$. Suppose that $\{\ell_n\}_{n=1}^{\infty}$ be an M^*-side sequence of leaves of $\mathscr{L}_1$. Then, there is $\{I_n\}_{n=1}^{\infty}$ such that for all $n \in \mathbb{N}$, $\ell_n = \ell(I_n)$, and

$$M^* \subseteq \liminf I_n \subseteq \limsup I_n \subseteq \overline{M^*}.$$

Choose $p_I \in I$ and $p_J \in J$. Since $M^* \subseteq \liminf I_n$ and $\{p_I, p_J\} \subseteq I \cup J \subseteq M^*$, there is $m \in \mathbb{N}$ such that $\{p_I, p_J\} \subseteq \bigcap_{n=m}^{\infty} I_n$. Therefore, for all $n \geq m$, $\{p_I, p_J\} \subseteq I_n$ and so $I_n \nsubseteq M$. Also, since $\limsup I_n \subseteq \overline{M^*}$ and $p \in M$, there is m' such that $p \notin \bigcup_{n=m'}^{\infty} I_n$. Therefore, for all $n \geq m'$, $p \notin I_n$ and so $M \nsubseteq I_n$. Hence, for all $n \geq \max\{m, m'\}$, $I_n \subseteq M^*$ or $M^* \subseteq I_n$. Moreover, for $n \geq \max\{m, m'\}$, since $\ell_n \neq \ell(M^*)$, $I_n \subsetneq M^*$ or $M^* \subsetneq I_n$ and since ℓ_n lies on M^*, $I_n \subsetneq M^*$ is the possible case. Thus, for all $n \geq \max\{m, m'\}$, $\{p_I, p_J\} \subseteq I_n \subsetneq M^*$.

Then, fix $n \geq \max\{m, m'\}$. If $I^* \subseteq I_n$ or $J^* \subseteq I_n$, then $M \subseteq I_n$, and this contradicts $I_n \subseteq M^*$. If $I_n \subseteq I^*$ or $I_n \subseteq J^*$, then $\{p_I, p_J\} \subseteq I^*$ or $\{p_I, p_J\} \subseteq J^*$, respectively, and so it is also a contradiction since $I \cap I^* = \emptyset$ and $J \cap J^* = \emptyset$. If $I_n \subseteq I$ or $I_n \subseteq J$, then $\{p_I, p_J\} \subseteq I$ or $\{p_I, p_J\} \subseteq J$, respectively, and so this is also a contradiction since $I \cap J = \emptyset$. Therefore, $I \cup J \subseteq I_n$. Hence $M \subsetneq I_n^* \subseteq I^* \cap J^*$. This contradicts the maximality of M on $C_p^{I^*} \cap C_p^{J^*}$. Thus, M^* is isolated.

Finally, by Lemma 10.7.12, there is a non-leaf gap $\mathscr{G}$ such that $M \in \mathscr{G}$. If there is $L \in \mathscr{G}$ such that $I^* \subseteq L$ or $J^* \subseteq L$, then $M \subseteq I^* \subseteq L$ or $M \subseteq J^* \subseteq L$ and so $M = I^* = L$ or $M = J^* = L$, respectively. Then, $J \subseteq I^* = M$ or $I \subseteq J^* = M$, respectively. However, this is a contradiction since $M \cap \{p_I, p_J\} = \emptyset$. Therefore, by the definition of gap, there are L and L' in $\mathscr{G}$ such that $I \subseteq L$ and $J \subseteq L'$. So, $\{I, J\}$ is separated.

Next, assume that $C_p^{I^*} \cap C_p^{J^*} = \emptyset$. Choose $K \in C_p^{I^*}$. If $K \subseteq J$, then this contradicts $p \in J^*$. If $K \subseteq J^*$, then $K \in C_p^{J^*}$, and so this contradicts the assumption. If $J^* \subseteq K$, then $J^* \subseteq K \subseteq I^*$, and $I \subseteq J$, then this contradicts the definition of distinct pairs. Therefore, $J \subsetneq K$, and so $(\{p\} \cup J) \subseteq K$. By Lemma 10.7.10, there is a sequence $\{F_n\}_{n=1}^{\infty}$ such that for all $n \in \mathbb{N}$, $F_{n+1} \subseteq F_n$ and $\bigcap_{n=1}^{\infty} F_n = \bigcap_{F \in C_p^{I^*}} F$ and since $(\{p\} \cup J) \subseteq F$ for all $F \in C_p^{I^*}$, $\bigcup_{n=1}^{\infty} F_n^*$ is a nondegenerate open interval and so $\bigcup_{n=1}^{\infty} F_n^* \in \mathscr{L}_1$. Therefore, by Lemma 10.7.5, $\operatorname{Int}\left(\bigcap_{n=1}^{\infty} F_n\right) = \operatorname{Int}\left(\bigcap_{F \in C_p^{I^*}} F\right)$ is a nondegenerate open interval N with $p \in \overline{N}$, $J \subseteq N$ and $N \in \mathscr{L}_1$. Since p is not in $E(\mathscr{L}_1)$, then $p \in N$.

Now, we want to show that N is isolated. Suppose that there is an N-side sequence $\{\ell_n\}_{n=1}^{\infty}$ on $\mathscr{L}_1$. So, there is a sequence $\{I_n\}_{n=1}^{\infty}$ on $\mathscr{L}_1$ such that for all $n \in \mathbb{N}$, $\ell_n = \ell(I_n)$ and

$$N \subseteq \liminf I_n \subseteq \limsup I_n \subseteq \overline{N}.$$

Choose $q \in I$. Since $\limsup I_n \subseteq \overline{N}$, there is $m \in \mathbb{N}$ such that $q \notin \bigcup_{n=m}^{\infty} I_n$. Also, since $N \subseteq \liminf I_n$, there is $m' \in \mathbb{N}$ such that $p \in \bigcap_{n=m'}^{\infty} I_n$. Therefore, for all $n \geq \max\{m, m'\}$, $p \in I_n$ and $q \in I_n^c$.

Fix $n \geq \max\{m, m'\}$. If $I_n \subseteq N^*$, then $p \in I_n \subseteq N^*$ and this is a contradiction since $p \in N$. If $N^* \subseteq I_n$, then $q \in I_n^c \subseteq (N^*)^c$ and this is a contradiction since $q \in I \subseteq N^*$. Therefore, $I_n \subseteq N$ or $N \subseteq I_n$. Moreover, since $\ell_n \neq \ell(N)$, $I_n \subsetneq N$ or $N \subsetneq I_n$ and since ℓ_n lies on N, $I_n \subsetneq N$ is the possible case. But $p \subseteq I_n \subsetneq N$, and this contradicts the minimality of N on $C_p^{I^*}$. So, by Lemma 10.7.12, there is a non-leaf gap $\mathscr{G}$ such that $N^* \in \mathscr{G}$. It is enough to show that there is K in $\mathscr{G}$ such that $J \subseteq K$. Suppose that there is K' in $\mathscr{G}$ such that $J^* \subseteq K'$. Then $N^* \subseteq J^* \subseteq K'$ and so $N^* = J^* = K'$. But, this implies $N = J$ and since $p \in N$ and $p \notin J$, this is a contradiction. Therefore, there is K in $\mathscr{G}$ such that $J \subseteq K$. Thus, $\{I, J\}$ is separated and so $\mathscr{L}_1$ is totally disconnected. For the same reason, $\mathscr{L}_2$ is also totally disconnected. □

We have introduced lamination systems as a model for laminations on the circle. In this perspective, laminar groups are groups acting on the circle with invariant lamination systems. We end this section, discussing about actions on lamination systems. A homeomorphism f on S^1 is *orientation preserving* if for any positively oriented triple (z_1, z_2, z_3), $(f(z_1), f(z_2), f(z_3))$ is a positively oriented triple. We denote the set of orientation preserving homeomorphisms on S^1 as $\text{Homeo}^+(S^1)$ and the set of fixed points of f as Fix_f. Note that if $f \in \text{Homeo}^+(S^1)$, for $u, v \in S^1$, we have $f((u, v)_{S^1}) = (f(u), f(v))_{S^1}$ and if $u \neq v$, then $f((u, v)^*_{S^1}) = f((u, v)_{S^1})^*$.

Definition 10.7.18 Let $\mathscr{L}$ be a lamination system, and G be a subgroup of $\text{Homeo}^+(S^1)$. $\mathscr{L}$ is called a *G-invariant lamination system* if for any $I \in \mathscr{L}$ and $g \in G$, $g(I) \in \mathscr{L}$. When $\mathscr{L}$ is a G-invariant lamination system, the action of G on $\mathscr{L}$ is said to be *minimal* if for any two leaves $\ell, \ell' \in \mathscr{L}$, there is a sequence $\{g_n\}_{n=1}^{\infty}$ on G such that $g_n(\ell') \to \ell$.

First, note that on a G-invariant lamination system, every gap is mapped to a gap and the converging property is preserved under the given G-action. Furthermore, when we consider a minimal action on a lamination system which has a non-leaf ideal polygon, the orbit of end points of the ideal polygon is usually dense in S^1. This denseness gives the following lemma which is useful to analyze the action.

Lemma 10.7.19 *Let G be a subgroup of* $\text{Homeo}^+(S^1)$*, and $\mathscr{L}$ a G-invariant lamination system. Assume that there is an ideal polygon $\mathscr{G}$ which is not a leaf, and that* $v_G(\mathscr{G}) = \bigcup_{g \in G} v(g(\mathscr{G}))$ *is dense in S^1. Then, for each $I \in \mathscr{G}$, there is an element g_I in G such that for any $J \in g_I(\mathscr{G})$, $\ell(J)$ properly lies on I, equivalently $v(g_I(\mathscr{G})) \subseteq I$.*

Proof Choose $I \in \mathscr{G}$. Since $v_G(\mathscr{G})$ is dense in S^1, there is $p \in v_G(\mathscr{G}) \cap I$. By the definition of $v_G(\mathscr{G})$, there is an element g in G such that $p \in v(g(\mathscr{G}))$. Note that $\mathscr{G} \neq g(\mathscr{G})$ since $v(\mathscr{G}) \neq v(g(\mathscr{G}))$. Since $\mathscr{G}$ is an ideal polygon and $3 \leq |\mathscr{G}|$, by Lemma 10.7.13, there are J in $\mathscr{G}$, and J' in $g(\mathscr{G})$ such that $(J')^* \subseteq J$. Hence, for all $K \in \mathscr{G}$, $\ell(K)$ lies on J', and for all $K' \in g(\mathscr{G})$, $\ell(K')$ lies on J. Since for all $K' \in g(\mathscr{G})$, $\ell(K')$ lies on J, then $v(g(\mathscr{G})) \subseteq \bar{J}$. Therefore, since $p \in I$, then J should be I. Since $(J')^* \subseteq J$, for each $I' \in g(\mathscr{G}) - \{J'\}$, $I' \subseteq J$. Choose I' in $g(\mathscr{G}) - \{J'\}$. If $I' = J$, then $J' \subseteq (I')^* = J^*$ and so $J' = (I')^* = J^*$since $(J')^* \subseteq J$. But in this case, $g(\mathscr{G})$ is the leaf $\ell(I')$, which contradicts the assumption. Therefore, $I' \subsetneq J = I$.

Then we do the same process in I'. Since $v_G(\mathscr{G})$ is dense in S^1, there is $p' \in v_G(\mathscr{G}) \cap I'$. By the definition of $v_G(\mathscr{G})$, there is an element g' in G such that $p' \in v(g'(\mathscr{G}))$. By Lemma 10.7.13, there are L in $g(\mathscr{G})$ and L' in $g'(\mathscr{G})$ such that $(L')^* \subseteq L$. Hence, for all $K \in g(\mathscr{G})$, $\ell(K)$ lies on L' and for all $K' \in g'(\mathscr{G})$, $\ell(K')$ lies on

L. Since for all $K' \in g'(\mathcal{G})$, $\ell(K')$ lies on L, then $v(g'(\mathcal{G})) \subseteq \overline{L}$. Therefore, since $p' \in I'$, L should be I'. If $\overline{I'} \subseteq I$, then g' is the element that we want. Assume that $\overline{I'} \nsubseteq I$. Since $I' \subsetneq I$, then $\partial I' \nsubseteq I$, that is, there is an element x in $\partial I'$ such that $x \notin I$. Since $\overline{I'} \subseteq \overline{I}$ then $\partial I' \subseteq \overline{I}$. Therefore, $x \in \partial I$ since $x \notin I$. So, $x \in \partial I \cap \partial I'$. If $\partial I = \partial I'$, then $I' = I$ or $I' = I^*$ since I and I' are nondegenerate open intervals. However, this is a contradiction since $I' \subsetneq I$. Thus, $\partial I' \cap \partial I = \{x\}$.

In this case, if $x \notin v(g'(\mathcal{G}))$, then we are done. Assume that $x \in v(g'(\mathcal{G}))$. Note that $v(g'(\mathcal{G})) \subseteq \overline{(L')^*}$. Since $(L')^* \subseteq I'$, $x \in \overline{(L')^*} \subseteq \overline{I'}$ and $x \in \partial I'$, so $x \in \partial L'$. There is a unique element M in $g'(\mathcal{G}) - \{L'\}$ such that $x \in \partial M$. Therefore, for any N in $g'(\mathcal{G}) - \{L', M\}$, $\overline{N} \subseteq I$. Choose I'' in $g'(\mathcal{G}) - \{L', M\}$. Finally, since $v_G(\mathcal{G})$ is dense in S^1, there is $p'' \in v_G(\mathcal{G}) \cap I''$. By the definition of $v_G(\mathcal{G})$, there is an element g'' in G such that $p'' \in v(g''(\mathcal{G}))$. Like the previous argument, we can conclude that $v(g''(\mathcal{G})) \subseteq \overline{I''}$. Therefore, $v(g''(\mathcal{G})) \subseteq I$. □

10.8 Not Virtually Abelian Laminar Groups

In this section, we prove the following theorem which gives a condition which guarantees that a laminar group is not virtually abelian.

Theorem 10.8.1 *Let G be a subgroup of* $\mathrm{Homeo}^+(S^1)$ *and $\mathcal{L}$ be a G-invariant lamination system. Suppose that there is an ideal polygon $\mathcal{G}$ on $\mathcal{L}$ which is not a leaf. If $v_G(\mathcal{G})$ is dense in S^1, then G is not virtually abelian.*

The denseness of $v_G(\mathcal{G})$ allows some movements of intervals by an element of G. So, the strategy of the proof of the above theorem is to analyze the fixed point set of some element of G and to find contradictory configurations of fixed points by using the denseness. Before proving the theorem, we define some notions about non-leaf ideal polygons.

Definition 10.8.2 Let G be a subgroup of $\mathrm{Homeo}^+(S^1)$ and $\mathcal{L}$ a G-invariant lamination system. For each $g \in G$ and for each ideal polygon $\mathcal{G}$ of $\mathcal{L}$, we define the g-*type* of $\mathcal{G}$ as follows:

(1) The g-type of $\mathcal{G}$ is g-*free* if $|v(\mathcal{G}) \cap Fix_g| = 0$
(2) The g-type of $\mathcal{G}$ is g-*sticky* if $|v(\mathcal{G}) \cap Fix_g| = 1$
(3) The g-type of $\mathcal{G}$ is g-*fixed* if $v(\mathcal{G}) \subseteq Fix_g$

In the following proposition, we can see that for each element g of G and for any ideal polygon $\mathcal{G}$ of a lamination system $\mathcal{L}$, $\mathcal{G}$ is one of these g-types.

Proposition 10.8.3 *Let G be a subgroup of* $\mathrm{Homeo}^+(S^1)$ *and $\mathcal{L}$ a G-invariant lamination system. Suppose that there is an ideal polygon $\mathcal{G}$ in $\mathcal{L}$ which is not a leaf. For $g \in G$, if there are three distinct elements I_1, I_2 and I_3 in $\mathcal{G}$ such that for all $i \in \mathbb{Z}_3$ ($i \in \mathbb{Z}_n$ means that the indices are modulo n), $\overline{I_i}$ contains a fixed point of g, then $\mathcal{G}$ is g-fixed.*

Proof By Lemma 10.7.13, $g(\mathscr{G}) = \mathscr{G}$ or there are I in $\mathscr{G}$ and I' in $g(\mathscr{G})$ such that $I^* \subseteq I'$. First, we consider the latter case. If $I' = g(I)$, then $I^* \subseteq g(I)$. We set $I = (a, b)_{S^1}$. Then $(b, a)_{S^1} \subseteq (g(a), g(b))_{S^1}$ and so a and b are not fixed points of g. Moreover, for $z \in (b, a)_{S^1}$, $z \in (b, a)_{S^1} \subseteq (g(a), g(b))_{S^1}$ and so $g^{-1}(z) \in (a, b)_{S^1}$. Then since $(a, b)_{S^1}$ and $(b, a)_{S^1}$ are disjoint, $z \neq g^{-1}(z)$ and so $g(z) \neq z$. Therefore, $(b, a)_{S^1} \subseteq S^1 - Fix_g$ and so $[b, a]_{S^1} \subseteq S^1 - Fix_g$. However, this implies that there is only one element in $\mathscr{G}$ whose closure contains a fixed point of g and this is a contradiction by the assumption.

So, we assume that $I' \neq g(I)$. Choose K in $\mathscr{G} - \{I, g^{-1}(I')\}$. Denote $K = (x, y)_{S^1}$. Since $(I')^* \subseteq I$ and $g(K)$ and I' are disjoint, $g(K) \subseteq (I')^* \subseteq I$ and since $I \subseteq K^*$, $g(K) \subseteq K^*$, that is, $(g(x), g(y))_{S^1} \subseteq (y, x)_{S^1}$. This implies that x and y are not fixed points of g. And for all $w \in (x, y)_{S^1}$, $g(w) \in (g(x), g(y))_{S^1} \subseteq (y, x)_{S^1}$ and since $(x, y)_{S^1}$ and $(y, x)_{S^1}$ are disjoint, $g(w) \neq w$. Therefore, $(x, y)_{S^1} \subseteq S^1 - Fix_g$ and so $[x, y]_{S^1} \subseteq S^1 - Fix_g$. However, there are exactly two elements I and $g^{-1}(I')$ of which the closures can contains fixed points of g and this is a contradiction by the assumption. Thus, $g(\mathscr{G}) = \mathscr{G}$ is the only possible case.

We set $\mathscr{G} = \{(x_1, x_2)_{S^1}, (x_2, x_3)_{S^1}, \cdots, (x_{n-1}, x_n)_{S^1}, (x_n, x_1)_{S^1}\}$ and use $\mathbb{Z}_n$ as the index set. Note that since $g(\mathscr{G}) = \mathscr{G}$, there is $k \in \mathbb{Z}_n$ such that $(g(x_i), g(x_{i+1}))_{S^1} = (x_{i+k}, x_{i+1+k})_{S^1}$. If $k \neq 0$ on $\mathbb{Z}_n$, then there is no fixed point of g since for all $i \in \mathbb{Z}_n$, $g(x_i) = x_{i+k} \neq x_i$ and $g((x_i, x_{i+1})_{S^1}) = (x_{i+k}, x_{i+1+k})_{S^1} \subseteq (x_i, x_{i+1})^*_{S^1}$. This is a contradiction since $Fix_g \neq \emptyset$ by the assumption. Therefore, $k = 0$ on $\mathbb{Z}_n$. Thus, for all $i \in \mathbb{Z}_n$, $g(x_i) = x_i$ and so $\mathscr{G}$ is g-fixed. □

Corollary 10.8.4 *Let G be a subgroup of* $\mathrm{Homeo}^+(S^1)$ *and $\mathscr{L}$ be a G-invariant lamination system. Let g be a nontrivial element of G with $Fix_g \neq \emptyset$. Suppose that there is an ideal polygon $\mathscr{G}$ of $\mathscr{L}$ with $2 \leq |v(G) \cap Fix_g|$. Then, $\mathscr{G}$ is g-fixed.*

By Corollary 10.8.4, we can see that all ideal polygons fall into one of the three g-types defined in Definition 10.8.2. The following proposition is very classical but since it will be frequently used, we include its proof for completeness.

Proposition 10.8.5 *Let g and h be two elements of* $\mathrm{Homeo}^+(S^1)$ *and x an element of S^1. Suppose that $gh = hg$. Then, x is a fixed point of g if and only if $h(x)$ is a fixed point of g.*

Proof Suppose that x is a fixed point of g. Then $g(h(x)) = h(g(x)) = h(x)$ and so $h(x)$ is a fixed point of g. Conversely, suppose that $h(x)$ is a fixed point of g. Then, $h(g(x)) = g(h(x)) = h(x)$ and since h is a bijection, $g(x) = x$. And so x is a fixed point of g. □

To start the proof of the main theorem, we should take a non-trivial element of G which has a fixed point. The following lemma shows that there is a non-trivial element of G under the conditions of the main theorem.

Lemma 10.8.6 *Let G be a subgroup of* $\mathrm{Homeo}^+(S^1)$ *and $\mathscr{L}$ be a G-invariant lamination system. Suppose that there is an ideal polygon $\mathscr{G}$ on $\mathscr{L}$ such that $\mathscr{G}$ is* *not a leaf and $v_G(\mathscr{G})$ is dense in S^1. Then, there is a nontrivial element g of G such that $Fix_g \neq \emptyset$.*

Proof Assume that there is no non-trivial element of G which has a fixed point. Choose $I \in \mathscr{G}$. By Lemma 10.7.19, there is an element g in G such that for any $K \in g(\mathscr{G})$, $\ell(K)$ properly lies on I. By Lemma 10.7.13, there is the element I' in $g(\mathscr{G})$ such that $(I')^* \subseteq I$ and since $\ell(I')$ properly lies on I, $\overline{(I')^*} \subseteq I$. If $I' \neq g(I)$, then $g(I) \subseteq (I')^*$ and so $g(\overline{I}) \subseteq \overline{(I')^*} \subseteq I \subseteq \overline{I}$. This implies that there is a fixed point of g in I, but this contradicts the assumption. So, $I' = g(I)$ is the possible case. Then, choose J in $g(\mathscr{G})$ such that $\overline{J} \subset I$. By Lemma 10.7.19, there is an element h in G such that for any $K \in hg(\mathscr{G})$, $\ell(K)$ properly lies on J. By Lemma 10.7.13, there is an element J' in $hg(\mathscr{G})$ such that $(J')^* \subseteq J$ and since $\ell(J')$ properly lies on J, $\overline{(J')^*} \subseteq J$. If $J' \neq h(J)$, then $h(J) \subseteq (J')^*$ and so $h(\overline{J}) \subseteq \overline{(J')^*} \subseteq J \subseteq \overline{J}$. This implies that there is a fixed point of h in J, but this contradicts the assumption. So $J' = h(J)$. Since $g(I) = I'$, $g(I) \subseteq J^*$ and so $h(g(I)) \subseteq h(J^*) = h(J)^* = (J')^*$. Then,

$$h(g(\overline{I})) = \overline{h(g(I))} \subseteq \overline{(J')^*} \subseteq J \subseteq \overline{J} \subseteq I \subseteq \overline{I}.$$

This implies that the nontrivial element hg has a fixed point in I but this contradicts the assumption. Thus, there is a nontrivial element of G which has a fixed point. □

Before proving the virtual case, we show that G is non-abelian under the condition of the main theorem.

Theorem 10.8.7 *Let G be a subgroup of* $\mathrm{Homeo}^+(S^1)$ *and $\mathscr{L}$ be a G-invariant lamination system. Suppose that there is an ideal polygon $\mathscr{G}$ on $\mathscr{L}$ which is not a leaf. If $v_G(\mathscr{G})$ is dense in S^1, then G is non-abelian.*

Proof Assume that G is abelian. By Lemma 10.8.6, there is a nontrivial element g in G with $Fix_g \neq \emptyset$. First, if there are three distinct elements in $\mathscr{G}$ such that the closure of each element contains a fixed point, then by Proposition 10.8.3, $\mathscr{G}$ is g-fixed. Since for all $h \in G$, $h(v(\mathscr{G})) \subseteq Fix_g$, so by Proposition 10.8.5, $v_G(\mathscr{G}) \subseteq Fix_g$. By the assumption, $v_G(\mathscr{G})$ is dense and so Fix_g is dense in S^1. Since Fix_g is closed in S^1, $Fix_g = S^1$, but this implies that g is the trivial element of G and so it is a contradiction.

If there are exactly two distinct elements I and J in $\mathscr{G}$ such that $\overline{I} \cap Fix_g \neq \emptyset$ and $\overline{J} \cap Fix_g \neq \emptyset$, then there is an element K in $\mathscr{G}$ such that $\overline{K} \cap Fix_g = \emptyset$. By Lemma 10.7.19, there is an element h in G such that for any $L \in h(\mathscr{G})$, $\ell(L)$ properly lies on K. By Lemma 10.7.13, there is an element K' in $h(\mathscr{G})$ such that $(K')^* \subseteq K$ and since $\ell(K')$ properly lies on K, $\overline{(K')^*} \subseteq K$. Then at least one of $h(I)$ and $h(J)$ is not K'. Without loss of generality, we may assume that $h(I) \neq K'$. Then $h(I) \subseteq (K')^* \subseteq K$ and so $h(\overline{I}) \subseteq \overline{K}$. However, by Proposition 10.8.5, $h(\overline{I}) \cap Fix_g \neq \emptyset$ since $\overline{I} \cap Fix_g \neq \emptyset$, and so $\overline{K} \cap Fix_g \neq \emptyset$. This is a contradiction since $\overline{K} \cap Fix_g = \emptyset$.

Finally, if there is a unique element M in $\mathscr{G}$ such that $\overline{M} \cap Fix_g \neq \emptyset$, that is, $Fix_g \subseteq \overline{M}$, then there are two distinct elements O_1 and O_2 in $\mathscr{G}$ such that $\overline{O_1} \cap Fix_g = \emptyset$ and $\overline{O_2} \cap Fix_g = \emptyset$. For each $i \in \mathbb{Z}_2$, by Lemma 10.7.19, there is an element f_i in G such that for any $P \in f_i(\mathscr{G})$, $\ell(P)$ properly lies on O_i. Fix $i \in \mathbb{Z}_2$. By Lemma 10.7.13, there is an element O_i' in $f_i(\mathscr{G})$ such that $(O_i')^* \subseteq O_i$. If $O_i' \neq f_i(M)$, then $f_i(M) \subseteq (O_i')^* \subseteq O_i$ and so $f_i(\overline{M}) = \overline{f_i(M)} \subseteq \overline{O_i}$. However, by Proposition 10.8.5, $f_i(\overline{M}) \cap Fix_g \neq \emptyset$ since $\overline{M} \cap Fix_g \neq \emptyset$ and so this is a contradiction since $\overline{O_i} \cap Fix_g = \emptyset$. Therefore, for all $i \in \mathbb{Z}_2$, $f_i(M) = O_i'$. Then, we get the following relations:

(1) $f_1(O_1) \subseteq f_1(M^*) = f_1(M)^* = (O_1')^* \subseteq O_1$
(2) $f_1(O_2) \subseteq f_1(M^*) = f_1(M)^* = (O_1')^* \subseteq O_1$
(3) $f_2(O_1) \subseteq f_2(M^*) = f_2(M)^* = (O_2')^* \subseteq O_2$
(4) $f_2(O_2) \subseteq f_2(M^*) = f_2(M)^* = (O_2')^* \subseteq O_2$

Let us consider two elements $f_1 f_2$ and $f_2 f_1$.

$$f_1 f_2(O_1) \subseteq f_1(O_2) \subseteq O_1$$

and

$$f_2 f_1(O_1) \subseteq f_2(O_1) \subseteq O_2.$$

However, this implies $f_1 f_2(O_1) \neq f_2 f_1(O_1)$ since O_1 and O_2 are disjoint, and so this contradicts the assumption that G is abelian. Thus, G is non-abelian. □

To improve this theorem, we need the following lemma. When we prove the virtual case, we will take a finite index subgroup H of G and construct new lamination system which is preserved by H. In this construction of the H-invariant lamination system, we will collapse the original circle on which the G-invariant lamination system is defined. The following lemma guarantees that there is a non-leaf gap of the H-invariant lamination system.

Lemma 10.8.8 *Let G be a subgroup of* $\mathrm{Homeo}^+(S^1)$ *and $\mathscr{L}$ a G-invariant lamination system in which there is a non-leaf ideal polygon $\mathscr{G}_0$. Suppose that $v_G(\mathscr{G}_0)$ is dense in S^1. If H is a finite index subgroup of G, then there is a non-leaf ideal polygon $\mathscr{G}$ in $\mathscr{L}$ which is $g(\mathscr{G}_0)$ for some $g \in G$ and has three elements I_1, I_2 and I_3 such that for all $i \in \mathbb{Z}_3$, $I_i \cap \overline{v_H(\mathscr{G})}$ has nonempty interior.*

Proof Since the case $G = H$ is obvious, we assume that H is a proper subgroup of G. Assume that for each $g \in G$, there are at most two elements in $g(\mathscr{G}_0)$ which contain interior points of $\overline{v_H(g(\mathscr{G}_0))}$. Since H has a finite index, we can denote

$H\backslash G = \{Hg_1, Hg_2, \cdots, Hg_n\}$ for some $\{g_1, g_2, \cdots, g_n\} \subseteq G$. Then, $v_G(\mathscr{G}_0) = \bigcup_{i=1}^{n} v_H(g_i(\mathscr{G}_0))$. So,

$$S^1 = \overline{v_G(\mathscr{G}_0)} = \overline{\bigcup_{i=1}^{n} v_H(g_i(\mathscr{G}_0))} = \bigcup_{i=1}^{n} \overline{v_H(g_i(\mathscr{G}_0))}$$

since $v_G(\mathscr{G}_0)$ is dense in S^1. Since a finite union of nowhere dense sets is nowhere dense and S^1 is not nowhere dense, there is $\alpha_1 \in \{1, 2, \cdots, n\}$ such that $\overline{v_H(g_{\alpha_1}(\mathscr{G}_0))}$ has non-empty interior. Without loss of generality, we may assume $\alpha_1 = 1$. Since $\overline{v_H(g_1(\mathscr{G}_0))}$ has non-empty interior, there is a nondegenerate interval J_1 in S^1 such that $J_1 \subseteq \overline{v_H(g_1(\mathscr{G}_0))}$. Denote $J_1 = (u_1, v_1)_{S^1}$. Since $J_1 \cap v_H(g_1(\mathscr{G}_0))$ is dense in J_1, there is a gap $\mathscr{G}_1$ such that $\mathscr{G}_1 = h_1 g_1(\mathscr{G}_0)$ for some $h_1 \in H$ and $J_1 \cap v(\mathscr{G}_1) \neq \emptyset$. Choose $p_1 \in J_1 \cap v(\mathscr{G}_1)$. By the assumption, there are exactly two elements in $\mathscr{G}_1$ which contain $(u_1, p_1)_{S^1}$ or $(p_1, v_1)_{S^1}$. Then, we can choose an elements K_1 in $\mathscr{G}_1$ such that $K_1 \cap v_H(g_1(\mathscr{G}_0))$ is nowhere dense in K_1. Then,

$$\begin{aligned} K_1 = K_1 \cap S^1 = K_1 \cap \overline{v_G(\mathscr{G}_0)} &= K_1 \cap \overline{\bigcup_{i=1}^{n} v_H(g_i(\mathscr{G}_0))} \\ = K_1 \cap \bigcup_{i=1}^{n} \overline{v_H(g_i(\mathscr{G}_0))} &= \bigcup_{i=1}^{n} K_1 \cap \overline{v_H(g_i(\mathscr{G}_0))}. \end{aligned}$$

Since a finite union of nowhere dense sets is nowhere dense and since K_1 is not nowhere dense, there is an element $\alpha_2 \in \{2, \cdots, n\}$ such that $K_1 \cap \overline{v_H(g_{\alpha_2}(\mathscr{G}_0))}$ has non-empty interior. Without loss of generality, we may assume $\alpha_2 = 2$. Since $K_1 \cap \overline{v_H(g_2(\mathscr{G}_0))}$ has non-empty interior, there is a nondegenerate interval J_2 on K_1 such that $J_2 \subseteq K_1 \cap \overline{v_H(g_2(\mathscr{G}_0))}$. We set $J_2 = (u_2, v_2)_{S^1}$. Since $v_G(\mathscr{G}_0)$ is dense in S^1 and so $E(\mathscr{L})$ is dense in S^1, there is a point q_1 in $E(\mathscr{L}) \cap J_2$. There is a leaf ℓ_1 such that $q_1 \in v(\ell_1)$. By Lemma 10.7.13, there is L_1 in ℓ_1 such that $L_1 \subseteq K_1$. Then, one of $L_1 \cap (u_2, q_1)_{S^1}$ and $L_1 \cap (q_1, v_2)_{S^1}$ is non-empty and so $J_2 \cap L_1$ is non-empty. Likewise, $J_2 \cap L_1^*$ is also non-empty. Since $J_2 \cap v_H(g_2(\mathscr{G}_0))$ is dense in J_2, there is a gap $\mathscr{G}_2$ such that $\mathscr{G}_2 = h_2 g_2(\mathscr{G}_0)$ for some $h_2 \in H$ and $J_2 \cap L_1 \cap v(\mathscr{G}_2) \neq \emptyset$. By Lemma 10.7.13, there is M_1 in $\mathscr{G}_2$ such that $M_1^* \subseteq L_1$. Since $J_2 \cap L_1^*$ is non-empty, $J_2 \cap L_1^* \subseteq L_1^* \subseteq M_1$ and so $M_1 \cap \overline{v_H(\mathscr{G}_2)}$ has non-empty interior. By assumption, this implies that there is K_2 in $\mathscr{G}_2$ such that $K_2 \subseteq M_1^*$ and $K_2 \cap v_H(\mathscr{G}_2)$ is nowhere dense in K_2. Moreover, $K_2 \subseteq M_1^* \subseteq L_1 \subseteq K_1$. Therefore, $K_2 \cap v_H(\mathscr{G}_1)$ and $K_2 \cap v_H(\mathscr{G}_2)$ are nowhere dense in K_2. If $n = 2$, then

$$\begin{aligned} K_2 &= K_2 \cap S^1 \\ &= K_2 \cap \overline{v_G(\mathscr{G}_0)} \end{aligned}$$

$$
\begin{aligned}
&= K_2 \cap \overline{\bigcup_{i=1}^{2} v_H(g_i(\mathscr{G}_0))} \\
&= K_2 \cap \bigcup_{i=1}^{2} \overline{v_H(g_i(\mathscr{G}_0))} \\
&= \bigcup_{i=1}^{2} K_2 \cap \overline{v_H(g_i(\mathscr{G}_0))} \\
&= [K_2 \cap \overline{v_H(g_1(\mathscr{G}_0))}] \cup [K_2 \cap \overline{v_H(g_2(\mathscr{G}_0))}] \\
&= \big[K_2 \cap \overline{v_H(\mathscr{G}_1)}] \cup [K_2 \cap \overline{v_H(\mathscr{G}_2)}].
\end{aligned}
$$

However, this is a contradiction since a finite union of nowhere dense sets is nowhere dense.

If n is greater than 3, choose $m \in \{2, \cdots, n-1\}$. Assume that for each $i \in \{1, 2, \cdots, m\}$, there is a gap $\mathscr{G}_i$ which is $h_i g_i(\mathscr{G}_0)$ for some $h_i \in H$ and that there is K_m in $\mathscr{G}_m$ such that for all $i \in \{1, 2, \cdots, m\}$, $K_m \cap v_H(\mathscr{G}_i)$ are nowhere dense in K_m. Then,

$$
\begin{aligned}
K_m &= K_m \cap S^1 \\
&= K_m \cap \overline{v_G(\mathscr{G}_0)} \\
&= K_m \cap \overline{\bigcup_{i=1}^{n} v_H(g_i(\mathscr{G}_0))} \\
&= K_m \cap \bigcup_{i=1}^{n} \overline{v_H(g_i(\mathscr{G}_0))} \\
&= \bigcup_{i=1}^{n} K_m \cap \overline{v_H(g_i(\mathscr{G}_0))}
\end{aligned}
$$

Since a finite union of nowhere dense sets is nowhere dense and since K_m is not nowhere dense, there is $\alpha_{m+1} \in \{m+1, \cdots, n\}$ such that $K_m \cap \overline{v_H(g_{\alpha_{m+1}}(\mathscr{G}_0))}$ has non-empty interior. Without loss of generality, we may assume $\alpha_{m+1} = m+1$. Since $K_m \cap \overline{v_H(g_{m+1}(\mathscr{G}_0))}$ has non-empty interior, there is a nondegenerate interval J_{m+1} on K_m such that $J_{m+1} \subseteq K_m \cap \overline{v_H(g_{m+1}(\mathscr{G}_0))}$. We set $J_{m+1} = (u_{m+1}, v_{m+1})_{S^1}$. Since $E(\mathscr{L})$ is dense in S^1, there is a point q_m in $E(\mathscr{L}) \cap J_{m+1}$. There is a leaf ℓ_m such that $q_m \in v(\ell_m)$. By Lemma 10.7.13, there exists L_m in ℓ_m such that $L_m \subseteq K_m$. Then, one of $L_m \cap (u_{m+1}, q_m)_{S^1}$ and $L_m \cap (q_m, v_{m+1})_{S^1}$ is non-empty and so $J_{m+1} \cap L_m$ is non-empty. Likewise, $J_{m+1} \cap L_m^*$ is also non-empty. Since $J_{m+1} \cap v_H(g_{m+1}(\mathscr{G}_0))$ is dense in J_{m+1}, there is a gap $\mathscr{G}_{m+1}$ such that $\mathscr{G}_{m+1} = h_{m+1} g_{m+1}(\mathscr{G}_0)$ for some $h_{m+1} \in H$ and $J_{m+1} \cap L_m \cap v(\mathscr{G}_{m+1}) \neq \emptyset$.

By Lemma 10.7.13, there is M_m in $\mathscr{G}_{m+1}$ such that $M_m^* \subseteq L_m$. Since $J_{m+1} \cap L_m^*$ is non-empty, $J_{m+1} \cap L_m^* \subseteq L_m^* \subseteq M_m$ and so $M_m \cap \overline{v_H(\mathscr{G}_{m+1})}$ has non-empty interior. By assumption, this implies that there is K_{m+1} in $\mathscr{G}_{m+1}$ such that $K_{m+1} \subseteq M_m^*$ and $K_{m+1} \cap v_H(\mathscr{G}_{m+1})$ is nowhere dense in K_{m+1}. Moreover, $K_{m+1} \subseteq M_m^* \subseteq L_m \subseteq K_m$. Therefore, for all $i \in \{1, 2, \cdots, m+1\}$, $K_{m+1} \cap v_H(\mathscr{G}_i)$ are nowhere dense in K_{m+1}.

Finally, for each $i \in \{1, 2, \cdots, n\}$, there is a gap $\mathscr{G}_i$ which is $h_i g_i(\mathscr{G}_0)$ for some $h_i \in H$ and there is K_n in $\mathscr{G}_n$ such that for all $i \in \{1, 2, \cdots, n\}$, $K_n \cap v_H(\mathscr{G}_i)$ are nowhere dense in K_n. However,

$$\begin{aligned}
K_n &= K_n \cap S^1 \\
&= K_n \cap \overline{v_G(\mathscr{G}_0)} \\
&= K_n \cap \overline{\bigcup_{i=1}^{n} v_H(g_i(\mathscr{G}_0))} \\
&= K_n \cap \bigcup_{i=1}^{n} \overline{v_H(g_i(\mathscr{G}_0))} \\
&= \bigcup_{i=1}^{n} K_n \cap \overline{v_H(g_i(\mathscr{G}_0))} \\
&= \bigcup_{i=1}^{n} K_n \cap \overline{v_H(\mathscr{G}_i)}
\end{aligned}$$

so this is a contradiction since a finite union of nowhere dense sets is nowhere dense. We are done. □

Let us prove the main theorem.

Theorem 10.8.9 *Let G be a subgroup of* $\text{Homeo}^+(S^1)$ *and $\mathscr{L}$ be a G-invariant lamination system. Suppose that there is an ideal polygon $\mathscr{G}_0$ on $\mathscr{L}$ which is not a leaf. If $v_G(\mathscr{G}_0)$ is dense in S^1, then G is not virtually abelian.*

Proof Suppose that H is a finite index subgroup of G. By Lemma 10.8.8, there is a gap $\mathscr{G}$ which is $g(\mathscr{G}_0)$ for some $g \in G$ and has three elements I_1, I_2 and I_3 such that for all $i \in \mathbb{Z}_3$, $I_i \cap \overline{v_H(\mathscr{G})}$ has non-empty interior on S^1. Since $\overline{v_H(\mathscr{G})}$ has non-empty interior, we can define $m : S^1 \to S^1$ as the monotone map which collapses each closure of connected component of $S^1 - \overline{v_H(\mathscr{G})}$. Then, for each element $h \in H$, there is a unique element g_h in $\text{Homeo}^+(S^1)$ which makes the following diagram commute:

$$\begin{array}{ccc}
S^1 & \xrightarrow{m} & S^1 \\
\downarrow{h} & & \downarrow{g_h} \\
S^1 & \xrightarrow{m} & S^1
\end{array}$$

since $\overline{v_H(S^1)}$ is preserved by the action of H. Define $G_H \equiv \{g_h \in \mathrm{Homeo}^+(S^1) : h \in H\}$. Then, G_H is isomorphic to some quotient group of H. Let us define $\mathscr{L}_H$ as the family of nondegenerate open intervals $(u, v)_{S^1}$ such that there is I in $\mathscr{L}$ such that $m(v(\ell(I))) = \{u, v\}$. By the construction of $\mathscr{L}_H$, $\mathscr{L}_H$ is a G_H-invariant lamination system. Moreover, since $\mathscr{G}$ has three elements I_1, I_2 and I_3 such that for all $i \in \mathbb{Z}_3$, $I_i \cap \overline{v_H(\mathscr{G})}$ has non-empty interior, there is a non-leaf ideal polygon $\mathscr{G}_H$ in $\mathscr{L}_H$ such that $m(v(\mathscr{G})) = v(\mathscr{G}_H)$. By the construction of G_H and $\mathscr{L}_H$, $v_{G_H}(\mathscr{G}_H)$ is dense in S^1. Therefore, by Theorem 10.8.7, G_H is non-abelian and so H is also non-abelian. Thus, G is not virtually abelian. □

10.9 Existence of a Non-abelian Free Subgroup in the Tight Pairs

In 2001, Calegari wrote a set of lecture notes entitled 'Foliations and the geometrization of 3-manifolds' [8], and later a large chunk of these notes became the book [10]. In these notes, Calegari introduced the notion of a tight pair to study special types of laminar groups. We rephrase the definition below in terms of lamination systems.

Definition 10.9.1 Let G be a subgroup of $\mathrm{Homeo}^+(S^1)$, and $\mathscr{L}$ be a G-invariant lamination system. The pair $(\mathscr{L}, G)$ is *tight* if $\mathscr{L}$ is very full and totally disconnected and if for each $I \in \mathscr{L}$, $\ell(I)$ is not isolated, G acts on $\mathscr{L}$ minimally and the set of non-leaf gaps consists of finitely many orbit classes under this action.

There is a familiar example which comes from the geodesic invariant laminations of a peusdo-Anosov element of the mapping class group of a closed hyperbolic surface. For example, take a pseudo-Anosov element $[\varphi] \in \mathrm{Mod}(S_2)$ where S_2 is a genus 2 hyperbolic surface. By the definition of a pseudo-Anosov element, there are two invariant geodesic laminations Λ^+(unstable) and Λ^-(stable). Since S_2 is hyperbolic, the universal covering is $\mathbb{H}^2$. The preimages of $\Lambda^\pm$ under the covering map are two geodesic laminations of $\mathbb{H}^2$. Each of these laminations satisfies the following properties.

(1) It is very full,
(2) there is no isolated leaf,
(3) it is totally disconnected, that is, there is no foliated open disk on $\mathbb{H}^2$,
(4) the action of $\pi_1(S_2)$ is minimal, that is, the orbit of any leaf is dense in the lamination, and
(5) the set of complementary regions falls into finitely many orbit classes under the action of $\pi_1(S_2)$.

So, we can see that the lamination systems induced from these laminations with $\pi_1(S_2)$ are tight pairs.

In [8], Calegari showed that there are two types of tight pairs, sticky pairs and slippery pairs. $(\mathscr{L}, G)$ is a sticky pair if every gap of $\mathscr{L}$ has a vertex shared with another non-leaf gap, and is a slippery pair if no non-leaf gap of $\mathscr{L}$ shares a vertex with other gaps. He constructed a dual $\mathbb{R}$-tree to $\mathscr{L}$ in the case of a sticky pair. By analyzing the G-action on this dual tree, the following theorem was obtained.

Theorem 10.9.2 (Calegari [8]) *Suppose $(\mathscr{L}, \pi_1(M))$ is a sticky pair for some closed irreducible 3-manifold M. Then M is Haken.*

Full details of the proof of the above theorem are also presented in the Master's thesis of Te Winkel [32]. In this section, we study a general feature of tight pairs.

Proposition 10.9.3 *Let $(\mathscr{L}, G)$ be a tight pair. Then for any leaf ℓ, $v_G(\ell) = \bigcup_{g\in G} v(g(\ell))$ is dense in S^1.*

Proof Suppose that there is a leaf ℓ of $\mathscr{L}$ such that $v_G(\ell)$ is not dense in S^1. Then there is a connected component K of $S^1 - \overline{v_G(\ell)}$. Since, by Corollary 10.7.15, $E(\mathscr{L})$ is dense in S^1, there is a p in $E(\mathscr{L}) \cap K$ and so there is a leaf ℓ' of $\mathscr{L}$ with $p \in v(\ell')$. Since the action of G is minimal, there is a sequence $\{g_n\}_{n=1}^{\infty}$ of G such that $g_n(\ell) \to \ell'$. Then there is a sequence $\{J_n\}_{n=1}^{\infty}$ of $\mathscr{L}$ such that for all $n \in \mathbb{N}$, $g_n(\ell) = \ell(J_n)$ and

$$I' \subseteq \liminf J_n \subseteq \limsup J_n \subseteq \overline{I'}$$

for some $I' \in \ell'$. By the choice of K, for each $n \in \mathbb{N}$, we have either $K \subseteq J_n$ or $K \subseteq J_n^*$. Note that $I' \cap K$ is not empty. Choose $q \in I' \cap K$. Since $I' \subseteq \liminf J_n$, there is N in $\mathbb{N}$ such that $q \in \bigcap_{n=N}^{\infty} J_n$. Therefore, for any $n \geq N$, $q \in K \subseteq J_n$ and so $K \subseteq \bigcap_{n=N}^{\infty} J_n \subseteq \liminf J_n$. However, K is not contained in $\overline{I'}$ and so this is a contradiction. □

Corollary 10.9.4 *Let $(\mathscr{L}, G)$ be a tight pair. Then for any non-leaf gap $\mathscr{G}$, $v_G(\mathscr{G})$ is dense in S^1.*

Proposition 10.9.5 *Let $(\mathscr{L}, G)$ be a tight pair. There is a non-leaf gap $\mathscr{G}$.*

Proof Since $\mathscr{L}$ is not empty, there exists an element $I \in \mathscr{L}$. By the definition of a lamination system, $I^* \in \mathscr{L}$. By Corollary 10.7.15, there is a p in $E(\mathscr{L}) \cap I^*$. So there is a $J \in \mathscr{L}$ such that $p \in v(\ell(J))$. If $J \subseteq I$, then $p \in \overline{J} \subseteq \overline{I}$ and this is a contradiction since $p \in I^*$. If $J^* \subseteq I$, then $p \in \overline{J^*} \subseteq \overline{I}$ and this is also a contradiction since $p \in I^*$. Therefore, either $I \subsetneq J$ or $I \subsetneq J^*$. So, $\{I, J^*\}$ or $\{I, J\}$ is a distinct pair, respectively. Thus, since $\mathscr{L}$ is totally disconnected, there is a non-leaf gap which makes the distinct pairs separated. □

By Theorem 10.8.9, tight pairs are not virtually abelian. Our goal here is to show that a tight pair actually contains a non-abelian free subgroup as long as it does not admit a global fixed point. We will use the following famous theorem of Margulis which is analogous to the Tits alternative.

Theorem 10.9.6 (Margulis [25]) *Let G be a subgroup of* $\mathrm{Homeo}^+(S^1)$. *At least one of the following properties holds:*

(1) *G contains a non-abelian free subgroup.*
(2) *There is a Borel probability measure on the circle which is G-invariant.*

Let μ be a Borel probability measure on S^1. We define the *support* of μ as the complement of the union of measure zero open sets and denote it as $supp(\mu)$. We have the following facts:

(1) $supp(\mu)$ is a closed subset of S^1.
(2) For each $p \in supp(\mu)$ and each open neighborhood U of p, $\mu(U) > 0$.
(3) If μ is also G-invariant where G is a subgroup of $\mathrm{Homeo}^+(S^1)$, then $supp(\mu)$ is also G-invariant, that is, for each $g \in G$, $g(supp(\mu)) = supp(\mu)$.

Lemma 10.9.7 *Let $(\mathscr{L}, G)$ be a tight pair. Suppose that there is a Borel probability measure μ on S^1 which is G-invariant. Then for each non-leaf gap $\mathscr{G}$ of $\mathscr{L}$, there is a unique element I in $\mathscr{G}$ such that $\mu(I) = 1$.*

Proof Let $\mathscr{G}$ be a non-leaf gap. First, we want to show that there are at most two positive measure elements in $\mathscr{G}$. Suppose that there are three elements I_0, I_1 *and* I_2 in $\mathscr{G}$ which are positive measure. Say that $\{I_i\}_{i\in\mathbb{Z}_3}$ and choose $i \in \mathbb{Z}_3$. By Lemma 10.7.19, there is a $g_i \in G$ such that for any $J \in g_i(\mathscr{G})$, $\ell(J)$ properly lies on I_i and by Lemma 10.7.13, there exist L_i in $\mathscr{G}$ and L'_i in $g_i(\mathscr{G})$ such that $(L'_i)^* \subseteq L_i$. Since for all $J \in g_i(\mathscr{G})$ which is not L'_i, $J \subseteq L_i$ and so $\ell(J)$ lies on L_i, so $L_i = I_i$. If $g_i(I_i) \neq L'_i$, then $g_i(I_i) \subseteq (L'_i)^* \subseteq I_i$. Then, at least one of $g_i(I_{i+1})$ and $g_i(I_{i+2})$ is contained in I_i. If, for some $j \in \mathbb{Z}_3 - \{i\}$, $g(I_i) \cup g(I_j) \subseteq I_i$, then $\mu(g(I_i)) + \mu(g(I_j)) \leq \mu(I_i)$ and since μ is G-invariant, $\mu(I_i) + \mu(I_j) \leq \mu(I_i)$, and so $\mu(I_j) \leq 0$. This is a contradiction since $0 < \mu(I_j)$. Therefore, $g_i(I_i) = L'_i$. Then for all $i \in \mathbb{Z}_3$, $g_i(I_{i+1}) \cup g_i(I_{i+2}) \subseteq g_i(I_i)^* = (L'_i)^* \subseteq L_i = I_i$ and so $\mu(I_{i+1}) + \mu(I_{i+2}) = \mu(g_i(I_{i+1})) + \mu(g_i(I_{i+2})) \leq \mu(I_i)$. However,

$$\mu(I_1) \geq \mu(I_2) + \mu(I_3) \geq \{\mu(I_3) + \mu(I_1)\} + \{\mu(I_1) + \mu(I_2)\}$$

and so

$$0 \geq \mu(I_1) + \mu(I_2) + \mu(I_3).$$

This is a contradiction since $\mu(I_1) + \mu(I_2) + \mu(I_3) > 0$. Therefore, there are at most two positive measure elements in $\mathscr{G}$. This implies that there is at least one measure-zero element J in $\mathscr{G}$ since $\mathscr{G}$ is a non-leaf gap. Now, we show that there is an element I such that $\mu(I) = 1$. By Lemma 10.7.19, there is $g \in G$ such that for any $K \in g(\mathscr{G})$, $\ell(K)$ properly lies on J and by Lemma 10.7.13, there exist L in $\mathscr{G}$ and L' in $g(\mathscr{G})$ such that $(L')^* \subseteq L$. Then, $L = J$ and $\overline{(L')^*} \subseteq L = J$. So $\mu((L')^c) = \mu(\overline{(L')^*}) \leq \mu(J) = 0$ and this implies that $\mu(L') = 1$. Thus, L' is the element I which we wanted. ◻

Now, we can get the same result in a leaf as the following lemma says.

Lemma 10.9.8 *Let $(\mathscr{L}, G)$ be a tight pair. Suppose that there is a Borel probability measure μ on S^1 which is G-invariant. Then for each leaf ℓ, there is a unique element I in ℓ such that $\mu(I) = 1$.*

Proof Let ℓ be a leaf of $\mathscr{L}$. By Proposition 10.9.5, there is a non-leaf gap $\mathscr{G}$ of $\mathscr{L}$. By the definition of gaps, ℓ lies on an element J of $\mathscr{G}$. Say that $\ell = \ell(I)$ and $I \subseteq J$. Choose K in $\mathscr{G}$ which is not J. By Corollary 10.9.4 and Lemma 10.7.19, there is g in G such that $v(g(\mathscr{G})) \subseteq K$. There is K' in $g(\mathscr{G})$ such that $\overline{K'} \subseteq K$. By Proposition 10.9.3, there is g' such that $v(g'(\ell)) \cap K' \neq \emptyset$. Choose p in $v(g'(\ell)) \cap K'$. If $K' \subseteq g'(I)$, then $p \in \overline{g'(I^*)} \subseteq \overline{(K')^*} = (K')^c$ which is a contradiction since $p \in K'$. If $K' \subseteq g'(I^*)$, then $p \in \overline{g'(I)} \subseteq \overline{(K')^*}$ which is also a contradiction since $p \in K'$. Therefore, we have either $g'(I) \subsetneq K'$ or $g'(I^*) \subsetneq K'$. So, either $\overline{g'(I)} \subseteq K$ or $\overline{g'(I^*)} \subseteq K$.

First, if I is positive measure, then J is also positive measure and, by Lemma 10.9.7, $\mu(J) = 1$. Moreover, K is of measure zero. So $\overline{g'(I^*)} \subseteq K$ and $\mu(\overline{I^*}) = \mu(g'(\overline{I^*})) = \mu(\overline{g'(I^*)}) \leq \mu(K) = 0$. Therefore, $\mu(\overline{I^*}) = 0$ and so $\mu(I) = 1$.

Next, assume that $\mu(I) = 0$. If $\mu(J) = 0$, then by Lemma 10.9.7 $\mu(\overline{J}) = 0$, and so $\mu(\overline{I}) \leq \mu(\overline{J}) = 0$. Therefore, $\mu(\overline{I}) = 0$ and so $\mu(I^*) = 1$. If $\mu(J) = 1$, then $\mu(K) = 0$ by Lemma 10.9.7. Since $1 = \mu(\overline{I^*}) = \mu(g'(\overline{I^*})) = \mu(\overline{g'(I^*)})$, $\overline{g'(I^*)} \subseteq K$ is not possible and so $\overline{g'(I)} \subseteq K$ is the possible case. Therefore, $\mu(\overline{g'(I)}) \leq \mu(K) = 0$, and so $\mu(\overline{I}) = 0$. Thus, $\mu(I^*) = 1$. □

Finally, we prove the main theorem.

Theorem 10.9.9 *Let $(\mathscr{L}, G)$ be a tight pair. Suppose that there is a Borel probability measure μ on S^1 which is G-invariant. Then, the support $supp(\mu)$ of the measure μ is a one point set.*

Proof Let p be a point in $supp(\mu)$. First, if $p \in E(\mathscr{L})$, then there is a leaf ℓ with $p \in v(\ell)$. By Lemma 10.9.8, there is a unique element I in ℓ such that $\mu(I) = 1$. So, $supp(\mu) \cap I^* = \emptyset$ by the definition of the support.

By Proposition 10.9.5, there is a non-leaf gap $\mathscr{G}$ and by Corollary 10.9.4, $v_G(\mathscr{G})$ is dense in S^1. So, there is g in G such that $v(g(\mathscr{G})) \cap I^* \neq \emptyset$. Moreover, by Lemma 10.7.13 , there is J in $g(\mathscr{G})$ such that $J^* \subseteq I^*$. Since I^* is measure zero, $\mu(J) = 1$ by Lemma 10.9.8. And since $g(\mathscr{G})$ is a non-leaf gap, there is K in $g(\mathscr{G})$ such that $K \subseteq J^*$ and $\mu(K) = 0$. Then by Corollary 10.9.4 and Lemma 10.7.19, there is h in G such that $v(h(g(\mathscr{G}))) \subseteq K$. Therefore, we can choose L in $h(g(\mathscr{G}))$ such that $\overline{L} \subset K$ and so $\overline{L} \subset I^*$.

By Proposition 10.9.3, $v_G(\ell)$ is dense so there is k in G such that $v(k(\ell)) \cap L \neq \emptyset$. Then by Lemma 10.7.13, $M \subseteq L$ for some $M \in k(\ell)$ which implies $k(p) \in v(k(\ell)) \subset \overline{L} \subset I^*$. However, since $k(p) \in supp(\mu)$, $0 < \mu(I^*)$ which is a contradiction. Thus $p \notin E(\mathscr{L})$.

So, by Lemma 10.7.14, there is a rainbow $\{I_n\}_{n=1}^{\infty}$ at p. Applying Lemma 10.9.8 to each $\ell(I_n)$, since $p \in I_n$, $\mu(I_n) = 1$ for all $n \in \mathbb{N}$. Therefore, $\mu(\{p\}) = \mu(\bigcap_{n=1}^{\infty} I_n) = \lim_{n\to\infty} \mu(I_n) = 1$ since μ is a finite measure. Thus, $supp(\mu) = \{p\}$. □

The following is an immediate corollary of the above theorem, since if there are more than one global fixed point, one can find an invariant probability measure supported on those points.

Corollary 10.9.10 *Let $(\mathscr{L}, G)$ be a tight pair. There is at most one global fixed point.*

Now we state the main result of this section.

Corollary 10.9.11 *Let $(\mathscr{L}, G)$ be a tight pair without global fixed points. Then, G contains a non-abelian free subgroup.*

Proof Suppose that there is a G-invariant Borel probability measure μ. Then by Lemma 10.9.9 $supp(\mu)$ is a one point set. Since $supp(\mu)$ is G-invariant, so the element of $supp(\mu)$ is a global fixed point. By assumption, this is a contradiction. Therefore, there is no such measure. By Theorem 10.9.6, G contains a non-abelian free subgroup. □

10.10 Loose Laminations

A very full lamination system is *loose* if for any two non-leaf gaps $\mathscr{G}$ and $\mathscr{G}'$ with $\mathscr{G} \neq \mathscr{G}'$, $v(\mathscr{G}) \cap v(\mathscr{G}') = \emptyset$. There are equivalent conditions in totally disconnected very full lamination systems.

Lemma 10.10.1 ([1]) *Let $\mathscr{L}$ be a totally disconnected very full lamination system. Then $\mathscr{L}$ is loose if and only if the following conditions are satisfied:*

(1) *for each $p \in S^1$, at most finitely many leaves of $\mathscr{L}$ have p as an endpoint.*
(2) *There are no isolated leaves.*

A group acting on the circle with two loose invariant laminations with certain conditions is called a pseudo-fibered triple. It was observed in the first author's PhD thesis [2] that each nontrivial element in the pseudo-fibered triple has at most finitely many fixed points under the assumption that the fixed point set is countable, hence countability of the fixed point sets is an underlying assumption in [1]. This section should serve as an appendix to [1] in which we prove that additional assumption that the fixed point sets are countable is not necessary.

In this section, we consider a pseudo-fibered triple which is a triple $(\mathscr{L}_1, \mathscr{L}_2, G)$ in which G is a finitely generated subgroup of $\mathrm{Homeo}^+(S^1)$, each nontrivial element of G has at most countably many fixed points in S^1 and $\mathscr{L}_i$ are G-invariant very full loose lamination systems with $E(\mathscr{L}_1) \cap E(\mathscr{L}_2) = \emptyset$. Indeed, without the fixed point condition of G, we can induce the original definition, that is, each nontrivial element of G has finitely many fixed points. Let us begin with a weaker version of the definition of a pseudo-fibered triple.

Definition 10.10.2 Let G be a finitely generated subgroup of $\mathrm{Homeo}^+(S^1)$, and $\mathscr{L}_1$ and $\mathscr{L}_2$ be two G-invariant lamination system. Then a triple $(\mathscr{L}_1, \mathscr{L}_2, G)$ is *pseudo-fibered* if $\mathscr{L}_1$ and $\mathscr{L}_2$ are very full loose lamination systems with $E(\mathscr{L}_1) \cap E(\mathscr{L}_2) = \emptyset$.

The disjoint endpoints condition of two lamination systems implies totally disconnectedness on lamination systems.

Proposition 10.10.3 ([1]) *Let $(\mathscr{L}_1, \mathscr{L}_2, G)$ be a pseudo-fibered triple. Then $\mathscr{L}_1$ and $\mathscr{L}_2$ are totally disconnected.*

Proof This follows from Corollary 10.7.15 and Lemma 10.7.17. □

The following proposition says that there is no sticky leaf on two lamination systems.

Proposition 10.10.4 *Let $(\mathscr{L}_1, \mathscr{L}_2, G)$ be a pseudo-fibered triple, and $\mathscr{G}$ be a leaf in $\mathscr{L}_1$. For each $g \in G$, we have either $v(\mathscr{G}) \subseteq Fix_g$ or $v(\mathscr{G}) \subseteq S^1 - Fix_g$.*

Proof Fix g in G. If $Fix_g = \emptyset$ or $Fix_g = S^1$, then this is obvious. Assume that $Fix_g \neq \emptyset$ and $Fix_g \neq S^1$. We set $\mathscr{G} = \ell((u, v)_{S^1})$. If $u \in Fix_g$ and $v \in S^1 - Fix_g$, then for each $n \in \mathbb{Z}$, $g^n(\ell((u, v)_{S^1})) = \ell((g^n(u), g^n(v))_{S^1}) = \ell((u, g^n(v))_{S^1})$. Since $g(v) \neq v$, we have either $(u, v)_{S^1} \subsetneq g((u, v)_{S^1}) = (u, g(v))_{S^1}$ or $(u, g(v))_{S^1} = g((u, v)_{S^1}) \subsetneq (u, v)_{S^1}$. Therefore, there are infinitely many leaves $\{g^n(\mathscr{G}) |\ n \in \mathbb{Z}\}$ in which u is an endpoint. However, by Proposition 10.10.3, $\mathscr{L}_1$ and $\mathscr{L}_2$ are totally disconnected and so we can apply Lemma 10.10.1 to $\mathscr{L}_1$. This is a contradiction. If $v \in Fix_g$ and $u \in S^1 - Fix_g$, we can make a same argument with $\mathscr{G} = \ell((v, u)_{S^1})$. Thus, we are done. □

With this proposition, we analyze the complement of the fixed points set of a nontrivial element of G. First, we recall the following lemma.

Lemma 10.10.5 ([3]) *Let g be a non-trivial orientation-preserving homeomorphism on S^1 with $3 \leq |Fix_g|$. Then any very full lamination system $\mathscr{L}$ which is $\langle g \rangle$-invariant has a leaf ℓ in $\mathscr{L}$ such that $v(\ell) \subseteq Fix_g$. Moreover, for any connected component I of $S^1 - Fix_g$ with $I = (a, b)_{S^1}$, at least one of a and b is an endpoint of a leaf of $\mathscr{L}$.*

We can interpret this lemma in a pseudo-fibered triple setting as the following proposition.

Proposition 10.10.6 *Let $(\mathscr{L}_1, \mathscr{L}_2, G)$ be a pseudo-fibered triple and g a nontrivial element of G with $3 \leq |Fix_g|$. For any connected component $(u, v)_{S^1}$ of $S^1 - Fix_g$, $u \in E(\mathscr{L}_i)$ and $v \in E(\mathscr{L}_j)$ with $i \neq j \in \{1, 2\}$.*

Proof By Lemma 10.10.5 and the condition $E(\mathscr{L}_1) \cap E(\mathscr{L}_2) = \emptyset$, this is obvious. □

Proposition 10.10.7 *Let* $(\mathscr{L}_1, \mathscr{L}_2, G)$ *be a pseudo-fibered triple and* g *a nontrivial element of* G*. Suppose that there are two distinct connected components* I_1 *and* I_2 *of* $S^1 - Fix_g$ *such that* $\overline{I_1} \cap \overline{I_2} = \emptyset$*. Then, for each* $i \in \{1, 2\}$*, there is no leaf* ℓ *of* $\mathscr{L}_i$ *such that* $|v(\ell) \cap I_1| = |v(\ell) \cap I_2| = 1$.

Proof Let $I_1 = (u_1, v_1)_{S^1}$ and $I_2 = (u_2, v_2)_{S^1}$. Since $\overline{I_1} \cap \overline{I_2} = \emptyset$, $|\{u_1, v_1, u_2, v_2\}| = 4$ and since $\{u_1, v_1, u_2, v_2\} \subseteq Fix_g$, $4 \le |Fix_g|$. Then we can apply Proposition 10.10.6 to $(u_1, v_1)_{S^1}$ and so $u_1 \in E(\mathscr{L}_i)$ and $v_1 \in E(\mathscr{L}_j)$ with $i \neq j \in \{1, 2\}$. Assume that there is an element I in $\mathscr{L}_1 \cup \mathscr{L}_2$ such that $[v_1, u_2]_{S^1} \subseteq I \subset \overline{I} \subset (u_1, v_2)_{S^1}$. Since $I \in C_{v_1}^{(u_1,v_2)_{S^1}}$ and $C_{v_1}^{(u_1,v_2)_{S^1}}$ is preserved by g and linearly ordered by the set inclusion by Proposition 10.7.9, we have either $I \subseteq g(I)$ or $g(I) \subseteq I$. Since $\partial I \subseteq S^1 - Fix_g$, we have either $\overline{I} \subseteq g(I)$ or $g(\overline{I}) \subseteq I$. So one of the two sequences $\{g^n(\ell(I))\}_{n=1}^{\infty}$ and $\{g^{-n}(\ell(I))\}_{n=1}^{\infty}$ converges to $(v_1, u_2)_{S^1}$ and the other converges to $(u_1, v_2)_{S^1}$. This implies that $\{u_1, v_1\} \subseteq E(\mathscr{L}_i)$ for some $i \in \{1, 2\}$. But this is a contradiction since $E(\mathscr{L}_1) \cap E(\mathscr{L}_2) = \emptyset$. Therefore there is no such I. So we are done. □

Proposition 10.10.8 *Let* $(\mathscr{L}_1, \mathscr{L}_2, G)$ *be a pseudo-fibered triple and* g *a nontrivial element of* G *with* $3 \le |Fix_g|$*. For any connected component* I *of* $S^1 - Fix_g$*, each point of* ∂I *is isolated in* Fix_g.

Proof We set $I = (u, v)_{S^1}$. Then by Proposition 10.10.6, $u \in E(\mathscr{L}_i)$ and $v \in E(\mathscr{L}_j)$ with $i \neq j \in \{1, 2\}$. Let us say that $u \in E(\mathscr{L}_1)$ and $v \in E(\mathscr{L}_2)$. Since $E(\mathscr{L}_1) \cap E(\mathscr{L}_2) = \emptyset$, $u \notin E(\mathscr{L}_2)$ and so by Lemma 10.7.14, u has a rainbow $\{I_n\}_{n=1}^{\infty}$ in $\mathscr{L}_2$. Since $3 \le |Fix_g|$, there is a fixed point w in $Fix_g - \{u, v\}$. Note that $w \in (v, u)_{S^1}$. Since $\bigcap_{n=1}^{\infty} I_n = \{u\}$, there is I_N in $\{I_n\}_{n=1}^{\infty}$ such that $\{v, w\} \subseteq I_N^*$.

So, we can assume that $u \in I_n$ and $\{v, w\} \subseteq I_n^*$ for all $n \in \mathbb{N}$ and set $I_n = (u_n, v_n)_{S^1}$. Then, we want to show that for all $n \in \mathbb{N}$, $v_n \in (u, v)_{S^1}$ and $u_n \in (w, u)_{S^1}$. Since $\varphi(u_n, u, v_n) = 1$ and $\varphi(v_n, v, u_n) = \varphi(u_n, v_n, v) = 1$, by the cocycle condition on the four points (u_n, u, v_n, v),

$$\varphi(u, v_n, v) - \varphi(u_n, v_n, v) + \varphi(u_n, u, v) - \varphi(u_n, u, v_n) = \varphi(u, v_n, v) - 1 + \varphi(u_n, u, v) - 1 = 0$$

and so the only possible case is $\varphi(u, v_n, v) = \varphi(u_n, u, v) = 1$. On the other hand, since $\varphi(u_n, u, v_n) = 1$ and $\varphi(v_n, w, u_n) = \varphi(u_n, v_n, w) = 1$, by the cocycle conditions on the four points (u_n, u, v_n, w),

$$\varphi(u, v_n, w) - \varphi(u_n, v_n, w) + \varphi(u_n, u, w) - \varphi(u_n, u, v_n) = \varphi(u, v_n, w) - 1 + \varphi(u_n, u, w) - 1 = 0$$

and so the only possible case is $\varphi(u, v_n, w) = \varphi(u_n, u, w) = 1$. Therefore, for all $n \in \mathbb{N}$, $v_n \in (u, v)_{S^1}$ and $u_n \in (w, u)_{S^1}$ since $\varphi(u, v_n, v) = 1$ and $\varphi(w, u_n, u) = \varphi(u_n, u, w) = 1$.

Fix n in $\mathbb{N}$. Since $v_n \in (u, v)_{S^1} \subseteq S^1 - Fix_g$, by Proposition 10.10.4, $u_n \in S^1 - Fix_g$ and so there is a unique connected component J of $S^1 - Fix_g$ which contains

u_n. Since $u_n \in (w,u)_{S^1}$, we have $J \subseteq (w,u)_{S^1}$ and by Proposition 10.10.7, $\overline{J} \cap [u,v]_{S^1} \neq \emptyset$. Since $v \neq w$, $\overline{J} \cap [u,v]_{S^1} \subseteq [w,u]_{S^1} \cap [u,v]_{S^1} = \{u\}$ and so $\overline{J} \cap [u,v]_{S^1} = \{u\}$. Therefore, u is isolated in Fix_g. Likewise, v is isolated in Fix_g. □

Now, we prove lemmas which will be used in the proof of the main theorem.

Lemma 10.10.9 *Let $(\mathscr{L}_1, \mathscr{L}_2, G)$ be a pseudo-fibered triple and g a nontrivial element of G with $4 \leq |Fix_g|$. Suppose that there is an isolated fixed point p of g. Then there is an element I in $\mathscr{L}_1 \cup \mathscr{L}_2$ such that $p \in v(\ell(I))$ and $|I \cap Fix_g| = 1$.*

Proof Since p is an isolated fixed point and $4 \leq |Fix_g|$, there is a connected component $(p,q)_{S^1}$ of $S^1 - Fix_g$ which is a nondegenerate open interval. By Proposition 10.10.8, q is also an isolated fixed point. So there is a connected component $(q,r)_{S^1}$ of $S^1 - Fix_g$. Since $4 \leq |Fix_g|, r \neq p$. By Proposition 10.10.6, $\{p,r\} \subseteq E(\mathscr{L}_i)$ and $q \in E(\mathscr{L}_j)$ with $i \neq j \in \{1,2\}$. Say that $\{p,r\} \subseteq E(\mathscr{L}_1)$ and $q \in E(\mathscr{L}_2)$. Since $E(\mathscr{L}_1)$ and $E(\mathscr{L}_2)$ are disjoint, there is a rainbow $\{I_n\}_{n=1}^{\infty}$ at q in $\mathscr{L}_1$ by Theorem 10.7.14. Since $\bigcap_{n=1}^{\infty} I_n = \{q\}$, there is I_N in $\{I_n\}_{n=1}^{\infty}$ such that $\{p,r\} \subseteq I_N^*$. We obtain that $q \in I_N \subset \overline{I_N} \subseteq (p,r)_{S^1}$. Since $C_q^{(p,r)_{S^1}}$ on $\mathscr{L}_1$ is linearly ordered and preserved by g, $g(I_N) \subseteq I_N$ or $I_N \subseteq g(I_N)$. Note that ∂I_N is contained in $S^1 - Fix_g$. So, $g(\overline{I_N}) \subset I_N$ or $\overline{I_N} \subset g(I_N)$. Then, one of the two sequences $\{g^k(\ell(I_N))\}_{k=1}^{\infty}$ and $\{g^{-k}(\ell(I_N))\}_{k=1}^{\infty}$ converges to $(p,r)_{S^1}$ on $\mathscr{L}_1$. Therefore, $(p,r)_{S^1} \in \mathscr{L}_1$. □

Lemma 10.10.10 *Let $(\mathscr{L}_1, \mathscr{L}_2, G)$ be a pseudo-fibered triple and g a nontrivial element of G with $5 \leq |Fix_g|$. Suppose that there is an isolated fixed point p of g. If I is an element in $\mathscr{L}_1 \cup \mathscr{L}_2$ such that $p \in v(\ell(I))$ and $|I \cap Fix_g| = 1$, then I^* is isolated.*

Proof Without loss of generality, we can assume that there is such an element I in $\mathscr{L}_1$. By Proposition 10.10.4, $v(\ell(I)) \subset Fix_g$ since $p \in Fix_g$. Say $I = (u,v)_{S^1}$ and denote the fixed point in I by q. By Proposition 10.10.8, there are two connected components $(x,u)_{S^1}$ and $(v,y)_{S^1}$ of $S^1 - Fix_g$. Since $5 \leq |Fix_g|, x \neq y$.

Suppose that there is an I^*-side sequence $\{\ell_n\}_{n=1}^{\infty}$ on $\mathscr{L}_1$. There is a sequence $\{I_n\}_{n=1}^{\infty}$ on $\mathscr{L}_1$ such that $\ell_n = \ell(I_n)$ for all $n \in \mathbb{N}$ and $I^* \subseteq \liminf I_n \subseteq \limsup I_n \subseteq \overline{I^*}$. Since $I^* \subseteq \liminf I_n$, there is an N such that $\{x,y\} \subseteq \bigcap_{n=N}^{\infty} I_n$ and since $\limsup I_n \subseteq \overline{I^*}$, there is an N' such that $q \notin \bigcup_{n=N'}^{\infty} I_n$.

Fix n with $n > \max\{N, N'\}$. Then by the choice of n, $\{x,y\} \subseteq I_n$ and $q \notin I_n$. From now on, we show $I_n \subsetneq I^*$. If $I_n^* \subseteq I^*$, then $q \in I \subseteq I_n$ which is a contradiction since $q \notin I_n$. If $I^* \subseteq I_n^*$, then $\{x,y\} \subseteq I_n \subseteq I$ which is a contradiction since $\{x,y\} \subseteq I^*$. Hence, $I_n \subseteq I^*$ or $I^* \subseteq I_n$. Then, since

$\ell(I) \neq \ell(I_n)$, $I_n \subsetneq I^*$ or $I^* \subsetneq I_n$. Therefore, since $\ell(I_n)$ lies on I^*, $I_n \subsetneq I^*$ is the only possible case.

Say $I_n = (a, b)_{S^1}$. By assumption, $\{x, y\} \subseteq I_n$. So, we want to show that $[y, x]_{S^1} \subset (a, b)_{S^1}$. Choose $z \in (y, x)_{S^1}$. Since $z \in (y, x)_{S^1}$ and $q \in (x, y)_{S^1}$, we get $\varphi(y, z, x) = 1$ and $\varphi(x, q, y) = 1$, respectively. This implies that $\varphi(x, q, z) = 1$ and $\varphi(y, z, q) = 1$. Since $\{x, y\} \subseteq I_n$ and $q \in I \subseteq I_n^*$, we get $\varphi(a, x, b) = 1$ and $\varphi(b, q, a) = 1$, respectively. This implies $\varphi(b, q, x) = 1$. Likewise, since $\{x, y\} \subseteq I_n$ and $q \in I_n^*$, we get $\varphi(a, y, b) = 1$ and $\varphi(b, q, a) = 1$, respectively. This implies $\varphi(a, y, q) = 1$. Therefore, since $\varphi(x, q, z) = \varphi(b, q, x) = 1$, $\varphi(q, z, b) = 1$ and since $\varphi(y, z, q) = \varphi(a, y, q) = 1$, $\varphi(q, a, z) = 1$. Finally, by applying the cocycle condition to four points (a, z, q, b),

$$\varphi(z, q, b) - \varphi(a, q, b) + \varphi(a, z, b) - \varphi(a, z, q) = 0$$

and so

$$(-1) - (-1) + \varphi(a, z, b) - 1 = 0.$$

Hence, $\varphi(a, z, b) = 1$ and so we can conclude that $(y, x)_{S^1} \subset (a, b)_{S^1}$. Thus, since $\{x, y\} \subseteq I_n$, $[y, x]_{S^1} \subset (a, b)_{S^1}$.

We have shown that $[y, x]_{S^1} \subset (a, b)_{S^1} \subsetneq (v, u)_{S^1}$. Then $a \in [v, y)_{S^1}$ and $b \in (x, u]_{S^1}$. By Proposition 10.10.4, $\{a, b\} = \{u, v\}$ or $a \in (v, y)_{S^1}$ and $b \in (x, u)_{S^1}$. If $\{a, b\} = \{u, v\}$, this is a contradiction since $(a, b)_{S^1} \subsetneq (v, u)_{S^1}$. Therefore, $a \in (v, y)_{S^1}$ and $b \in (x, u)_{S^1}$. However, by Proposition 10.10.7, this is a contradiction. Thus, I^* is isolated. ◻

Lemma 10.10.11 *Let $(\mathscr{L}_1, \mathscr{L}_2, G)$ be a pseudo-fibered triple and g a nontrivial element of G with $Fix_g \neq \emptyset$. Suppose that there is a non-leaf gap $\mathscr{G}$ of $\mathscr{L}_i$ for some $i \in \{1, 2\}$. If there is an isolated fixed point in $v(\mathscr{G})$, then $v(\mathscr{G}) \subseteq Fix_g$ and for all $I \in \mathscr{G}$, $|I \cap Fix_g| = 1$.*

Proof Without loss of generality, we can assume that $\mathscr{G}$ is a gap on $\mathscr{L}_1$. Denote the isolated fixed point by p. By Proposition 10.10.4 we can derive $v(\mathscr{G}) \subseteq Fix_g$. Since $\mathscr{G}$ is a non-leaf gap, $3 \leq |Fix_g|$. Then, for each $J \in \mathscr{G}$, $1 \leq |J \cap Fix_g|$. If not, there is an element J in $\mathscr{G}$ such that $J \cap Fix_g = \emptyset$. This implies that J is a connected component of $S^1 - Fix_g$. But this contradicts Proposition 10.10.6. So, we also conclude $6 \leq |Fix_g|$.

Let us set $p = x_1$ and $\mathscr{G} = \{(x_1, x_2)_{S^1}, (x_2, x_3)_{S^1}, \cdots, (x_{n-1}, x_n)_{S^1}, (x_n, x_1)_{S^1}\}$. We use $\mathbb{Z}_n$ as the index set. First, assume that x_i is an isolated fixed point for some $i \in \mathbb{Z}_n$. Since x_i is isolated, there is a connected component $(x_i, x_i')_{S^1}$. Since $(x_i, x_i')_{S^1} \subseteq (x_i, x_{i+1})_{S^1}$ and $1 \leq |(x_i, x_{i+1})_{S^1} \cap Fix_g|$, so $x_i \neq x_i'$ and $(x_i, x_i')_{S^1} \subsetneq (x_i, x_{i+1})_{S^1}$. By Proposition 10.10.8, x_i' is also an isolated fixed point and so there is a connected component (x_i', x_i'') of $S^1 - Fix_g$. Then there are two cases. One is $(x_i, x_i'')_{S^1} = (x_i, x_{i+1})_{S^1}$ and the other is $(x_i, x_i'')_{S^1} \subsetneq (x_i, x_{i+1})_{S^1}$. Assume that $(x_i, x_i'')_{S^1} \subsetneq (x_i, x_{i+1})_{S^1}$. By Proposition 10.10.6, $\{x_i, x_i''\} \subseteq E(\mathscr{L}_1)$

and $x_i' \in E(\mathscr{L}_2)$. Then by Theorem 10.7.14, there is a rainbow $\{I_n\}_{n=1}^{\infty}$ at x_i' on $\mathscr{L}_1$ since $E(\mathscr{L}_1) \cap E(\mathscr{L}_2) = \emptyset$. Since $\bigcap_{n=1}^{\infty} I_n = \{x_i'\}$, there is I_N in $\{I_n\}_{n=1}^{\infty}$ such that $\{x_i, x_i''\} \subseteq I_N^*$. Then $I_N \in C_{x_i'}^{(x_i, x_i'')_{S^1}}$ with $\overline{I_N} \subset (x_i, x_i'')_{S^1}$. Since $C_{x_i'}^{(x_i, x_i'')_{S^1}}$ is preserved by g and $v(\ell(I_N)) \subseteq S^1 - Fix_g$, $g(\overline{I_N}) \subset I_N$ or $\overline{I_N} \subset g(I_N)$. So, one of two sequences $\{\ell(g^k(I_N))\}_{k=1}^{\infty}$ and $\{\ell(g^{-k}(I_N))\}_{k=1}^{\infty}$ converges to $(x_i, x_i'')_{S^1}$ on $\mathscr{L}_1$. Therefore, $(x_i, x_i'') \in \mathscr{L}_1$. By Lemma 10.10.10, $(x_i, x_i'')_{S^1}^*$ is isolated. Therefore, by Lemma 10.7.12, there is a non-leaf gap $\mathscr{G}'$ of $\mathscr{L}_1$ which contains $(x_i, x_i'')_{S^1}$, However, by the definition of looseness, this is a contradiction since $\mathscr{G}' \neq \mathscr{G}$ and $x_i \in v(\mathscr{G}) \cap v(\mathscr{G}')$. Thus, $(x_i, x_i'')_{S^1} = (x_i, x_{i+1})_{S^1}$ is the possible case. So, $(x_i, x_{i+1})_{S^1}$ contains only one fixed point and by Proposition 10.10.8 x_{i+1} is an isolated fixed point. Therefore, since x_1 is an isolated fixed point, we are done.

□

Let us prove the main theorem.

Theorem 10.10.12 *Let $(\mathscr{L}_1, \mathscr{L}_2, G)$ be a pseudo-fibered triple and g a nontrivial element of G. Then $|Fix_g| < \infty$.*

Proof It is enough to show the case $5 \leq |Fix_g|$. Assume $5 \leq |Fix_g|$. Since g is nontrivial, $S^1 - Fix_g$ is nonempty and so there is a connected component I of $S^1 - Fix_g$ which is a nondegenerate open interval. By Proposition 10.10.8, for each $p \in \partial I$, p is an isolated fixed point. Choose $p \in \partial I$. By Lemma 10.10.9, there is an element J in $\mathscr{L}_1 \cup \mathscr{L}_2$ such that $p \in v(\ell(J))$ and $|J \cap Fix_g| = 1$. Without loss generality, say $J \in \mathscr{L}_1$. Then, by Lemma 10.10.10, J^* is isolated and so by Lemma 10.7.12, there is a non-leaf gap $\mathscr{G}$ such that $J \in \mathscr{G}$. Therefore, by Lemma 10.10.11, $v(\mathscr{G}) \subseteq Fix_g$ and for all $K \in \mathscr{G}$, $|K \cap Fix_g| = 1$. This implies $|Fix_g| < \infty$. □

10.11 Future Directions

We conclude the chapter by suggesting some future directions. As we saw in Corollary 10.9.11, if $(\mathscr{L}, G)$ is a tight pair, then G contains a nonabelian free subgroup as long as it does not admit a global fixed point. Indeed, one can show that a sticky pair has no global fixed point, hence the group of the sticky pair necessarily contains a nonabelian free subgroup (the proof will be contained in an upcoming paper of the authors). Although it seems more difficult to determine if a slippery pair has no global fixed point, we propose the following conjecture.

Conjecture 10.11.1 *Suppose $(\mathscr{L}, G)$ is a tight pair, then G admits no global fixed point (therefore, it contains a nonabelian free subgroup).*

Another direction is to study further properties of pseudo-fibered triples. In [1], the following conjecture based on observations in [2, 3] was proposed.

Conjecture 10.11.2 ([1]) *Let* $(G, \mathscr{L}_1, \mathscr{L}_2)$ *be a pseudo-fibered triple. Suppose G is finitely generated, torsion-free, and freely indecomposable. Then one of the following three possibilities holds:*

1. *G is virtually abelian.*
2. *G is topologically conjugate into* $\mathrm{PSL}_2(\mathbb{R})$.
3. *G is isomorphic to a closed hyperbolic 3-manifold group.*

By Theorem 10.5.2 (or its simplified version), the second possibility of Conjecture 10.11.2 holds if there exists a third invariant lamination which is compatible with other $\mathscr{L}_1, \mathscr{L}_2$. The following theorem is a combination of two main theorems of [1] on pseudo-fibered triples.

Theorem 10.11.3 ([1]) *Suppose G is a group as in Conjecture 10.11.2.*

1. *G satisfies a type of Tits alternative. Namely, each subgroup of G either contains a nonabelian free subgroup or is virtually abelian.*
2. *If G purely consists of hyperbolic elements, then G acts on* S^2 *as a convergence group.*

The 2-sphere that appears in the second part of Theorem 10.11.3 is obtained as a quotient of the circle on which the group G acts. The quotient map is the map collapsing laminations $\mathscr{L}_1$ and $\mathscr{L}_2$ which is analogous to the famous Cannon–Thurston map constructed in the seminal paper [13]. The study of the induced action on S^2 in our work was largely influenced by Fenley's work [16].

One strategy to achieve the third possibility of Conjecture 10.11.2 is to first strengthen the second part of Theorem 10.11.3. Namely, one may try to show that if G contains both hyperbolic and non-hyperbolic elements, then G acts on S^2 as a uniform convergence group. Then by a theorem of Bowditch [6], G is word-hyperbolic and S^2 is equivariantly homeomorphic to its boundary. Hence, if one can prove Cannon's conjecture in this setting, one ends up with the third possibility of Conjecture 10.11.2. Perhaps as an intermediate step, one may try the following conjecture.

Conjecture 10.11.4 *Suppose G is a group as in Conjecture 10.11.2, and assume G is not virtually abelian. Then G is word-hyperbolic.*

Acknowledgments We thank Michele Triestino, Steven Boyer, Sanghyun Kim, Thierry Barbot and Michel Boileau for fruitful conversations. We would like to give special thanks to Hongtaek Jung for a careful reading of an earlier draft and giving helpful comments regarding Sect. 10.9. Finally we greatly appreciate Athanase Papadopoulos and the anonymous referee for their valuable comments which improved the exposition of this chapter. The first half of the chapter was written based on the mini-course the first author gave in the workshop "Low dimensional actions of 3-manifold groups" at Université de Bourgogne in November, 2019. The first author was partially supported by Samsung Science & Technology Foundation grant No. SSTF-BA1702-01, and the second author was partially supported by the Mid-Career Researcher Program (2018R1A2B6004003) through the National Research Foundation funded by the government of Korea.

References

1. J. Alonso, H. Baik, E. Samperton, On laminar groups, tits alternatives and convergence group actions on S^2. J. Group Theory **22**(3), 359–381 (2019)
2. H. Baik, Laminations on the circle and hyperbolic geometry. Ph.D. Thesis. Cornell University (2014)
3. H. Baik, Fuchsian groups, circularly ordered groups, and dense invariant laminations on the circle. Geom. Topol. **19**(4), 2081–2115 (2015)
4. H. Baik, E. Samperton, Spaces of invariant circular orders of groups. Groups Geom. Dyn. **12**(2), 721–764 (2018)
5. T. Barbot, S. Fenley, Free seifert pieces of pseudo-Anosov flows (2015). Preprint. arXiv:1512.06341
6. B. Bowditch, A topological characterization of hyperbolic groups. J. Amer. Math. Soc. **11**, 643–667 (1998)
7. D. Calegari, The geometry of R-covered foliations. Geom. Topol. **4**, 457–515 (2000)
8. D. Calegari, *Foliations and the Geometrization of 3-Manifolds*. Notes from graduate course at Harvard (2001)
9. D. Calegari, Universal circles for quasigeodesic flows. Geom. Topol. **10**(4), 2271–2298 (2006)
10. D. Calegari, *Foliations and the Geometry of 3-Manifolds* (Oxford Science Publications, Oxford, 2007)
11. D. Calegari, N. Dunfield, Laminations and groups of homeomorphisms of the circle. Invent. Math. **152**, 149–207 (2003)
12. A. Candel, Uniformization of surface laminations, in *Annales scientifiques de l'Ecole normale supérieure*, vol. 26 (1993), pp. 489–516
13. J.W. Cannon, W.P. Thurston, Group invariant Peano curves. Geom. Topol. **11**, 1315–1355 (2007)
14. S.R. Fenley, Anosov flows in 3-manifolds. Ann. Math. **139**(1), 79–115 (1994)
15. S.R. Fenley, The structure of branching in Anosov flows of 3-manifolds. Comment. Math. Helv. **73**(2), 259–297 (1998)
16. S. Fenley, Ideal boundaries of pseudo-Anosov flows and uniform convergence groups with connections and applications to large scale geometry. Geom. Topol. **16**(1), 1–110 (2012)
17. S.R. Fenley, Quasigeodesic pseudo-Anosov flows in hyperbolic 3-manifolds and connections with large scale geometry. Adv. Math. **303**, 192–278 (2016)
18. S. Frankel, From quasigeodesic to pseudo-Anosov: laminar decompositions and orbit semi-stability. Preprint
19. S. Frankel, Coarse hyperbolicity and closed orbits for quasigeodesic flows. Ann. Math. **188**(1), 1–48 (2018)
20. D. Gabai, W.H. Kazez, Order trees and laminations of the plane. Math. Res. Lett. **4**(4), 603–616 (1997)
21. D. Gabai, W.H. Kazez, Group negative curvature for 3-manifolds with genuine laminations. Geom. Topol. **2**(1), 65–77 (1998)
22. É. Ghys, Flots d'anosov sur les 3-variétés fibrées en cercles. Ergodic Theory Dynam. Systems **4**(1), 67–80 (1984)
23. R.S. Hamilton, The formation of singularities in the Ricci flow. *Surveys in Differential Geometry (Cambridge, MA, 1993)*, vol. ii (International Press, Cambridge, 1995), pp. 7–136
24. M. Hamm, Filling essential laminations. Ph.D. Thesis. Washington University in St. Louis (2009)
25. G. Margulis, Free subgroups of the homeomorphism group of the circle. C. R. Acad. Sci. Ser. I Math. **331**(9), 669–674 (2000)
26. L. Mosher, Laminations and flows transverse to finite depth foliations, Part I: Branched surfaces and dynamics. Preprint (1996)
27. S.P. Novikov, Topology of foliations. Trans. Moscow Math. Soc. **14**, 248–278 (1965)

28. G. Perelman, The entropy formula for the Ricci flow and its geometric applications (2002). arXiv Preprint math/0211159
29. G. Perelman, Ricci flow with surgery on three-manifolds (2003). arXiv Preprint math/0303109
30. H. Rosenberg, Foliations by planes. Topology **7**(2), 131–138 (1968)
31. W.P. Thurston, Three-manifolds, Foliations and Circles, I. arXiv:math/9712268v1 [math.GT] (1997)
32. E.T. Winkel, Sticky and slippery laminations on the circle. Master's Thesis, University of Bonn (2016)

Chapter 11
Length Functions on Currents and Applications to Dynamics and Counting

Viveka Erlandsson and Caglar Uyanik

Abstract The aim of this chapter is twofold. We first explore a variety of length functions on the space of currents, and we survey recent work regarding applications of length functions to counting problems. Secondly, we use length functions to provide a proof of a folklore theorem which states that pseudo-Anosov homeomorphisms of closed hyperbolic surfaces act on the space of projective geodesic currents with uniform North-South dynamics.

Keywords Teichmüller spaces · Geodesic laminations · Currents · Mapping class group

2010 Mathematics Subject Classification 37E30 (primary); 30F60, 57M50 (secondary)

11.1 Introduction

Geodesic currents are measure theoretic generalizations of closed curves on hyperbolic surfaces and they play an important role, among many other things, in the study of the geometry of Teichmüller space (see, for example, [5, 6]). The set of all closed curves sits naturally as a subset of the space of currents, and various fundamental notions such as geometric intersection number and length of curves extend to this more general setting of currents.

The aim of this (mostly expository) chapter is twofold. We first explore a variety of length functions on the space of currents, and we survey recent work regarding applications of length functions to counting problems. Secondly, we use length

V. Erlandsson (✉)
School of Mathematics, University of Bristol, Bristol, UK
e-mail: v.erlandsson@bristol.ac.uk

C. Uyanik
Department of Mathematics, Yale University, New Haven, CT, USA
e-mail: caglar.uyanik@yale.edu

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_11

functions to provide a proof of a folklore theorem which states that pseudo-Anosov homeomorphisms of closed hyperbolic surfaces act on the space of projective geodesic currents with uniform North-South dynamics, see Theorem 11.5.1.

More precisely, let S be a closed, orientable, finite type surface of genus $g \geq 2$ and denote the space of geodesic currents on S by $\text{Curr}(S)$. By a *length function* on $\text{Curr}(S)$ we mean a function that is homogeneous and positive (see Definition 11.2.2). There are many ways to define a length of a closed curve on S: a Riemannian metric on S naturally induces a notion of length, a generating set of $\pi_1(S)$ gives the notion of word length of a curve, and given a fixed (filling) curve γ one can consider a combinatorial length given by the curves' intersection number with γ. We will see that all these notions of length give rise to *continuous* length functions on $\text{Curr}(S)$.

As a first example, in Sect. 11.2.6, we will see that Bonahon's intersection form, which is an extension of the geometric intersection number of curves to currents, induces a continuous length function on $\text{Curr}(S)$. We also use this intersection form to prove the uniform North-South dynamics result mentioned above (see Sect. 11.5).

In Sect. 11.3 we explore other notions of length of curves that have continuous extensions to length functions on $\text{Curr}(S)$. In Sect. 11.3.1 we use *Liouville currents* to extend the length of curves given by any (possibly singular) non-positively curved Riemannian metric on S to a continuous length function on $\text{Curr}(S)$, as well as the word length with respect to so-called simple generating sets of $\pi_1(S)$ (see Theorem 11.3.3 and Corollary 11.3.5). Next, we explore length functions with respect to any Riemannian metric on S (respectively the word length with respect to an arbitrary generating set) and explain why the corresponding *stable lengths* extend to continuous length functions on $\text{Curr}(S)$, see Theorem 11.3.6.

In Sect. 11.4, we apply the results of Sect. 11.3 to problems regarding counting curves on surfaces. Mirzakhani [24, 25] proved that the asymptotic growth rate of the number of curves of bounded hyperbolic length, in each mapping class group orbit, is polynomial in the length (see Theorem 11.4.1 for the precise statement). We explain how to use continuous length functions on $\text{Curr}(S)$ to generalize her result to other notions of length of curves, and show that the same asymptotic behavior holds for all lengths discussed above (see Theorem 11.4.3 and Corollary 11.4.9). These results appeared in [11–13] and here we attempt to give a clear outline of the logic behind these proofs.

11.2 Background

11.2.1 Curves on Surfaces

Throughout this chapter, we let S be a closed, orientable, finite type surface of genus $g \geq 2$. By a *curve* γ on S we mean a (free) homotopy class of an immersed, essential, closed curve. That is, the homotopy class of the image of an immersion of the unit circle $S^1 \to S$, where the image is not homotopic to a point. We say the curve is *simple* if the immersion is homotopic to an embedding. We identify a curve with its corresponding conjugacy class, denoted $[\gamma]$, in the fundamental group $\pi_1(S)$. Furthermore, we assume curves to be primitive, that is $\gamma \neq \eta^k$ for any $k > 1$ and $\eta \in \pi_1(S)$. By a *multicurve* we mean a union of finitely many weighted curves, that is

$$\bigcup_{i=1}^{n} a_i \gamma_i$$

where $a_i > 0$ and γ_i is a curve for each i. We say the multicurve is *integral* if $a_i \in \mathbb{Z}$ for all i, and that it is *simple* if the curves γ_i are simple and pairwise disjoint.

11.2.2 Teichmüller Space and the Mapping Class Group

A *hyperbolic structure* on a surface S is a collection of charts $\{(U_i, \psi_i)\}$ such that

(1) $\{U_i\}$ is an open cover of S,
(2) the map $\psi_i : U_i \to \mathbb{H}^2$ is an orientation preserving homeomorphism onto its image for each i,
(3) For each i, j such that $U_i \cap U_j \neq \emptyset$ the restriction of $\psi_j \circ \psi_i^{-1}$ to each component of $U_i \cap U_j$ is an element of $\mathrm{Isom}^+(\mathbb{H}^2)$.

The surface S together with a hyperbolic structure is called a *hyperbolic surface*. The Cartan–Hadamard theorem asserts that a closed hyperbolic surface is isometrically diffeomorphic to $\mathbb{H}^2/\Gamma$ where Γ is a torsion-free discrete subgroup of $\mathrm{Isom}^+(\mathbb{H}^2)$.

A *marked hyperbolic surface* is a pair (X, f) where

(1) $X = \mathbb{H}^2/\Gamma$ is a hyperbolic surface, and
(2) $f : S \to X$ is an orientation-preserving homeomorphism.

Given a marked hyperbolic surface (X, f), we can pull back the hyperbolic structure on X by f to one on S. Conversely, given a hyperbolic structure on S, the identity map $id : S \to S$ makes (S, id) into a marked hyperbolic surface.

The *Teichmüller space* of S is the set $\mathrm{Teich}(S) = \{(X, f)\}/\sim$ of equivalence classes of marked hyperbolic surfaces, where two hyperbolic surfaces (X, f) and (Y, g) are *equivalent* if $g \circ f^{-1}$ is homotopic to an isometry from X to Y.

The *mapping class group* $\mathrm{Mod}(S)$ of S is the group of isotopy classes of orientation-preserving homeomorphisms of S; in other words,

$$\mathrm{Mod}(S) = \mathrm{Homeo}^+(S)/\mathrm{Homeo}_0(S)$$

where $\mathrm{Homeo}_0(S)$ is the connected component of the identity in the orientation-preserving homeomorphism group $\mathrm{Homeo}^+(S)$.

The mapping class group $\mathrm{Mod}(S)$ acts on $\mathrm{Teich}(S)$ naturally by precomposing the marking map, i.e. for $\varphi \in \mathrm{Mod}(S)$, and $[(X, f)] \in \mathrm{Teich}(S)$ choose a lift $\Phi \in \mathrm{Homeo}^+(S)$ of φ and define

$$\varphi[(X, f)] = [(X, f \circ \Phi^{-1})].$$

11.2.3 Measured Laminations

A *geodesic lamination* on a hyperbolic surface S is a closed subset $\mathcal{L}$ of S that is a union of simple, pairwise disjoint, complete geodesics on S. The geodesics in $\mathcal{L}$ are called the *leaves* of the lamination. A *transverse measure* λ on $\mathcal{L}$ is an assignment of a locally finite Borel (Radon) measure $\lambda_{|k}$ on each arc k transverse to $\mathcal{L}$ so that

(1) If k' is a subarc of an arc k, then $\lambda_{|k'}$ is the restriction to k' of $\lambda_{|k}$;
(2) Transverse arcs which are transversely isotopic have the same measure.

A *measured lamination* is a pair $(\mathcal{L}, \lambda)$ where $\mathcal{L}$ is a geodesic lamination and λ is a transverse measure. In what follows, we will suppress $\mathcal{L}$ and write λ for brevity. The set of measured laminations on S is denoted by $\mathrm{ML}(S)$, and endowed with the weak-* topology: a sequence $\lambda_n \in \mathrm{ML}(S)$ converges to $\lambda \in \mathrm{ML}(S)$ if and only if

$$\int_k f d\lambda_n \longrightarrow \int_k f d\lambda$$

for any compactly supported continuous function f defined on a generic transverse arc k on S.

An easy example of a measured geodesic lamination is given by a simple curve γ on S, together with the transverse measure λ_γ: for each transverse arc k the transverse measure is the Dirac measure $\lambda_{\gamma|k}$ which counts the number of intersections with γ, i.e. for any Borel subset B of k, we have $\lambda_{\gamma|k}(B) = |B \cap \gamma|$. The set of measured geodesic laminations coming from *weighted simple curves* λ_γ is dense in $\mathrm{ML}(S)$, see [27]. We denote the subset of $\mathrm{ML}(S)$ coming from simple integral multicurves by $\mathrm{ML}_{\mathbb{Z}}(S)$.

Note that $\mathbb{R}^+$ acts naturally on $\mathrm{ML}(S)$ by scaling the transverse measure. The space of *projective measured laminations* is defined as the quotient

$$\mathbb{P}\,\mathrm{ML}(S) = \mathrm{ML}(S)/\mathbb{R}_+$$

and the equivalence class of a measured geodesic lamination in $\mathbb{P}\,\mathrm{ML}(S)$ is denoted by $[\lambda]$. We endow $\mathbb{P}\,\mathrm{ML}(S)$ with the quotient topology.

The space $\mathrm{ML}(S)$ of measured laminations is homeomorphic to $\mathbb{R}^{6g-6}$ and has a $\mathrm{Mod}(S)$-invariant piecewise linear manifold structure (see, for example, [27]). This piecewise linear structure is given by train track coordinates. We refer the reader to [29] and [27] for a detailed discussion of train tracks and only recall the relevant notions for our purposes.

Let τ be a smoothly embedded 1-complex in S, i.e. an embedded complex whose edges are smoothly embedded arcs with well-defined tangent lines at the end-points. A *complementary region* of τ is the metric completion of a connected component of $S \setminus \tau$. We say τ is a *train track* on S if in addition it satisfies the following properties:

(1) at each vertex, called *switch*, the tangent lines to all adjacent edges agree
(2) at each vertex the set of adjacent edges can be divided into two sets according to the direction of the tangent line; we require each of these sets to be non-empty at every vertex
(3) doubling each complementary region gives a surface with singular points having negative Euler characteristic $\chi = 2 - 2g - p$ (where g and p represent the genus and the number of singular points, respectively).

A train track is called *maximal* if the complementary regions to τ are all triangles.

A simple closed curve, or more generally a measured lamination $(\mathcal{L}, \lambda)$, is *carried* by τ if there is a smooth map $g : S \to S$ such that

(1) $g : S \to S$ is isotopic to the identity,
(2) the restriction of g to $\mathcal{L}$ is an immersion,
(3) $g(\mathcal{L}) \subset \tau$.

There is a finite collection of train tracks $\mathcal{T} = \{\tau_1, \tau_2, \ldots, \tau_n\}$ such that for all $\lambda \in \mathrm{ML}(S)$ there is a $\tau_i \in \mathcal{T}$ which λ is carried by. The set of measured laminations carried by a train track is full dimensional if and only if the train track is maximal and recurrent. Now, each maximal train track τ determines a cone $C(\tau)$ in $\mathbb{R}^{6g-6}$, given by the solutions to the so called *switch equations*, and we have a homeomorphism between all laminations carried by τ and $C(\tau)$. Moreover, the set of integer points in $C(\tau)$ is in one to one correspondence with the simple integral multicurves (i.e. elements of $\mathrm{ML}_{\mathbb{Z}}(S)$) carried by τ.

11.2.4 Geodesic Currents

Consider a hyperbolic metric ρ on S and let $\tilde{S}$ be the universal cover equipped with the pullback metric $\tilde{\rho}$. Let $\mathcal{G}(\tilde{\rho})$ denote the set of complete geodesics in $\tilde{S}$. Let S^1_∞ denote the boundary at infinity of $\tilde{S}$. Note that since $\tilde{\rho}$ is hyperbolic and complete, $\tilde{S}$ is isometric to $\mathbb{H}^2$ and its boundary is homeomorphic to the unit circle S^1. Each geodesic is uniquely determined by its pair of endpoints on S^1_∞. Hence we can identify the set of geodesics with the *double boundary*

$$\mathcal{G}(\tilde{S}) = \left(S^1_\infty \times S^1_\infty\right) \setminus \Delta)/(x, y) \sim (y, x)$$

where Δ denotes the diagonal. That is, $\mathcal{G}(\tilde{S})$ consists of unordered pairs of distinct boundary points, and we refer to it as the *space of geodesics of* $\tilde{S}$. Note that $\mathcal{G}(\tilde{S})$ is independent of the metric ρ. Indeed, if ρ' is another geodesic metric on S, then the universal cover $\tilde{S}$ equipped with the pullback metric $\tilde{\rho}'$ is quasi-isometric to $\mathbb{H}^2$ and this quasi-isometry extends to a homeomorphism of the boundaries at infinity (see [1] for the details). Hence $\mathcal{G}(\tilde{S})$ is well-defined without a reference to a metric.

The fundamental group $\pi_1(S)$ acts naturally on $\tilde{S}$ by deck transformations, and this action extends continuously to S^1_∞ and $\mathcal{G}(\tilde{S})$. For any (geodesic) metric ρ the map

$$\partial_\rho : \mathcal{G}(\tilde{\rho}) \to \mathcal{G}(\tilde{S})$$

that maps each geodesic to its pair of endpoints is continuous, surjective and $\pi_1(S)$-invariant, and a homeomorphism when ρ is negatively curved.

A *geodesic current* on S is a Radon measure on $\mathcal{G}(\tilde{S})$ which is invariant under the action of $\pi_1(S)$. We denote the set of all geodesic currents on S by $\mathrm{Curr}(S)$ and endow it with the weak-* topology: A sequence $\mu_n \in \mathrm{Curr}(S)$ of currents converges to $\mu \in \mathrm{Curr}(S)$ if and only if

$$\int f d\mu_n \longrightarrow \int f d\mu$$

for all continuous, compactly supported functions $f : \mathcal{G}(\tilde{S}) \to \mathbb{R}$.

As a first example of a geodesic current, consider the preimage under the covering map in $\tilde{S}$ of any closed curve γ on S, which is a collection of complete geodesics in $\tilde{S}$. This defines a discrete subset of $\mathcal{G}(\tilde{S})$ which is invariant under the action of $\pi_1(S)$. The Dirac (counting) measure associated with this set on $\mathcal{G}(\tilde{S})$ gives a geodesic current on S.

The map from the set of curves on S to $\mathrm{Curr}(S)$ that sends each curve to its corresponding geodesic current, as above, is injective. Hence, we view the set of curves on S as a subset of $\mathrm{Curr}(S)$. In fact, Bonahon showed that the set of all *weighted curves* is dense in $\mathrm{Curr}(S)$ [5]. We identify a curve γ with the current it defines, and by abuse of notation we denote both by γ.

Another important subset of geodesic currents is given by measured laminations. Let $(\mathcal{L}, \lambda)$ be a measured lamination and consider its preimage $\tilde{\mathcal{L}}$ in $\tilde{S}$ which is a collection of pairwise disjoint complete geodesics. The lift $\tilde{\mathcal{L}}$ is a discrete subset of $\mathcal{G}(\tilde{S})$ which is $\pi_1(S)$-invariant. Hence the associated Dirac measure on $\tilde{\mathcal{L}}$ defines a geodesic current on S. Moreover this measure agrees with the transverse measure λ, see [1] for the details. Hence we view $\mathrm{ML}(S)$ as a subset of $\mathrm{Curr}(S)$ as well.

A current $\nu \in \mathrm{Curr}(S)$ is called *filling* if every complete geodesic in $\tilde{S}$ transversely intersects a geodesic in the support of ν in $\mathcal{G}(\tilde{S})$. Note that this definition agrees with the classical notion of *filling curves*: a curve γ defines a filling current if and only if γ is filling as a curve, i.e. $S \setminus \gamma$ is a union of topological disks.

11.2.5 Nielsen–Thurston Classification

Thurston defined a Mod(S)-equivariant compactification of the Teichmüller space by the space of projective measured laminations and using the action of Mod(S) on $\overline{\text{Teich}(S)} = \text{Teich}(S) \cup \mathbb{P}\,\text{ML}(S)$ showed:

Theorem 11.2.1 (Nielsen-Thurston Classification) *[14, 30] Each $\varphi \in \text{Mod}(S)$ is either periodic, reducible or pseudo-Anosov. Furthermore, pseudo-Anosov mapping classes are neither periodic nor reducible.*

Here $\varphi \in \text{Mod}(S)$ is called *periodic* if there exist a $k > 0$ such that φ^k is the identity. The map φ is called *reducible* if there is a collection C of disjoint simple curves on S and a representative φ' of φ such that $\varphi'(C)$ is isotopic to C. Finally, $\varphi \in \text{Mod}(S)$ is called *pseudo-Anosov* if there exists a filling pair of transverse, measured laminations $(\mathcal{L}^+, \lambda_+)$ and $(\mathcal{L}^-, \lambda_-)$, a number $\alpha > 1$ called the *stretch factor*, and a representative homeomorphism φ' of φ such that

$$\varphi'(\mathcal{L}^+, \lambda_+) = (\mathcal{L}^+, \alpha\lambda_+)$$

and

$$\varphi'(\mathcal{L}^-, \lambda_-) = (\mathcal{L}^-, \frac{1}{\alpha}\lambda_-).$$

The measured laminations $(\mathcal{L}^+, \lambda_+)$ and $(\mathcal{L}^-, \lambda_-)$ are called the *unstable lamination* and the *stable lamination* respectively. We will suppress the $\mathcal{L}$ and write λ_+ and λ_- respectively.

11.2.6 Length Functions and the Intersection Number

Definition 11.2.2 A *length function* on the space of geodesic currents is a map $\ell : \text{Curr}(S) \to \mathbb{R}$ which is homogeneous and positive, i.e.

$$\ell(a\mu) = a\ell(\mu)$$

for any $a > 0$ and $\mu \in \text{Curr}(S)$, $\ell(\mu) \geq 0$ for all $\mu \in \text{Curr}(S)$ and $\ell(\mu) = 0$ if and only if $\mu = 0$.

We say that a map ℓ on the set of curves on S is a length function if it is a positive function, i.e. $\ell(\gamma) > 0$ for all curves γ on S. Note that this agrees with the definition above, when viewing the set of curves as a subset of the space of geodesic currents.

Given two curves γ, η on S, their *(geometric) intersection number*, denoted $i(\gamma, \eta)$, is defined as the minimum number of transverse intersections between transverse representatives of the homotopy classes of γ and η. That is

$$i(\gamma, \eta) = \min \left\{ |\gamma' \pitchfork \eta'| \mid \gamma' \sim \gamma, \eta' \sim \eta \right\}$$

where $\sim$ denotes homotopic.

We note that $i(\gamma, \gamma) = 0$ if and only if γ is a simple curve. Moreover, an equivalent description of the intersection number of two distinct curves γ and η is the following. Let ρ be a hyperbolic metric on S and $\tilde{S}$ be the universal cover equipped with the pullback metric $\tilde{\rho}$. Let $\tilde{\gamma}$ be a geodesic representative of a lift of γ to $\tilde{S}$. Consider the set of lifts of η and take their geodesic representatives. Let x be a point on $\tilde{\gamma}$ that does not lie on a geodesic representative of any lift of η, and consider the bounded segment δ_γ of $\tilde{\gamma}$ between x and $\gamma(x)$. Then the intersection number $i(\gamma, \eta)$ is exactly the same as the number of the lifts of η that intersect (necessarily transversely) δ_γ. This description of the intersection number will be helpful below.

Viewing the set of curves as a subspace of the space of geodesic currents, it is natural to ask if the intersection number extends, in a nice way, to Curr(S). Indeed, Bonahon [6] showed that there is a unique continuous extension of the intersection number to the space of geodesic currents:

Theorem 11.2.3 ([6, Proposition 4.5]) *There is a unique continuous, symmetric, bilinear form*

$$i(\cdot, \cdot) : \mathrm{Curr}(S) \times \mathrm{Curr}(S) \to \mathbb{R}_{\geq 0}$$

such that $i(\gamma, \eta)$ *agrees with the geometric intersection number whenever* γ, η *are curves on* S.

Here we give the definition of this intersection form and explain how it induces length functions on Curr(S). For the definition we follow the exposition presented in [1] and refer to that paper for more details. Let $\mathcal{G}^{\pitchfork}(\tilde{S}) \subset \mathcal{G}(\tilde{S}) \times \mathcal{G}(\tilde{S})$ be the subset defined by

$$\mathcal{G}^{\pitchfork}(\tilde{S}) = \left\{ (\{x, y\}, \{z, w\}) \in \mathcal{G}(\tilde{S}) \times \mathcal{G}(\tilde{S}) \setminus \Delta \mid \{x, y\}, \{c, d\} \text{ link} \right\}$$

where Δ represents the diagonal and we say that $\{x, y\}$ and $\{z, w\}$ *link* if x and y belong to different components of $S^1_\infty \setminus \{z, w\}$. Equivalently, $\mathcal{G}^{\pitchfork}(\tilde{S})$ consists of pairs of geodesics in $\tilde{S}$ that intersect transversely. The action of $\pi_1(S)$ on $\tilde{S}$ induces a free and properly discontinuous action on $\mathcal{G}^{\pitchfork}(\tilde{S})$ and hence the quotient map

$$\mathcal{G}^{\pitchfork}(\tilde{S}) \to \mathcal{G}^{\pitchfork}(\tilde{S})/\pi_1(S)$$

is a covering map. We define

$$\mathcal{G}^{\pitchfork}(S) = \mathcal{G}^{\pitchfork}(\tilde{S})/\pi_1(S).$$

Now, let $\mu, \nu \in \mathrm{Curr}(S)$. Then $\mu \times \nu$ is a product measure on $\mathcal{G}(\tilde{S}) \times \mathcal{G}(\tilde{S})$ and hence on $\mathcal{G}^{\pitchfork}(\tilde{S})$. This descends to a measure on $\mathcal{G}^{\pitchfork}(S)$ by locally pushing forward $\mu \times \nu$ through the covering map, and the intersection of μ and ν is defined as the $\mu \times \nu$-mass on $\mathcal{G}(S)$, that is,

$$i(\mu, \nu) = \int_{\mathcal{G}^{\pitchfork}(S)} d\mu \times d\nu.$$

Let γ be a curve and identify it with the current it defines. Let $\mu \in \mathrm{Curr}(S)$. Then $i(\gamma, \mu)$ can be defined as follows. As above, choose a hyperbolic metric ρ on S and consider the universal cover $\tilde{S}$ equipped with the pullback metric $\tilde{\rho}$. Take a lift of γ and let $\tilde{\gamma}$ be its geodesic representative. Let x be a point on $\tilde{\gamma}$ and consider the geodesic segment η_γ from x to $\gamma(x)$. Let $\mathcal{G}^{\pitchfork}(\eta_\gamma)$ denote the set of geodesics that transversely intersect η_γ and $\partial_\rho \mathcal{G}^{\pitchfork}(\eta_\gamma)$ the subset of $\mathcal{G}(\tilde{S})$ obtained by identifying each geodesic in $\mathcal{G}^{\pitchfork}(\eta_\gamma)$ with its pair of endpoints on S^1_∞. Then

$$i(\gamma, \mu) = \mu\left(\partial_\rho \mathcal{G}^{\pitchfork}(\eta_\gamma)\right),$$

see Fig. 11.1. In particular, we see that when μ is also (the current associated with) a curve on S, then the intersection form agrees with the geometric intersection number of curves on S.

We record some useful facts about the intersection form:

(1) If $\nu \in \mathrm{Curr}(S)$ is filling, then $i(\nu, \mu) \neq 0$ for all $\mu \in \mathrm{Curr}(S) \setminus \{0\}$.
(2) The intersection form is invariant under $\mathrm{Mod}(S)$. That is, if $g \in \mathrm{Mod}(S)$ then $i(\mu, \nu) = i(g(\mu), g(\nu))$ for any $\mu, \nu \in \mathrm{Curr}(S)$.
(3) $i(\mu, \mu) = 0$ if and only if $\mu \in \mathrm{ML}(S)$.

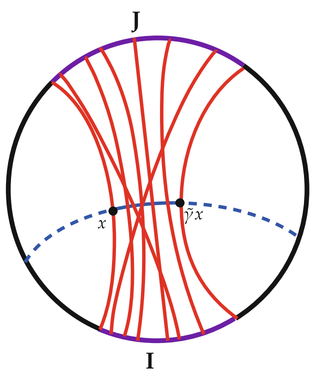


Fig. 11.1 Intersection of a curve γ with a current μ: $i(\gamma, \mu) = \mu(I \times J)$. Here η_γ is represented by the solid blue segment and $\partial_\rho \mathcal{G}^{\pitchfork}(\eta_\gamma) = I \times J$

(4) If $\nu \in \mathrm{Curr}(S)$ is filling and $L > 0$, then the set

$$\{\mu \in \mathrm{Curr}(S) \mid i(\mu, \nu) \leq L\}$$

is a compact set.

The first two statements follow from the definition of the intersection form, while the last two are results by Bonahon, see [5].

We define the space of *projective geodesic currents* to be

$$\mathbb{P}\mathrm{Curr}(S) = (\mathrm{Curr}(S) \setminus \{0\}) / \mathbb{R}_+.$$

It follows from (4) above that $\mathbb{P}\mathrm{Curr}(S)$ is a compact space.

Next we show how to obtain continuous length functions on $\mathrm{Curr}(S)$ from the intersection form. Fix a filling current $\nu \in \mathrm{Curr}(S)$. Define

$$\ell_\nu(\mu) = i(\nu, \mu)$$

for all $\mu \in \mathrm{Curr}(S)$. By the linearity and continuity of the intersection form, ℓ_ν is continuous and homogenous on $\mathrm{Curr}(S)$. Furthermore, since ν is filling, it follows from (1) above that ℓ_ν is positive. Hence the function above defines a continuous length function

$$\ell_\nu : \mathrm{Curr}(S) \to \mathbb{R}.$$

Moreover, this is the unique continuous extension of the length function on the set of curves defined by

$$\ell_\nu(\gamma) = i(\nu, \gamma)$$

for all curves γ on S. In Sect. 11.3 we will see that many other notions of lengths of curves have unique continuous extensions to length functions on $\mathrm{Curr}(S)$.

We end this section by noting that the intersection form can also be defined for geodesic currents on surfaces with boundary, and we refer the reader to [10] for the definitions. For simplicity of the exposition we assume throughout that S is a closed surface although the results presented here have generalizations that also hold for the case of compact surfaces.

11.3 Length Functions on Space of Currents

In Sect. 11.2.6 we saw that the geometric intersection number on the set of curves extend continuously to a bilinear form on pairs of currents and hence, fixing a filling curve (or current) ν, the length function

$$\ell_\nu(\gamma) = i(\nu, \gamma)$$

on the set of curves, extends *continuously* to a length function on currents, defined by

$$\ell_\nu(\mu) = i(\nu, \mu)$$

for all $\mu \in \mathrm{Curr}(S)$. There are many ways to define the length of a curve on a surface and it is natural to ask which other notions of length extends continuously to the space of currents. More concretely, let ρ be a (possibly singular) Riemannian metric on S. Then ρ naturally induces a length function $\ell_\rho(\cdot)$ on the set of curves where the length of a curve γ is defined to be the ρ-length of a shortest representative of γ. In the case when ρ is a negatively curved metric, this is the length of the unique geodesic representative in the homotopy class of γ. Another natural length function on the set of curves is given by first identifying a curve on S with a conjugacy class in the fundamental group $\pi_1(S)$ and, for a fixed a generating set of $\pi_1(S)$, defining the length of a curve to be the minimal number of generators needed to represent the corresponding conjugacy class. In general, given a geodesic metric space (X, d) on which $\pi_1(S)$ acts discretely and cocompactly by isometries, one can ask whether the translation length of a curve γ

$$\ell_X(\gamma) = \inf_{x \in X} d(x, \gamma(x)) \tag{11.1}$$

extends continuously to a length function on the space of currents. Note that when X is the universal cover of S equipped with a Riemannian metric, or X is the Cayley graph with respect to a generating set of $\pi_1(S)$, this length agrees with the notions described above.

We will see that in many cases such a continuous extension exists. In particular, in Sect. 11.3.1 below, we explain why it exists for any (possibly singular) non-positively curved Riemannian metric on S, through the use of Liouville currents and their relation to the intersection form on $\mathrm{Curr}(S)$. Similar arguments show that the word length with respect to certain (well-chosen) generating sets extends continuously to a length function on the space of currents.

Alas, such a continuous extension does not always exist. However, as we will see in Sect. 11.3.2, for any length function ℓ_X on curves as above, the *stable length* function defined by

$$sl_X(\gamma) = \lim_{n \to \infty} \frac{1}{n} \ell_X(\gamma^n)$$

always extends continuously to a length function on $\mathrm{Curr}(S)$.

Finally, in Sect. 11.3.3, we will see that the two approaches of defining an extension (using intersection with a special current, and considering the stable length) are intimately related.

11.3.1 Length of Currents Through Liouville Currents

In this section we explain how the length function of curves with respect to any (possibly singular) non-positively curved Riemannian metric on S can be extended continuously to a length function on the space of currents on S.

First, fix a hyperbolic metric ρ on S. The hyperbolic length of a homotopy class of a closed curve γ is defined as the length of the ρ-geodesic representative, and denoted by $\ell_\rho(\gamma)$. There exists a current associated with ρ, called its *Liouville current* and denoted by L_ρ, whose intersection form with curves on S determines the length function induced by ρ, that is:

$$i(\gamma, L_\rho) = \ell_\rho(\gamma) \tag{11.2}$$

for all curves γ on S.

Here we describe two equivalent definitions of the Liouville current and we refer the reader to [1, 6, 17, 26] for more details.

First we define the *Liouville measure* L on the hyperbolic plane $\mathbb{H}^2$. Let $\mathcal{G}(\mathbb{H}^2)$ denote the set of all bi-infinite geodesics in $\mathbb{H}^2$, which we identify with their endpoints on the unit circle S^1. Let $[a, b]$ and $[c, d]$ be two non-empty disjoint intervals on S^1. Define

$$L([a, b] \times [c, d]) = \left|\log\left|\frac{(a-c)(b-d)}{(a-d)(b-c)}\right|\right| \tag{11.3}$$

whenever a, b, c, d are four distinct points, and set $L([a, b] \times [c, d]) = 0$ if one of the intervals is a singleton. The map L extends uniquely to a Radon measure on $\mathcal{G}(\mathbb{H}^2)$ (see [6]) and is invariant under the action of $\pi_1(S)$ since the right-hand side in (11.3) is invariant under this action. In the disk model of $\mathbb{H}^2$ we have, using local coordinates $(e^{i\alpha}, e^{i\beta})$,

$$L_\rho = \frac{d\alpha d\beta}{|e^{i\alpha} - e^{i\beta}|^2}$$

where $d\alpha d\beta$ is the Lebesgue measure defined by the Euclidean metric on S^1, and in particular L is absolutely continuous with respect to the Lebesgue measure (see, for example, [1]). Now, given a hyperbolic metric ρ on S the universal cover $\tilde{S}$ with the pull-back metric $\tilde{\rho}$ is isometric to $\mathbb{H}^2$ and the boundary S^1_∞ is homeomorphic to S^1. We define L_ρ, the *Liouville current* with respect to ρ, to be the pull-back of L through this homeomorphism.

Alternatively, one can define the Liouville current in the following way. Let $\tilde{\eta}$ be a $\tilde{\rho}$-geodesic arc in $\tilde{S}$, parametrized at unit-speed by $\tilde{\eta} : (-a, a) \to \tilde{\eta}(t)$. Let $\mathcal{G}(\tilde{\eta})$ denote the set of all $\tilde{\rho}$-geodesics in $\tilde{S}$ that intersect $\tilde{\eta}$ transversely. Note that each geodesic in $\mathcal{G}(\tilde{\eta})$ is uniquely determined by its point of intersection $\tilde{\eta}(t)$ with $\tilde{\eta}$ and its angle of intersection (chosen in an arbitrary but consistent way). This gives rise to a homeomorphism

$$h_\eta : (-a, a) \times (0, \pi) \to \mathcal{G}(\tilde{\eta}).$$

Consider the measure on $(-a, a) \times (0, \pi)$ defined by

$$ds = \frac{1}{2} \sin(\theta) d\theta dt.$$

We push forward this measure through h_η to obtain a measure on $\mathcal{G}(\tilde{\eta})$. Lastly, we further push the measure forward through the homeomorphism

$$\partial_\rho : \mathcal{G}(\tilde{\rho}) \to \mathcal{G}(\tilde{S})$$

which maps each geodesic in $\mathcal{G}(\tilde{\eta})$ to its endpoints. The resulting measure is a Radon measure on $\mathcal{G}(\tilde{S})$. Furthermore, since $\pi_1(S)$ acts by isometries on $\tilde{S}$, the measure is invariant under its action. This measure is the Liouville measure L_ρ and agrees with the previous definition.

While the closed formula in the first definition makes it easier to state, the construction involved in the latter makes (11.2) more natural to see. Indeed, integrating ds over $\mathcal{G}(\tilde{\eta})$ for a unit-speed parametrized geodesic arc $\tilde{\eta}$ gives exactly the length of $\tilde{\eta}$.

The existence of Liouville currents for hyperbolic metrics allows us to embed the Teichmüller space of S into the space of geodesics currents, as shown by Bonahon [6]. More precisely, let (X, f) be a point in the Teichmüller space, and ℓ_X and L_X be the corresponding length function on curves and the Liouville current, respectively. Then, we have:

Theorem 11.3.1 ([6]) *The map*

$$(X, f) \mapsto L_X$$

defines an embedding $\mathrm{Teich}(S) \hookrightarrow \mathrm{Curr}(S)$ *satisfying*

$$i(\gamma, L_X) = \ell_X(\gamma)$$

for all curves γ on S.

Note that, since the intersection form is continuous and bilinear on $\mathrm{Curr}(S) \times \mathrm{Curr}(S)$, as discussed in Sect. 11.2.6, the hyperbolic length function has a continuous extension to a length function on $\mathrm{Curr}(S)$ by setting

$$\ell_\rho(\mu) = i(\mu, L_\rho)$$

for all $\mu \in \text{Curr}(S)$. The positivity of this function follows from the fact that the Liouville current is filling and hence $i(\mu, L_\rho) = 0$ if and only if μ is the 0-current. Moreover, this extension is unique due to the following theorem by Otal [26].

Theorem 11.3.2 ([26]) *Suppose $\mu_1, \mu_2 \in \text{Curr}(S)$. If $i(\mu_1, \gamma) = i(\mu_2, \gamma)$ for all curves γ on S, then $\mu_1 = \mu_2$.*

More generally, let ρ be any metric on S, and let $\ell_\rho(\gamma)$ denote the length of a shortest representative in the homotopy class of a curve γ. We say L_ρ is a *Liouville current for* ρ if Eq. (11.2) holds, that is

$$i(\gamma, L_\rho) = \ell_\rho(\gamma)$$

for all curves γ on S. Note that when such a current exists it must be unique and is necessarily a filling current, by the same theorem by Otal.

As explained above, a Liouville current exists for any hyperbolic metric on S. Otal [26] showed the existence of a Liouville current for any (variable) negatively curved metric on S. By work of Duchin-Leininger-Rafi [10] and Bankovic-Leininger [3] such a current also exists for any non-positively curved Euclidean cone metric on S. Finally, Constantine [9] extended these results to any non-positively curved (singular) Riemannian metric, giving the Liouville current associated with any such metric (in fact, also for the larger class of so-called *no conjugate points cone metrics*, see [9] for the definition). We record a consequence of this sequence of results here:

Theorem 11.3.3 ([9, Proposition 4.4]) *Let ρ be any (possibly singular) non-positively curved Riemannian metric on S and let $\ell_\rho(\gamma)$ denote the ρ-length of a shortest representative in the homotopy class of γ. Then the length function ℓ_ρ on the set of curves extends continuously to a length function*

$$\ell_\rho : \text{Curr}(S) \to \mathbb{R}.$$

Moreover, this extension is unique.

We note that Liouville currents also exist in other settings. Notably, Martone–Zhang proved the existence of such currents in the context of a large class of representations, including Hitchin and maximal ones, see [22] for details.

From a more algebraic viewpoint, one can consider the word metric on $\pi_1(S)$ with respect to a fixed generating set: We choose a base point p on S and identify the elements of $\pi_1(S)$ with loops based at p. Since this group is finitely generated, we choose a finite, symmetric generating set $G = \{g_1^{\pm 1}, g_2^{\pm 1}, \ldots, g_n^{\pm 1}\}$. Given a conjugacy class $[\gamma]$ (or, equivalently, a homotopy class of a curve γ) we define the word length of the conjugacy class $[\gamma]$ with respect to G to be

$$\ell_G([\gamma]) = \min \left\{ |k_1| + |k_2| + \cdots |k_m| \mid g_{i_1}^{k_1} g_{i_2}^{k_2} \cdots g_{i_m}^{k_m} \in [\gamma] \right\}.$$

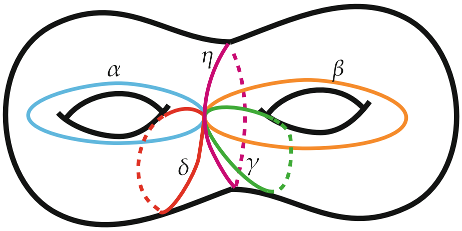


Fig. 11.2 A genus 2 surface S with a simple (non-minimal) generating set $G = \{\alpha^{\pm 1}, \beta^{\pm 1}, \gamma^{\pm 1}, \delta^{\pm 1}, \eta^{\pm 1}\}$

We say a generating set G is *simple* if the loops g_i in G are simple and pairwise disjoint except at the base point p (see Fig. 11.2 for an example). Note that there are many such generating sets, including any one-vertex triangulation of S or the standard generating set for a genus g surface $\{a_1, b_1, a_2, b_2, \ldots a_g, b_g\}$ with the relation $[a_1, b_1] \cdots [a_g, b_g] = 1$.

In [11] it is shown that, given a simple generating set G, there exists a collection of curves $\nu = \nu(G)$, depending only on the generating set, such that the word length of a curve is given exactly by its geometric intersection number with this curve:

Theorem 11.3.4 ([11, Theorem 1.2]) *Let G be a simple generating set for $\pi_1(S)$. Then there exists a collection of curves $\nu = \nu(G)$ on S such that*

$$\ell_G(\gamma) = i(\nu, \gamma)$$

for all curves γ in S. Moreover, ν is unique with this property.

By viewing the set of curves as a subset of $\mathrm{Curr}(S)$, if G is a simple generating set for $\pi_1(S)$, then the above result says that there exists a (unique) Liouville current associated with the corresponding word metric. In particular, it follows that the word length extends continuously to the space of currents:

Corollary 11.3.5 ([11, Corollary 1.3]) *Let G be a simple generating set for $\pi_1(S)$. Then the word length with respect to G on the set of curves extends continuously to a length function*

$$\ell_G : \mathrm{Curr}(S) \to \mathbb{R}.$$

Moreover, this extension is unique.

11.3.2 Stable Length of Currents

There are many notions of lengths not covered by the Liouville currents explained above. Two such occasions are the length of a curve with respect to a Riemannian metric which attains positive curvature values at places, and the word length with respect to a non-simple generating set. In fact, in these settings such currents do not necessarily exist. For instance, if we consider the word metric with respect to a non-simple generating set then we observe that the length function *cannot* extend continuously to a length function on the space of geodesic currents. To see this, consider the case where S is the once-punctured torus and let a, b be the standard generators for the free group $\pi_1(S)$. Then, the word length with respect to the generating set $G = \{a^{\pm 1}, b^{\pm 1}, a^{\pm 2}\}$, does not extend to a continuous homogeneous function on Curr(S). Indeed, the sequence of currents $\left(\frac{1}{2n}a^{2n}b\right)$ converges to the current a as $n \to \infty$ and hence if such a function ℓ_G existed, continuity would imply that

$$\ell_G\left(\frac{1}{2n}a^{2n}b\right) \to \ell_G(a) = 1$$

while, on the other hand, homogeneity would imply

$$\ell_G\left(\frac{1}{2n}a^{2n}b\right) = \frac{1}{2n}\ell_G(a^{2n}b) = \frac{n+1}{2n} \to \frac{1}{2}$$

as $n \to \infty$, a contradiction.

However, as shown in [13], if we consider the *stable length* of curves instead, which we describe below, this length function always extends continuously to the space of geodesic currents.

Let X be any geodesic metric space on which $\pi_1(S)$ acts discretely and cocompactly by isometries. For a conjugacy class $[\gamma]$ in $\pi_1(S)$ (or, equivalently, a curve γ on S), define its translation length $\ell_X(\gamma)$ with respect to X as in (11.1). Then the *stable length* of $[\gamma]$ is defined to be

$$sl_X(\gamma) = \lim_{n\to\infty} \frac{1}{n}\ell_X(\gamma^n) = \lim_{n\to\infty} \frac{1}{n} \inf_{x\in X} d(x, \gamma^n(x)).$$

Again, this definition is independent of the choice of the representative in the conjugacy class. In [13] it is shown that, with X as above, this notion of length always extends continuously to Curr(S):

Theorem 11.3.6 ([13, Theorem 1.5]) *Let X be a geodesic metric space on which $\pi_1(S)$ acts discretely and cocompactly by isometries. Then the stable length function sl_X on the set of curves extends continuously to a length function*

$$sl_X : \mathrm{Curr}(S) \to \mathbb{R}.$$

Moreover, this extension is unique.

The proof of Theorem 11.3.6 is rather involved, with the main difficulty being how to define the stable length of a current, and we will not explain it here, but we refer the reader to [13]. Instead we give some consequences of Theorem 11.3.6.

If we equip S with any Riemannian metric and let X be its universal cover $\tilde{S}$ we immediately get the following corollary:

Corollary 11.3.7 *Let ρ be any Riemannian metric on S. For a curve γ, let $\ell_\rho(\gamma)$ be the ρ-length of a shortest representative. Then the stable length defined by*

$$st_\rho(\gamma) = \lim_{n\to\infty} \frac{1}{n}\ell_\rho(\gamma^n)$$

has a unique continuous extension to a length function

$$st_\rho : \mathrm{Curr}(S) \to \mathbb{R}_+.$$

Similarly, if we let X be the Cayley graph with respect to a finite generating set of $\pi_1(S)$ we also have:

Corollary 11.3.8 *Let G be any finite generating set for $\pi_1(S)$. Let $\ell_G(\gamma)$ denote the shortest word length of a representative in the conjugacy class of γ. Then the stable length defined by*

$$st_G(\gamma) = \lim_{n\to\infty} \frac{1}{n}\ell_G(\gamma^n)$$

has a unique continuous extension to a length function

$$st_G : \mathrm{Curr}(S) \to \mathbb{R}_+.$$

We remark that in [13] Theorem 11.3.6 was proved in a more general setting, namely when the surface group is replaced by any torsion free Gromov hyperbolic group Γ. It is shown that in this setting, the corresponding stable length of a conjugacy class extends to a continuous length function on the *space of (oriented) currents on* Γ. This space, introduced by Bonahon [7] and denoted $\mathrm{Curr}(\Gamma)$, is defined to be the set of Γ-invariant Radon measures on the double boundary

$$(\partial\Gamma \times \partial\Gamma \setminus \Delta) / \sim$$

where $\partial\Gamma$ is the Gromov boundary of Γ and where we identify (γ_1, γ_2) with (γ_2, γ_1) (see, for example, [19]). Since we will not use this more general setting here we refer to [7] and [13] for the precise definitions.

We also remark that Theorem 11.3.6 was proved by Bonahon [7] in the case when X is "uniquely geodesic at infinity", i.e. any two points on the (Gromov) boundary at infinity of X determine a unique geodesic between them. However, this condition is not satisfied in general for the universal cover of Riemannian metrics, nor for Cayley graphs.

Finally we note that, in [7], Bonahon remarks that it should be possible to remove not only the uniquely geodesic hypothesis, which Theorem 11.3.6 proves, but also the cocompact assumption. The proof of Theorem 11.3.6 in [13] still requires Γ to act cocompactly on X and it is an interesting question whether it is a necessary condition.

Question 11.3.9 *Does Theorem 11.3.6 still hold for a surface group that acts discretely, but not cocompactly on X?*

It should be noted that the assumption on the action to be discrete cannot be removed, as shown by Bonahon [7].

11.3.3 Stable Length as a Generalization of Intersection Length

At first glance, extending length of curves to length functions on currents through the intersection length or by considering the stable length might seem like very different approaches. However, as we will observe below, the two notions can be unified: given a filling current ν one can construct a metric space (X, d) on which $\pi_1(S)$ acts discretely and cocompactly by isometries, and such that

$$sl_X(\gamma) = i(\nu, \gamma)$$

for all curves γ on S. The basis for our metric is a semi-distance presented by Glorieux in [16], described below.

Fix a hyperbolic metric ρ on S and let $\tilde{S}$ be the universal covering equipped with the pull-back metric. We define a metric space (X, d) in the following way. As in Sect. 11.3.1, for a geodesic arc $\tilde{\eta}$ let $\mathcal{G}(\tilde{\eta})$ denote the set of geodesics in $\tilde{S}$ that intersect $\tilde{\eta}$ transversely. Let $\partial_\rho \mathcal{G}(\tilde{\eta})$ denote the image of $\mathcal{G}(\tilde{\eta})$ under the homeomorphism that maps each geodesic to its pair of endpoints. Let ν be a filling current in Curr(S). For two distinct points $x, y \in \tilde{S}$, define

$$d'(x, y) = \nu(\partial_\rho \mathcal{G}(\tilde{\eta}))$$

where $\tilde{\eta}$ is the geodesic arc connecting x and y. Set $d'(x, x) = 0$ for all $x \in \tilde{S}$.

Note that d' is symmetric, i.e. $d'(x, y) = d'(y, x)$, and $d'(x, y) \geq 0$ for all $x, y \in \tilde{S}$ (although d' might not separate points, i.e. $d'(x, y) = 0$ need not imply that $x = y$). Furthermore, by the definition of the intersection number (see Sect. 11.2.6), if x lies on the axis of an element $\gamma \in \pi_1(S)$ then

$$d'(x, \gamma(x)) = i(\nu, \gamma). \tag{11.4}$$

Moreover, in [16] it is shown that

(1) d' satisfies the triangle inequality, i.e. $d'(x, y) \leq d'(x, z) + d'(z, y)$ for all $x, y, x \in \tilde{S}$,
(2) $i(\nu, \gamma) \leq d'(x, \gamma(x))$ for all $x \in \tilde{S}$ and $\gamma \in \pi_1(S)$.

In [16] d' was used to find the critical exponent for geodesic currents, here we use it to construct our desired metric space. Define

$$X = \tilde{S}/\sim$$

where $x \sim y$ if and only if $d'(x, y) = 0$, equipped with the metric d induced by d'. That is,

$$d([x], [y]) = d'(x, y)$$

for all $[x], [y] \in X$, where x and y are any representatives of $[x]$ and $[y]$, respectively. Using 11.3.3 above and Eq. (11.4) we see that the stable length with respect to X agrees with the length function defined by the intersection with ν:

$$sl_X(\gamma) = \lim_{n\to\infty} \frac{1}{n} \inf_{x\in X} d(x, \gamma^n(x)) = \lim_{n\to\infty} \frac{1}{n} i(\nu, \gamma^n) = i(\nu, \gamma)$$

for any conjugacy class $[\gamma]$ in $\pi_1(S)$ (or, equivalently, any curve γ on S).

Since ν is $\pi_1(S)$-invariant, $\pi_1(S)$ acts by isometries on (X, d) and, since the action is cocompact on $\tilde{S}$ it is also cocompact on (X, d). Moreover, it is not hard to see that $\pi_1(S)$ acts discretely on (X, d) since ν is filling: if there exists a sequence (γ_n) in $\pi_1(S)$ and $x \in \tilde{S}$ such that $d(x, \gamma_n(x)) \to 0$ as $n \to \infty$, then, by 11.3.3, $i(\nu, \gamma_n) \to 0$ as $n \to \infty$, contradicting the fact that ν is filling.

We have the following result:

Theorem 11.3.10 *Let ν be any filling current. Then there exists a metric space X on which $\pi_1(S)$ acts discretely and cocompactly by isometries such that*

$$sl_X(\gamma) = i(\nu, \gamma)$$

for all curves γ on S. □

11.4 Applications to Counting Curves

In [24, 25] Mirzakhani gives the asymptotic growth rate of the number of curves of bounded length, in each Mod(S)-orbit, as the length grows.

Theorem 11.4.1 ([24, 25, Theorem 1.1]) *Let γ_0 be a curve on S, and ρ be a hyperbolic metric on S. Then*

$$\lim_{L\to\infty} \frac{\#\{\gamma \in \mathrm{Mod}(S)\cdot\gamma_0 \mid \ell_\rho(\gamma) \leq L\}}{L^{6g-6}} = C_{\gamma_0}\cdot m_\rho$$

for some $C_{\gamma_0} > 0$, and $m_\rho = m_{Th}(\{\lambda \in \mathrm{ML}(S) \mid \ell_\rho(\lambda) \leq 1\})$ where m_{Th} is the Thurston measure on ML(S).

The constant C_{γ_0} in Theorem 11.4.1 is independent of the hyperbolic metric ρ. In fact, Mirzakhani [24] showed that

$$C_{\gamma_0} = \frac{n_{\gamma_0}}{\boldsymbol{m}_g}$$

where $n_{\gamma_0} > 0$ depends only on γ_0 and

$$\boldsymbol{m}_g = \int_{\mathcal{M}} m_\rho \, d\mathrm{vol}_{WP} \tag{11.5}$$

where the integral is taken over the moduli space

$$\mathcal{M} = \mathrm{Teich}(S)/\,\mathrm{Mod}(S)$$

with respect to the Weil–Petersson volume form.

The *Thurston measure* is the natural Mod(S)-invariant locally finite measure on ML(S) given by the piecewise linear structure coming from train track coordinates. See Sect. 11.4.1 for details.

The purpose of this section is to discuss a generalization of the theorem of Mirzakhani above, based on the previous section (see Theorem 11.4.3). We will explain why the same asymptotic behavior as in Theorem 11.4.1 holds for other metrics on S, in particular for any Riemannian metric. The results presented are contained in [11, 12] and [13]. The idea behind the proof of the generalization to other metrics crystallized over the above series of papers, so we provide a unified but brief explanation for the statements and proofs of these results.

Remark 11.4.2 Theorem 11.4.1, as well as its generalization Theorem 11.4.3 below, holds for any finite type, orientable surface of negative Euler characteristic (other than the thrice punctured sphere). That is, we can allow S to have n punctures or boundary components, and the same asymptotic behavior holds (where we replace $6g-6$ in the exponent by $6g-6+2n$). However, somewhat surprisingly, orientability is a necessary condition. For non-orientable surfaces the theorems fail, see [15, 21].

11.4.1 Thurston Measure

Recall, from Sect. 11.2.3, that the space $\mathrm{ML}(S)$ of measured laminations has a $\mathrm{Mod}(S)$-invariant piecewise linear manifold structure. Moreover, the PL-manifold is equipped with a $\mathrm{Mod}(S)$-invariant symplectic structure, which gives rise to a $\mathrm{Mod}(S)$-invariant measure in the Lebesgue class. This is the Thurston measure m_{Th}. It is infinite, but locally finite, and satisfies

$$m_{Th}(L \cdot U) = L^{6g-6} \cdot m_{Th}(U)$$

for every Borel set $U \subset \mathrm{ML}(S)$ and $L > 0$ (see [29]). Furthermore, as shown by Masur [23], the Thurston measure m_{Th} is ergodic with respect to the $\mathrm{Mod}(S)$-action on $\mathrm{ML}(S)$, and is the only (up to scaling) invariant measure in the Lebesgue class. Recall that a measure m is said to be ergodic with respect to $\mathrm{Mod}(S)$ if for every $\mathrm{Mod}(S)$-invariant Borel set U we have that either $m(U) = 0$ or $m(U^c) = 0$.

In this section we explain how one can see the Thurston measure (up to scaling) as a limit of a sequence of measures, which gives perhaps a more intuitive feeling of what this measure is.

For each L, define a measure on $\mathrm{ML}(S)$ by

$$m^L = \frac{1}{L^{6g-6}} \sum_{\gamma \in \mathrm{ML}_{\mathbb{Z}}(S)} \delta_{\frac{1}{L}\gamma}$$

where δ_x denotes the Dirac measure centered at x and $\mathrm{ML}_{\mathbb{Z}}(S)$ is the subset of $\mathrm{ML}(S)$ corresponding to integral multicurves. We will show that, as $L \to \infty$, these measures converge to a multiple of the Thurston measure, i.e.

$$\lim_{L \to \infty} \frac{1}{L^{6g-6}} \sum_{\gamma \in \mathrm{ML}_{\mathbb{Z}}(S)} \delta_{\frac{1}{L}\gamma} = c \cdot m_{Th} \tag{11.6}$$

for some $c > 0$. Note that each m^L is $\mathrm{Mod}(S)$-invariant, and hence so is any limit.

It is enough to show the convergence of the measures in each chart given by the linear piecewise structure on $\mathrm{ML}(S)$. Hence we fix a maximal train track τ and let $C(\tau)$ be the solution set to the switch equations of τ. The set $C(\tau)$ is a rational cone of dimension $6g-6$ in $\mathbb{R}^E$, where E is the number of edges of τ, and defines an open set in $\mathrm{ML}(S)$ given by all measured laminations carried by τ. The integral simple multicurves carried by τ correspond exactly to the integer points in $C(\tau)$ which in turn, by the rationality of $C(\tau)$, we identify with a subset of $\mathbb{Z}^{6g-6}$. Accordingly, we identify $C(\tau)$ with a cone $C'(\tau)$ in $\mathbb{R}^{6g-6}$ such that $C'(\tau) \cap \mathbb{Z}^{6g-6}$ correspond to the integral multicurves carried by τ. Finally, we push forward m^L through these identifications to a measure on $\mathbb{R}^{6g-6} \cap C'(\tau)$ which is the restriction of the measure

$$m_\tau^L = \frac{1}{L^{6g-6}} \sum_{p \in \mathbb{Z}^{6g-6}} \delta_{\frac{1}{L}p}$$

(viewed as a measure on $\mathbb{R}^{6g-6}$) to the cone $C'(\tau)$.

It is not hard to see that m_τ^L converges to the Lebegue measure as $L \to \infty$. However, we include an outline for a proof of this statement here, since we will use a similar argument in Sect. 11.4.2 concerning convergence of a family of measures on the space of currents.

Note that the family $(m_\tau^L)_L$ is precompact in the space of Radon measures on $\mathbb{R}^{6g-6}$, meaning that any sequence of measures has a subsequence that weakly converges to a measure. Indeed, since the space of probability measures on a compact metric space is compact, it is enough to show that

$$\limsup_{L \to \infty} m_\tau^L(R_s) < \infty \tag{11.7}$$

where R_s is a (closed) cube of side length s in $\mathbb{R}^{6g-6}$. Clearly we have

$$(s-1)^{6g-6} \leq \# R_s \cap \mathbb{Z}^{6g-6} \leq (s+1)^{6g-6} \tag{11.8}$$

and so

$$m_\tau^L(R_s) = \frac{\#\{p \in \mathbb{Z}^{6g-6} \mid p \in R_{s \cdot L}\}}{L^{6g-6}} \leq \frac{(sL+1)^{6g-6}}{L^{6g-6}}$$

and the limit (superior) of the right hand side is finite. Hence (11.7) holds. Now let m be any limit point of $(m_\tau^L)_L$. Note that for each L, the measure m_τ^L is invariant under translation in the lattice $(\frac{1}{L}\mathbb{Z})^{6g-6}$. It follows that m is translation invariant in $\mathbb{R}^{6g-6}$ and hence must be a multiple of the Lebesgue measure (since this is the unique measure, up to scaling, with this property). For any $(L_n)_n$ with $L_n \to \infty$ there exists a subsequence $(L_{n_k})_k$ such that

$$m_\tau^{L_{n_k}} \to c \cdot \mathfrak{L}$$

for some $c > 0$ as $k \to \infty$, where $\mathfrak{L}$ denotes the Lebesgue measure. Hence, to prove (11.6) we need to show that c is independent of the subsequence. Note that, as above,

$$m_\tau^L(R_1) = \frac{\#\left\{p \in \mathbb{Z}^{6g-6} \mid p \in R_L\right\}}{L^{6g-6}}$$

and the right hand side converges as $L \to \infty$ by (11.8) to 1, that is, to the Lebesgue measure of the unit cube R_1. Hence the limit of the right hand side does not depend on the subsequence and (11.6) follows.

11.4.2 Counting with Respect to Length Functions

Given a hyperbolic metric ρ on S and its corresponding Liouville current L_ρ, one can replace the length function $\ell_\rho(\cdot)$ in Theorem 11.4.1 with the intersection function $i(L_\rho, \cdot)$. In view of this, one can consider the following generalization of the limit appearing in the mentioned theorem:

$$\lim_{L\to\infty} \frac{\#\{\gamma \in \mathrm{Mod}(S)\cdot\gamma_0 \mid i(\nu,\gamma) \leq L\}}{L^{6g-6}} \tag{11.9}$$

where γ_0 is a curve on S and ν is any filling current. (Note that we require ν to be filling to guarantee that there are only finitely many curves with bounded intersection number with ν). In particular, by letting ν be a Liouville current for another metric, such as a variable negatively curved or Euclidean cone metric, this is equivalent to asking if the limit (11.9) exists with respect to this metric.

In [12] it was shown that the limit in (11.9) exists for any filling current ν, and in fact, more generally when the intersection function $i(\nu, \cdot)$ is replaced by any continuous length function $\ell(\cdot)$ defined on the space of currents. Recall that we say ℓ is a length function on $\mathrm{Curr}(S)$ if it is homogeneous and $\ell(\mu) \geq 0$ for all currents μ and $\ell(\mu) = 0$ if and only if $\mu = 0$.

Theorem 11.4.3 ([12]) *Let $\ell : \mathrm{Curr}(S) \to \mathbb{R}$ be any continuous length function and γ_0 a curve on S. Then*

$$\lim_{L\to\infty} \frac{\#\{\gamma \in \mathrm{Mod}(S)\cdot\gamma_0 \mid \ell(\gamma) \leq L\}}{L^{6g-6}} = C_{\gamma_0}\cdot m_\ell$$

where $C_{\gamma_0} > 0$ is the same constant as in Theorem 11.4.1, and

$$m_\ell = m_{Th}\left(\{\lambda \in \mathrm{ML}(S) \mid \ell(\lambda) \leq 1\}\right).$$

Here we give an outline of the arguments involved in proving Theorem 11.4.3, and refer to [12] for the details.

The main idea to prove the convergence of the limit

$$\lim_{L\to\infty} \frac{\#\{\gamma \in \mathrm{Mod}(S)\cdot\gamma_0 \mid \ell(\gamma) \leq L\}}{L^{6g-6}} \tag{11.10}$$

is to consider a sequence of measures on $\mathrm{Curr}(S)$ analogous to the measures on $\mathrm{ML}(S)$ in Sect. 11.4.1. Let $\gamma_0 \in S$ be a curve and define, for each $L > 0$, a measure on Curr(S) by

$$m_{\gamma_0}^L = \frac{1}{L^{6g-6}} \sum_{\gamma \in \mathrm{Mod}(S)\cdot\gamma_0} \delta_{\frac{1}{L}\gamma}.$$

Note that each $m_{\gamma_0}^L$ is locally finite and invariant under the action of Mod(S). In fact, we will see that, as $L \to \infty$ they converge to a Mod(S)-invariant measure on ML(S) that is absolutely continuous with respect to the Thurston measure, and hence, using the ergodicity of m_{Th}, they must converge to a multiple of this measure:

Theorem 11.4.4 ([12, 13, Theorem 5.1]) *Let γ_0 be any curve on S. Then*

$$\lim_{L\to\infty} m_{\gamma_0}^L = C_{\gamma_0} \cdot m_{Th}$$

where $C_{\gamma_0} > 0$ is the constant in Theorem 11.4.1.

First we explain why Theorem 11.4.4 implies Theorem 11.4.3. Fix a continuous length function $\ell : \mathrm{Curr}(S) \to \mathbb{R}$ and let

$$B_\ell = \{\mu \in \mathrm{Curr}(S) \mid \ell(\mu) \le 1\}.$$

Note that the limit (11.10) is equivalent to

$$\lim_{L\to\infty} m_{\gamma_0}^L(B_\ell).$$

The continuity of ℓ implies that B_ℓ is a closed set. Also, for any measurable set U satisfying $U \cap L \cdot U = \emptyset$ for any positive $L \neq 1$, the scaling properties of the Thurston measure imply that $m_{Th}(U) = 0$. To see this, note that for all $L \neq 1$

$$m_{Th}(U \cup L\cdot U) = m_{Th}(U) + m_{Th}(L\cdot U) = m_{Th}(U)(1 + L^{6g-6})$$

and letting $L \to 1$ we get $m_{Th}(U) = 2m_{Th}(U)$, i.e. $m_{Th}(U) = 0$. In particular, $m_{Th}(\partial B_\ell) = 0$. Hence, by the Portmanteau Theorem, see [4],

$$\lim_{L\to\infty} m_{\gamma_0}^L = C_{\gamma_0} \cdot m_{Th}$$

implies that

$$\lim_{L\to\infty} m_{\gamma_0}^L(B_\ell) = C_{\gamma_0} \cdot m_{Th}(B_\ell)$$

where we view m_{Th} as a measure on Curr(S) with full support on the subspace ML(S). Theorem 11.4.3 follows.

Next we outline the arguments proving Theorem 11.4.4. In an attempt to aid the reader we first outline the main steps involved in the proof:

(1) Let m_{γ_0} be any limit point of the family $(m^L_{\gamma_0})_L$, and note that it is $\mathrm{Mod}(S)$-invariant.
(2) We show that m_{γ_0} is supported on $\mathrm{ML}(S)$, and
(3) that m_{γ_0} is absolutely continuous with respect to the Thurston measure m_{Th} on $\mathrm{ML}(S)$.
(4) Ergodicity of m_{Th} with respect to $\mathrm{Mod}(S)$ together with the steps above, implies that $m_{\gamma_0} = C \cdot m_{Th}$ for some $C > 0$.
(5) Finally, using Mirzakhani's theorem (Theorem 11.4.1) we show that the constant C above does not depend on the subsequence and is in fact equal to C_{γ_0}. Hence $m^L_{\gamma_0} \to C_{\gamma_0} \cdot m_{Th}$.

We formalize the conclusion of step 11.4.2 below:

Proposition 11.4.5 ([12, Proposition 4.1]) *Let $(L_n)_n$ be any sequence of positive numbers such that $L_n \to \infty$. Then there is a subsequence $(L_{n_k})_k$ such that*

$$m_{\gamma_0}^{L_{n_k}} \to C \cdot m_{Th}$$

for some $C > 0$, as $k \to \infty$.

As above, due to the Portmanteau Theorem, we get the following consequence:

Corollary 11.4.6 *Let $\ell : \mathrm{Curr}(S) \to \mathbb{R}$ be a continuous length function and let $(L_n)_n$ be any sequence of positive numbers such that $L_n \to \infty$. Then there is a subsequence $(L_{n_k})_k$ such that*

$$m_{\gamma_0}^{L_{n_k}}(B_\ell) \to C \cdot m_{Th}(B_\ell)$$

for some $C > 0$, as $k \to \infty$.

The key idea behind proving Proposition 11.4.5 is to associate to each (generic) curve in $\mathrm{Mod}(S) \cdot \gamma_0$ a simple multi-curve. Specifically, we define a map

$$\pi^\epsilon_{\gamma_0} : \Sigma^\epsilon_{\gamma_0} \to \mathrm{ML}_{\mathbb{Z}}(S)$$

where $\Sigma^\epsilon_{\gamma_0} \subset \mathrm{Mod}(S) \cdot \gamma_0$ is a generic subset such that

$$(1-\epsilon)\ell(\gamma) < \ell(\pi^\epsilon_{\gamma_0}(\gamma)) < (1+\epsilon)\ell(\gamma). \tag{11.11}$$

We say a set Σ is generic if

$$\frac{\#\Sigma}{L^{6g-6}} \to 0$$

as $L \to \infty$. The existence of such a map results from the following observation, which says that the expected angle of self-intersection of a long curve is arbitrarily small.

Theorem 11.4.7 ([12, Theorem 1.2]) *Let $\angle(\gamma)$ denote the largest angle among the self-intersection angles of a curve γ. Let $\gamma_0 \subset S$ be a curve and ρ a hyperbolic metric. Then*

$$\lim_{L\to\infty} \frac{\#\{\gamma \in \mathrm{Mod}(S)\cdot\gamma_0 \mid \ell_\rho(\gamma) \leq L, \angle(\gamma) \geq \delta\}}{L^{6g-6}} = 0$$

for all $\delta > 0$.

The proof of Theorem 11.4.7 is quite involved (see [12, Section 3.3]), but the general idea is that large self-intersection angles result in ideal 4-gons on the surface which most of the curves have to avoid. The set of curves on S which do not intersect a 4-gon must live on a proper subsurface and hence the number of these curves of length bounded by L must grow at a slower rate than L^{6g-6}. This idea is inspired by the fact that the subspace of $\mathrm{ML}(S)$ of measured laminations carried by non-maximal train tracks (i.e. train tracks that have complementary regions larger than triangles) has dimension strictly less than $6g-6$.

Armed with Theorem 11.4.7, we can resolve the self-intersections and end up with a simple multi-curve whose length is close to the length of the original curve (see [12, Section 3.4] for details), and this is the idea for the map $\pi^\epsilon_{\gamma_0}$. In particular, for any $\epsilon > 0$ there is an angle bound $\delta > 0$ such that any curve γ with self-intersection angles less than δ is mapped to a simple multi-curve $\pi^\epsilon_{\gamma_0}(\gamma)$ satisfying (11.11). These curves are what make up the generic set $S^\epsilon_{\gamma_0}$.

We fix $\epsilon > 0$ and suppress the superscript in $\pi^\epsilon_{\gamma_0}$ for ease of notation. It is clear that π_{γ_0} is finite-to-one, but the main useful property of the map, and the key technical difficulty of the proof (details of which will be omitted here, see [12, Section 2.4]) is that it is uniformly bounded-to-1. That is:

Lemma 11.4.8 ([12, Proposition 3.9]) *There exists a constant $K = K(\gamma_0) > 0$ such that*

$$|\pi^{-1}_{\gamma_0}(\lambda)| < K \tag{11.12}$$

for all $\lambda \in \mathrm{ML}_{\mathbb{Z}}(S)$.

We note that any limit point m_{γ_0} is locally finite and $\mathrm{Mod}(S)$-invariant since this is true for each $m^L_{\gamma_0}$. We then use Lemma 11.4.8 to show that any limit point is also uniformly continuous with respect to the Thurston measure. To do so, we first push forward the measure $m^L_{\gamma_0}$ via π_{γ_0} resulting in the following measure supported on $\mathrm{ML}(S)$:

$$n^L_{\gamma_0} = \frac{1}{L^{6g-6}} \sum_{\lambda\in\mathrm{ML}_{\mathbb{Z}}(S)} |\pi^{-1}_{\gamma_0}(\lambda)|\delta_{\frac{1}{L}\lambda}.$$

It is not difficult to see that m_{γ_0} is a limit point of the family $(m^L_{\gamma_0})_L$ if an only if it is a limit point of the family $(n^L_{\gamma_0})_L$. In particular, any limit point is supported on $\mathrm{ML}(S)$, completing step 11.4.2.

Now, (11.12) implies that

$$n^L_{\gamma_0} = \frac{1}{L^{6g-6}} \sum_{\lambda \in \mathrm{ML}_{\mathbb{Z}}(S)} |\pi^{-1}_{\gamma_0}(\lambda)| \delta_{\frac{1}{L}\lambda} < K \cdot \frac{1}{L^{6g-6}} \sum_{\lambda \in \mathrm{ML}_{\mathbb{Z}}(S)} \delta_{\frac{1}{L}\lambda}$$

and the right hand side converges to a multiple of m_{Th} as $L \to \infty$ (see (11.6)). In particular, any limit point of $(n^L_{\gamma_0})_L$, and hence of $(m^L_{\gamma_0})_L$, is absolutely continuous with respect to the Thurston measure, completing step 11.4.2.

Next, recall that, by a result of Masur [23], the Thurston measure is ergodic with respect to the action of $\mathrm{Mod}(S)$ on $\mathrm{ML}(S)$. Hence, since any limit m_{γ_0} of $(m^L_{\gamma_0})_L$ is invariant under this action and absolutely continuous with respect to the Thurston measure, the only choice for m_{γ_0} is a positive multiple of the Thurston measure. This completes the argument for proving Proposition 11.4.5 (and hence step 11.4.2).

Finally, we use Mirzakhani's result (Theorem 11.4.1) to complete the outline of the proof of Theorem 11.4.4. We need to show that the constant C in Proposition 11.4.5 is independent of the subsequence and that C is in fact equal to the constant C_{γ_0}.

Let ρ be a hyperbolic metric on S and L_ρ the corresponding Liouville current. Let $\ell_\rho : \mathrm{Curr}(S) \to \mathbb{R}$ be the length function defined by

$$\ell_\rho(\mu) = i(\mu, L_\rho)$$

which agrees with the hyperbolic length on curves. Following the notation above, let

$$B_{\ell_\rho} = \{\mu \in \mathrm{Curr}(S) \mid i(\mu, L_\rho) \leq 1\}.$$

By definition,

$$m^L_{\gamma_0}(B_{\ell_\rho}) = \frac{\#\{\gamma \in \mathrm{Mod}(S) \cdot \gamma_0 \mid \ell_\rho(\gamma) \leq L\}}{L^{6g-6}}.$$

By Theorem 11.4.1 we know that the right hand side converges to

$$C_{\gamma_0} \cdot m_{Th}(B_{\ell_\rho}).$$

In particular, $m^L_{\gamma_0}(B_{\ell_\rho})$ converges and by Corollary 11.4.6 it must converge to $C \cdot m_{Th}(B_{\ell_\rho})$ for some $C > 0$. Hence we have $C = C_{\gamma_0}$, completing step 11.4.2, and Theorem 11.4.4 follows.

Lastly, we note Theorem 11.4.4 tells in particular that we have the asymptotic growth

$$\#\{\gamma \in \mathrm{Mod}(S) \cdot \gamma_0 \mid \ell(\gamma) \le L\} \sim const \cdot L^{6g-6}$$

for any of the length functions ℓ discussed in Sect. 11.3. In particular, this result is true for the length induced by any non-positive (singular) Riemannian metric on S as well as for the stable length with respect to a geodesic metric space X on which $\pi_1(S)$ acts discretely and cocompactly by isometries (see Theorem 11.3.6). However, in [13] it is shown that it is enough for the stable length to extend to $\mathrm{Curr}(S)$ to conclude that the asymptotics above hold for the actual (translation) length. In particular, it holds for *any* Riemannian metric on S.

Corollary 11.4.9 ([13, Corollaries 1.3 and 1.4]) *Let γ_0 be a curve on S. If ρ is any (possibly singular) Riemannian metric on S and ℓ_ρ is the corresponding length function on curves, then*

$$\lim_{L\to\infty} \frac{\#\{\gamma \in \mathrm{Mod}(S) \cdot \gamma_0 \mid \ell_\rho(\gamma) \le L\}}{L^{6g-6}}$$

exists and is positive. Similarly, if we replace the length function with the word length with respect to any finite generating set of $\pi_1(S)$ then the corresponding limit also exists and is positive.

11.4.3 Orbits of Currents

We end by remarking that one could also ask whether the limit in Theorem 11.4.3 exists if we look at the $\mathrm{Mod}(S)$-orbit of any *current* instead of a curve. Rafi-Souto proved that this is indeed the case:

Theorem 11.4.10 ([28, Main Theorem]) *Let $\ell : \mathrm{Curr}(S) \to \mathbb{R}$ be a continuous length function. For any filling current $\nu \in \mathrm{Curr}(S)$ we have*

$$\lim_{L\to\infty} \frac{\#\{\mu \in \mathrm{Mod}(S) \cdot \nu \mid \ell(\mu) \le L\}}{L^{6g-6}} = C_\nu \cdot m_\ell$$

where $C_\nu > 0$ and $m_\ell = m_{Th}(\{\lambda \in \mathrm{ML} \mid \ell(\lambda) \le 1\})$.

The constant C_ν, as in Theorem 11.4.1 is independent of ℓ and can be written as

$$C_\nu = \frac{n_\nu}{\boldsymbol{m}_g}$$

where $\boldsymbol{m}_g$ is the same constant as in (11.5). However, in [28] the constant n_ν, in the case when ν is filling, is also described:

$$n_\nu = m_{Th}(\{\lambda \in \mathrm{ML} \mid i(\nu, \lambda) \le 1\}).$$

The proof of Theorem 11.4.10 follows a similar logic to the proof of Theorem 11.4.3 above. However, in order to generalize Proposition 11.4.5 to hold also when γ_0 is a filling current, Rafi and Souto combine this proposition together with a deep result of Lindenstrauss–Mirzakhani [20] about the classifications of invariant measures on $\mathrm{ML}(S)$.

We note that Theorem 11.4.10 holds also for surfaces with boundary, as do Theorems 11.4.1 and 11.4.3, but unlike the latter two which also work for surfaces with cusps, Theorem 11.4.10 requires S to be compact (or alternatively, that ν has compact support).

As an application to Theorem 11.4.10, Rafi–Souto prove the asymptotic growth of lattice points in Teichmüller space with respect to the Thurston metric. As before, for a length function $f : \mathrm{Curr}(S) \to \mathbb{R}$ we let m_f denote the constant

$$m_f = m_{Th}(\{\lambda \in \mathrm{ML}(S) \mid f(\lambda) \leq 1\})$$

and we let m_X denote the corresponding constant when $f = \ell_X$, the hyperbolic length on $X \in \mathrm{Teich}(S)$.

Theorem 11.4.11 ([28, Theorem 1.1]) *Let* $X, Y \in \mathrm{Teich}(S)$. *Then*

$$\lim_{R\to\infty} \frac{\#\{\varphi \in \mathrm{Mod}(S) \mid d_{Th}(X, \varphi(Y)) \leq R\}}{e^{(6g-6)R}} = \frac{m_{D_X} m_Y}{\boldsymbol{m}_g}$$

where d_{Th} *denotes the Thurston metric on* $\mathrm{Teich}(S)$, $\boldsymbol{m}_g$ *is as above, and*

$$D_X(\mu) = \max_{\lambda\in\mathrm{ML}(S)} \frac{i(\lambda, \mu)}{\ell_X(\lambda)}.$$

The analogous result of Theorem 11.4.11 when the Thurston metric is replaced by the Teichmüller metric was proved, using different methods, by Athreya-Bufetov–Eskin–Mirzakhani in [2].

11.5 Dynamics of Pseudo-Anosov Homeomorphisms

The purpose of this section is to give a concise proof of a folklore result using Bonahon's intersection function on the space of currents: pseudo-Anosov homeomorphisms of closed hyperbolic surfaces act on the space of projective geodesic currents with uniform North-South dynamics.

Theorem 11.5.1 *Let S be closed hyperbolic surface and $\varphi : S \to S$ be a pseudo-Anosov homeomorphism. Then φ acts on the space of projective geodesic currents* $\mathbb{P}\mathrm{Curr}(S)$ *with uniform North-South dynamics: The action of φ on* $\mathbb{P}\mathrm{Curr}(S)$ *has exactly two fixed points* $[\lambda_+]$ *and* $[\lambda_-]$ *and for any open neighborhood* $U_\pm$ *of* $[\lambda_\pm]$

and a compact set $K_{\pm} \subset \mathbb{P}\mathrm{Curr}(S) \setminus [\lambda_{\mp}]$*, there exist an exponent* $M \geq 1$ *such that* $\varphi^{\pm n}(K_{\pm}) \subset U_{\pm}$ *for all* $n \geq M$.

The idea of the proof is as follows: The set of non-zero currents that have zero intersection with the *stable* current/lamination is precisely the positive scalar multiples of the *stable* current/lamination. Similarly, the set of non-zero currents that has zero intersection with the *unstable* current/lamination is precisely the positive scalar multiples of the *unstable* current/lamination, see Lemma 11.5.2.

Using Lemma 11.5.2 we define continuous functions J_+ and J_- on the space of projective currents which take the value 0 only on $[\lambda_+]$ and $[\lambda_-]$ respectively. We then use these functions to construct neighborhoods of $[\lambda_+]$ and $[\lambda_-]$ and use the properties of intersection function to get convergence estimates.

The proof we present here is motivated by Ivanov's proof of North-South dynamics in the setting of projective measured laminations [18], and consists of putting together a series of lemmas, which we first state and prove.

Lemma 11.5.2 *Let* $\varphi : S \to S$ *be a pseudo-Anosov homeomorphism on a closed hyperbolic surface and* λ_+ *and* λ_- *be the corresponding unstable and stable laminations for* φ*. Then,*

$$i(\lambda_{\pm}, \mu) = 0 \text{ if and only if } \mu = k\lambda_{\pm}$$

for some $k \geq 0$.

Proof Here we give a brief idea of the proof and refer the reader to proof of [31, Proposition 3.1] for details in the case of non-closed surfaces, where the proof is more involved. Let λ_+ be the unstable lamination on S corresponding to the pseudo-Anosov homeomorphism f. The proof for λ_- is almost identical.

We first prove the easy direction of the statement. Namely, let $\mu = k\lambda_+$, and $\alpha > 1$ be such that $\varphi(\lambda_+) = \alpha\lambda_+$. Then, by properties of the intersection number we have

$$\begin{aligned} i(k\lambda_+, \lambda_+) &= i(\varphi^n(k\lambda_+), \varphi^n(\lambda_+)) \\ &= i(\alpha^n k\lambda_+, \alpha^n\lambda_+) \\ &= \alpha^{2n} i(k\lambda_+, \lambda_+) \end{aligned}$$

which implies $i(k\lambda_+, \lambda_+) = 0$.

For the forward implication, we first cut the surface along the leaves of the unstable lamination. The complementary regions are finite sided ideal polygons, [8, Proposition 5.3]. Let μ be any current such that $i(\mu, \lambda_+) = 0$. Let ℓ be any leaf in the support of μ. Since the projection of this leaf onto the surface cannot intersect the leaves of the unstable lamination transversely, there are two possibilities for this projection. Either ℓ projects onto a leaf of the lamination λ or it is a complete geodesic that is asymptotic to two different sides of a complementary polygon. In the second case, this leaf cannot support any measure, otherwise the corresponding current would not be locally finite. Hence μ and λ_+ have the same support, and unique ergodicity of λ_+ implies that $\mu = k\lambda_+$. □

Fix a filling current ν on S, and consider the following two functions $J_+, J_- : \mathbb{P}\mathrm{Curr}(S) \to \mathbb{R}_{\geq 0}$ defined by

$$J_+([\mu]) = \frac{i(\mu, \lambda_+)}{i(\mu, \nu)}, \quad J_-[\mu] = \frac{i(\mu, \lambda_-)}{i(\mu, \nu)}$$

where μ is any representative of $[\mu]$. Note that J_+, J_- are well defined and continuous since the intersection function is continuous and homogeneous, and the denominator is non-zero by the choice of ν.

Lemma 11.5.3 *Let α be the stretch factor for the pseudo-Anosov element φ and let ν be a filling current. If K is a compact set in $\mathbb{P}\mathrm{Curr}(S) \setminus [\lambda_-]$, then there exist $C > 0$ such that*

$$\frac{1}{i(\varphi^n(\mu), \nu)} \leq \frac{C}{\alpha^n i(\mu, \nu)}$$

for all μ such that $[\mu] \in K$.

Proof Since $\mathbb{P}\mathrm{Curr}(S)$ is compact, there exist $0 < C_1 < \infty$ such that

$$J_-([\mu]) = \frac{i(\mu, \lambda_-)}{i(\mu, \nu)} \leq C_1$$

i.e.

$$i(\mu, \lambda_-) \leq C_1 i(\mu, \nu)$$

for all nonzero $\mu \in \mathrm{Curr}(S)$.

Furthermore, by Lemma 11.5.2 the quantity $i(\mu, \lambda_-)$ is non-zero for any μ such that $[\mu] \in \mathbb{P}\mathrm{Curr}(S) \setminus [\lambda_-]$. Therefore, by compactness of K, there exist $C_2 > 0$ such that

$$i(\mu, \lambda_-) \geq C_2 i(\mu, \nu)$$

for all μ with $[\mu] \in K$.

From these two inequalities we obtain, for all μ such that $[\mu] \in K$,

$$\begin{aligned} i(\varphi^n(\mu), \nu) \geq \frac{1}{C_1} i(\varphi^n(\mu), \lambda_-) &= \frac{1}{C_1} i(\mu, \varphi^{-n}(\lambda_-)) \\ &= \frac{1}{C_1} i(\mu, \alpha^n \lambda_-) \end{aligned}$$

$$= \frac{1}{C_1} \alpha^n i(\mu, \lambda_-)$$
$$\geq \frac{C_2}{C_1} \alpha^n i(\mu, \nu).$$

Setting $C = \frac{C_1}{C_2}$, the conclusion of the lemma follows.

□

Lemma 11.5.4 *Let U be an open neighborhood of $[\lambda_+]$ and K be a compact set in $\mathbb{P}\mathrm{Curr}(S) \setminus [\lambda_-]$. There exist $M_1 > 0$ such that*

$$\varphi^n(K) \subset U$$

for all $n \geq M_1$.

Proof Since $i(\mu, \lambda_+) = 0$ if and only if $[\mu] = [\lambda_+]$, and $\mathbb{P}\mathrm{Curr}(S) \setminus U$ is compact, the function $J_+([\mu])$ has a positive absolute minimum on the set $\mathbb{P}\mathrm{Curr}(S) \setminus U$, say $\epsilon > 0$. Therefore, it suffices to prove that $J_+(\varphi^n[\mu]) < \epsilon$ for all $[\mu] \in K$, and for all large n in order to obtain the conclusion of the lemma.

On the other hand, $\mathbb{P}\mathrm{Curr}(S)$ is compact, so the function $J_+([\mu])$ has an upper bound, i.e. there exists $0 < D < \infty$ such that

$$\frac{i(\mu, \lambda_+)}{i(\mu, \nu)} \leq D$$

for all μ.

Let $\epsilon > 0$ be as above, and choose $M_1 > 0$ such that $\frac{DC}{\alpha^{2M_1}} < \epsilon$ where α is the stretch factor of φ and C is the constant given by Lemma 11.5.3. Then, for all $[\mu] \in K$ we have

$$J_+(\varphi^n[\mu]) = \frac{i(\varphi^n(\mu), \lambda_+)}{i(\varphi^n(\mu), \nu)} = \frac{i(\mu, \varphi^{-n}(\lambda_+))}{i(\varphi^n(\mu), \nu)}$$
$$= \frac{\alpha^{-n} i(\mu, \lambda_+)}{i(\varphi^n(\mu), \nu)}$$
$$\leq \frac{C\alpha^{-n} i(\mu, \lambda_+)}{\alpha^n i(\mu, \nu)}$$
$$\leq \frac{DC}{\alpha^{2n}} < \epsilon$$

for all $n \geq M_1$. □

We are now ready to prove the theorem:

Proof of Theorem 11.5.1 Using an argument symmetric to the one in the proof of Lemma 11.5.4, we can show that given any compact set $K \subset \mathbb{P}\mathrm{Curr}(S) \setminus [\lambda_-]$ and an open neighborhood U of $[\lambda_-]$ there exist $M_2 > 0$ such that $\varphi^{-n}(K) \subset U$ for all $n \geq M_2$. The theorem now follows by setting $M = \max\{M_1, M_2\}$. □

In fact, we have much more precise information in terms of pointwise dynamics:

Theorem 11.5.5 *Let $\alpha > 1$ be the stretch factor for φ. Then, for any $[\mu] \neq [\lambda_-]$,*

$$\lim_{n\to\infty} \alpha^{-n}\varphi^n(\mu) = c_\mu \lambda_+$$

for some $c_\mu > 0$, and for any $[\mu'] \neq [\lambda_+]$

$$\lim_{n\to\infty} \alpha^{-n}\varphi^{-n}(\mu') = c_{\mu'} \lambda_-$$

for some $c_{\mu'} > 0$.

Proof The proof builds on the analogous result in the case of laminations, and nearly identical to the case where S has boundary components, see the proof of [31, Theorem 3.4]. □

Recall, from Sect. 11.3.1, that given a hyperbolic metric ρ on S the hyperbolic length extends to a continuous length function ℓ on $\mathrm{Curr}(S)$ given by

$$\ell_\rho(\mu) = i(L_\rho, \mu)$$

for all $\mu \in \mathrm{Curr}(S)$, where L_ρ is the Liouville currents associated with ρ. As an application to this we get as a corollary to the North-South dynamics the following generalization, to *all* curves, of a well known result about simple closed curves:

Corollary 11.5.6 *For any pseudo-Anosov homeomorphism $\varphi : S \to S$ of a closed, orientable hyperbolic surface S, there exists $M > 0$ such that for any essential (not necessarily simple) closed curve γ on S, either*

$$\ell_\rho(\varphi^k\gamma) > \ell_\rho(\gamma) \ \textit{ or } \ \ell_\rho(\varphi^{-k}\gamma) > \ell_\rho(\gamma)$$

for all $k \geq M$.

Proof We will show that there exists $M > 0$ such that for all γ on S, the following holds:

$$\frac{\ell(\varphi^k\gamma) + \ell(\varphi^{-k}\gamma)}{\ell(\gamma)} > 2$$

for all $k \geq M$. Let L_ρ be the Liouville current associated with the hyperbolic metric ρ. Note that, for all k,

$$\ell_\rho(\varphi^k \gamma) = i(\varphi^k \gamma, L_\rho) = i(\gamma, \varphi^{-k} L_\rho)$$

and

$$\ell_\rho(\varphi^{-k} \gamma) = i(\varphi^{-k} \gamma, L_\rho) = i(\gamma, \varphi^k L_\rho).$$

Hence it suffices to prove that there exist $M > 0$ such that

$$\frac{i(\gamma, \varphi^k L_\rho) + i(\gamma, \varphi^{-k} L_\rho)}{\ell_\rho(\gamma)} > 2 \tag{11.13}$$

for all γ and for all $k \geq M$.

Let $\alpha > 1$ be the stretch factor for φ and λ_+ and λ_- its unstable and stable laminations, respectively. Using the properties of the intersection form we have

$$\begin{aligned}\frac{i(\gamma_k, \varphi^k L_\rho) + i(\gamma_k, \varphi^{-k} L_\rho)}{\ell_\rho(\gamma_k)} &= i\left(\frac{\gamma_k}{\ell_\rho(\gamma_k)}, \varphi^k L_\rho\right) + i\left(\frac{\gamma_k}{\ell_\rho(\gamma_k)}, \varphi^{-k} L_\rho\right) \\ &= \alpha^k i\left(\frac{\gamma_k}{\ell_\rho(\gamma_k)}, \alpha^{-k}\varphi^k L_\rho\right) + \alpha^k i\left(\frac{\gamma_k}{\ell_\rho(\gamma_k)}, \alpha^{-k}\varphi^{-k} L_\rho\right)\end{aligned}$$

Since the length of $\dfrac{\gamma_k}{\ell_\rho(\gamma_k)}$ is 1 for all k, they lie in a compact set and hence there exist $\mu \in \mathrm{Curr}(S)$ such that (up to passing to a subsequence)

$$\lim_{k\to\infty} \frac{\gamma_k}{\ell(\gamma_k)} = \mu.$$

On the other hand, Theorem 11.5.5 implies that, for some $c_0, c_1 > 0$

$$\lim_{k\to\infty} \alpha^{-k}\varphi^k L_\rho = c_0 \lambda_+ \text{ and } \lim_{k\to\infty} \alpha^{-k}\varphi^{-k} L_\rho = c_1 \lambda_-.$$

Therefore, as $k \to \infty$,

$$i\left(\frac{\gamma_k}{\ell_\rho(\gamma_k)}, \alpha^{-k}\varphi^k L_\rho\right) \to c_0 i(\mu, \lambda_+) \text{ and } i\left(\frac{\gamma_k}{\ell_\rho(\gamma_k)}, \alpha^{-k}\varphi^{-k} L_\rho\right) \to c_1 i(\mu, \lambda_-).$$

Since $\{\lambda_+, \lambda_-\}$ is a filling pair of currents, at least one of the quantities $i(\mu, \lambda_+)$ or $i(\mu, \lambda_-)$ must be positive. Hence, since $\alpha^k \to \infty$ as $k \to \infty$, we have

$$\frac{i(\gamma_k, \varphi^k L_\rho) + i(\gamma_k, \varphi^{-k} L_\rho)}{\ell_\rho(\gamma_k)} \to \infty.$$

In particular, there exists $M > 0$ such that (11.13) holds for all $k \geq M$.

□

Acknowledgments We are grateful to David Constantine, Spencer Dowdall, Ilya Kapovich, Chris Leininger, Kasra Rafi, Juan Souto and Weixu Su for interesting conversations throughout this project. We thank Dave Futer for asking a question that led to Corollary 11.5.6. The first author also thanks the School of Mathematics at Fudan University and Vanderbilt University for their hospitality, and gratefully acknowledges support from the NSF grant DMS-1500180 (A. Olshanskii and M. Sapir). Finally, we thank the anonymous referee for helpful comments and corrections.

References

1. J. Aramayona, C.J. Leininger, Hyperbolic structures on surfaces and geodesic currents, in *Algorithmic and Geometric Topics Around Free Groups and Automorphisms*. Advanced Courses in Mathematics CRM Barcelona (Birkhäuser, Springer, Cham, 2017), pp. 111–149
2. J. Athreya, A. Bufetov, A. Eskin, M. Mirzakhani, Lattice point asymptotics and volume growth on Teichmüller space. Duke Math. J. **161**(6), 1055–1111 (2012)
3. A. Bankovic, C.J. Leininger, Marked-length-spectral rigidity for flat metrics. Trans. Amer. Math. Soc. **370**(3), 1867–1884 (2018)
4. P. Billingsley, *Convergence of Probability Measures*, 2nd ed. Wiley Series in Probability and Statistics: Probability and Statistics (John Wiley & Sons, Inc., New York, 1999). A Wiley-Interscience Publication
5. F. Bonahon, Bouts des variétés hyperboliques de dimension 3. Ann. Math. (2) **124**(1), 71–158 (1986)
6. F. Bonahon, The geometry of Teichmüller space via geodesic currents. Invent. Math. **92**(1), 139–162 (1988)
7. F. Bonahon, Geodesic currents on negatively curved groups, in *Arboreal Group Theory (Berkeley, 1988)*, vol. 19. Mathematical Sciences Research Institute Publications (Springer, New York, 1991), pp. 143–168
8. A.J. Casson, S.A. Bleiler, *Automorphisms of Surfaces After Nielsen and Thurston*, vol. 9. London Mathematical Society Student Texts. (Cambridge University Press, Cambridge, 1988)
9. D. Constantine, Marked length spectrum rigidity in non-positive curvature with singularities. Indiana Univ. Math. J. **67**(6), 2337–2361 (2018)
10. M. Duchin, C.J. Leininger, K. Rafi, Length spectra and degeneration of flat metrics. Invent. Math. **182**(2), 231–277 (2010)
11. V. Erlandsson, A remark on the word length in surface groups. Trans. Amer. Math. Soc. **372**(1), 441–455 (2019)
12. V. Erlandsson, J. Souto, Counting curves in hyperbolic surfaces. Geom. Funct. Anal. **26**(3), 729–777 (2016)
13. V. Erlandsson, H. Parlier, J. Souto, Counting curves, and the stable length of currents. J. Eur. Math. Soc. (JEMS) **22**(6), 1675–1702 (2020)

14. A. Fathi, F. Laudenbach, V. Poénaru, *Thurston's Work on Surfaces*, vol. 48. Mathematical Notes (Princeton University Press, Princeton, 2012). Translated from the 1979 French original by Djun M. Kim and Dan Margalit
15. M. Gendulphe, What's wrong with the growth of simple closed geodesics on nonorientable hyperbolic surfaces (2017). arXiv:1706.08798
16. O. Glorieux, Critical exponent for geodesic currents (2017). arXiv:1704.06541
17. S. Hersonsky, F. Paulin, On the rigidity of discrete isometry groups of negatively curved spaces. Comment. Math. Helv. **72**(3), 349–388 (1997)
18. N.V. Ivanov, *Subgroups of Teichmüller Modular Groups*, vol. 115. Translations of Mathematical Monographs (American Mathematical Society, Providence, 1992). Translated from the Russian by E.J.F. Primrose and revised by the author
19. I. Kapovich, N. Benakli, Boundaries of hyperbolic groups, in *Combinatorial and Geometric Group Theory (New York, 2000/Hoboken, NJ, 2001)*, vol. 296 (Contemp. Math. Amer. Math. Soc., Providence, 2002), pp. 39–93
20. E. Lindenstrauss, M. Mirzakhani, Ergodic theory of the space of measured laminations. Int. Math. Res. Not. IMRN **4**, Art. ID rnm126, 49 (2008)
21. M. Magee, Counting one-sided simple closed geodesics on Fuchsian thrice punctured projective planes. Int. Math. Res. Not. IMRN **6**, Art. ID rnm112 (2018)
22. G. Martone, T. Zhang, Positively ratioed representations. Comment. Math. Helv. **94**(2), 273–345 (2019)
23. H. Masur, Ergodic actions of the mapping class group. Proc. Am. Math. Soc. **94**(3), 455–459 (1985)
24. M. Mirzakhani, Growth of the number of simple closed geodesics on hyperbolic surfaces. Ann. Math. (2) **168**(1), 97–125 (2008)
25. M. Mirzakhani, Counting mapping class group orbits on hyperbolic surfaces (2016). arXiv:1601.03342
26. J.-P. Otal, Le spectre marqué des longueurs des surfaces à courbure négative. Ann. Math. (2) **131**(1), 151–162 (1990)
27. R.C. Penner, J.L. Harer, *Combinatorics of Train Tracks*, vol. 125. Annals of Mathematics Studies (Princeton University Press, Princeton, 1992)
28. K. Rafi, J. Souto, Geodesic currents and counting problems. Geom. Funct. Anal. **29**(3), 871–889 (2019)
29. W.P. Thurston, The geometry and topology of 3-manifolds (1980). Unpublished notes
30. W.P. Thurston, On the geometry and dynamics of diffeomorphisms of surfaces. Bull. Amer. Math. Soc. (N.S.) **19**(2), 417–431 (1988)
31. C. Uyanik, Generalized north-south dynamics on the space of geodesic currents. Geom. Dedicata. **177**, 129–148 (2015)

Chapter 12
Big Mapping Class Groups: An Overview

Javier Aramayona and Nicholas G. Vlamis

A Domingo, in memoriam.

Abstract We survey recent developments on mapping class groups of surfaces of infinite topological type.

12.1 Introduction

In the blogpost [30], D. Calegari proposed the study of the mapping class group $\mathrm{Map}(\mathbb{R}^2 \setminus C)$, where C denotes a Cantor set. More concretely, he posed the question of whether this group has an infinite-dimensional space of quasimorphisms, as is the case with the mapping class group of a surface of finite topological type, after a celebrated result of Bestvina–Fujiwara [19]. In addition, Calegari suggested a line of attack on the problem, in analogy with Bestvina–Fujiwara's original argument; in a nutshell, the first idea is to prove that a certain *complex of arcs* on which $\mathrm{Map}(\mathbb{R}^2 \setminus C)$ acts is hyperbolic and has infinite diameter, and then exhibit elements which act *weakly properly discontinously* [19] on this complex.

This strategy was successfully implemented by J. Bavard in her thesis [13] (English translation: [14]), and has since caused a surge of interest in mapping class groups of infinite-type surfaces (or *big mapping class groups*, in the terminology coined by Calegari) among the geometric group theory and low-dimensional topology communities. Most of the results to date have focused on the basic structure of big mapping class groups, as well as on the similarities and differences with mapping class groups of finite-type surfaces.

J. Aramayona (✉)
ICMAT (CSIC-UAM-UC3M-UCM), Madrid, Spain
e-mail: Javier.aramayona@icmat.es

N. G. Vlamis
Queens College, City University of New York, Flushing, NY, USA
e-mail: nicholas.vlamis@qc.cuny.edu

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_12

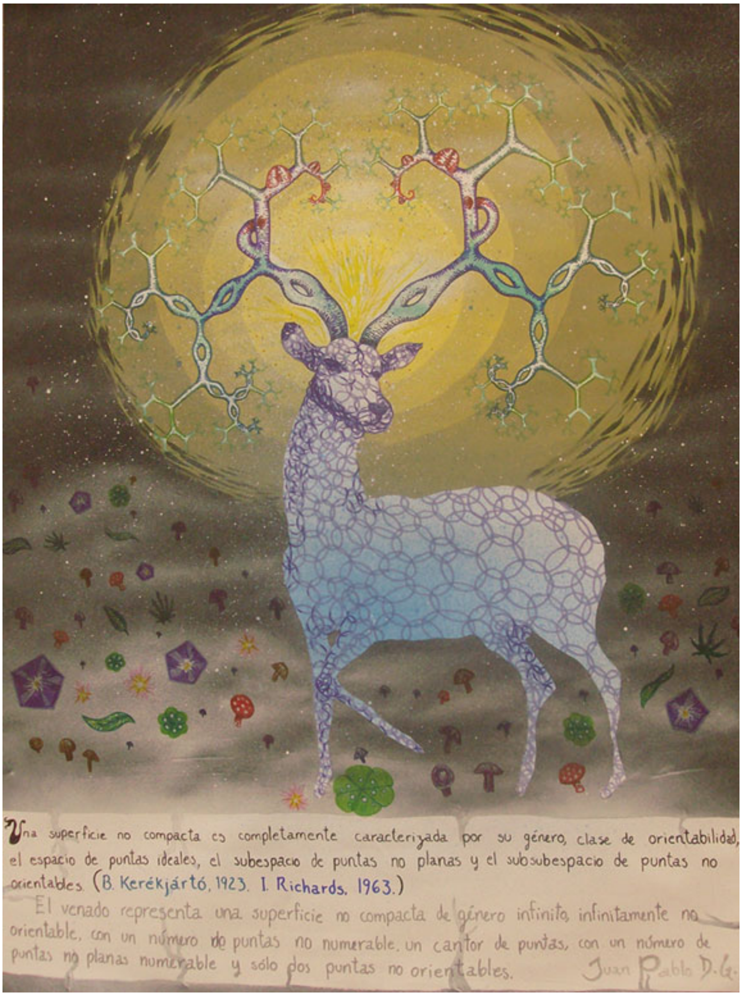


Fig. 12.1 *Infinite-type deer*, by Juan Pablo Díaz González, UNAM

This said, big mapping class groups made their appearance in other related areas of mathematics quite a long time ago. For instance, big mapping class groups arise naturally in the context of stable properties of mapping class groups [84]; infinite-type surfaces are intimately related to the study of quasiconformal maps [18]; the so-called *braided Thompson's group* BV of Brin [28] and Dehornoy [35] is naturally a subgroup of the mapping class group of a sphere minus a Cantor set; etc.

The aim of this survey is to give an overview of the recent developments around big mapping class groups, mainly from the point of view of geometric group theory, and to describe some of the connections to other areas of mathematics, such as Polish groups and Thompson's groups. Along the way, we will offer open problems related to the topics covered.

Plan of the Chapter All the objects and definitions needed in the exposition are introduced in Sect. 12.2. In Sect. 12.3, we present two results which are crucial to a large number of the results discussed in subsequent sections. Section 12.4 deals with topological aspects of big mapping class groups: generation, Polish structure, etc. Section 12.5 concerns algebraic results: automorphisms, homology, relation with Thompson's groups, etc. Finally, in Sect. 12.6 we will concentrate on the action of big mapping class groups on various hyperbolic complexes constructed from arcs and/or curves on the surface.

Big Absences There are a number of interesting topics related to big mapping class groups which are not covered in this survey. Notably, the relation between mapping class groups and dynamics [29], the theory of Teichmüller spaces of infinite-type surfaces (see [73, 81] and the references therein), and the theory of infinite translation surfaces (see for instance [91] and the references therein).

12.2 Preliminaries

In this section we introduce the background material needed for the rest of the chapter.

12.2.1 Surfaces and Their Classification

Throughout this chapter, all surfaces considered will be assumed to be second countable, connected, orientable, and have compact (possibly empty) boundary. If the fundamental group of S is finitely generated, we will say that S is of *finite type*; otherwise, we will say that S is of *infinite type*.

It is well-known that the homeomorphism type of a finite-type surface is determined by the triple (g, p, b), where $g \geq 0$ is the genus, and $p, b \geq 0$ are, respectively, the number of punctures and boundary components of the surface. Because of this fact, we will use the standard notation $S^b_{g,p}$ for the surface specified by these data; as usual, we will drop p and b from the notation whenever they are equal to zero.

There is also a similar classification for infinite-type surfaces [69, 96], in terms of genus, number of boundary components, and the topology of the *space of ends*, which we now define. First, an *exiting sequence* is a sequence $\{U_n\}_{n\in\mathbb{N}}$ of connected open subsets of S with the following properties:

1. $U_n \subset U_m$ whenever $m < n$,
2. U_n is not relatively compact for any $n \in \mathbb{N}$,
3. U_n has compact boundary for all $n \in \mathbb{N}$, and
4. any relatively compact subset of S is disjoint from all but finitely many U_n's.

Two exiting sequences are *equivalent* if every element of the first is eventually contained in some element of the second, and vice versa. We denote by Ends(S) the set of all equivalence classes of exiting sequences of S; an element of Ends(S) is referred to as an *end* of S. The set Ends(S) becomes a topological space, called *the space of ends of* S, by specifying the following basis: given a subset $U \subset S$ with compact boundary, consider the set U^* of all ends represented by an exiting sequence eventually contained in U; the set $\{U^* : U \subset S$ open with compact boundary$\}$ is the desired basis. If U is an open set with compact boundary and $e \in U^*$, then we say that U is a *neighborhood* of the end e.

Given the above basis, it is not difficult to see that Ends(S) is Hausdorff, totally disconnected, and second countable. Moreover, the definition above can be reframed to describe Ends(S) in terms of an inverse limit of compact spaces; in particular, Tychonoff's theorem implies Ends(S) is compact. (For a reference, see [3, Chapter 1].)

Theorem 12.2.1 *For any surface S, the space* Ends(S) *is totally disconnected, second countable, and compact. In particular,* Ends(S) *is homeomorphic to a closed subset of a Cantor set.*

We now proceed to describe the classification of infinite-type surfaces up to homeomorphism. To this end, we will say that an end is *planar* if it admits a neighborhood that is embeddable in the plane; otherwise an end is *non-planar* (or *accumulated by genus*) and every neighborhood of the end has infinite genus. Denote by $\mathrm{Ends}_{np}(S)$ the subspace of Ends(S) consisting of non-planar ends, noting that it is closed in the subspace topology. The following result was proved by Kerékjártó [69] and Richards [96].

Theorem 12.2.2 (Classification, [69, 96]) *Let S_1, S_2 be surfaces, and write g_i and b_i, respectively, for the genus and number of boundary components of S_i. Then $S_1 \cong S_2$ if and only if $g_1 = g_2$, $b_1 = b_2$ and there is a homeomorphism*

$$\mathrm{Ends}(S_1) \to \mathrm{Ends}(S_2)$$

that restricts to a homeomorphism

$$\mathrm{Ends}_{np}(S_1) \to \mathrm{Ends}_{np}(S_2).$$

In light of the above result, an obvious question is: given two closed subsets X, Y of a Cantor set, with $Y \subset X$, can they be realized as the spaces of ends (resp. ends accumulated by genus) of some surface? The following theorem, due to Richards [96], states that the answer is "yes":

Theorem 12.2.3 (Realization, [96]) *Let X, Y be closed subsets of a Cantor set with $Y \subset X$. Then there exists a surface S such that* Ends(S) $\cong X$ *and* $\mathrm{Ends}_{np}(S) \cong Y$.

With the classification and realization theorems at hand, we make a quick note about cardinality: there are exactly $\aleph_0$ many homeomorphism classes of compact surfaces, but $2^{\aleph_0}$ many homeomorphism classes of second-countable surfaces. The second statement follows from a count on the homeomorphism classes of closed subsets of the Cantor set [95]. Interestingly, if one drops the condition of second countability, then there are $2^{\aleph_1}$ many homeomorphism classes of surfaces [49].

12.2.1.1 Some Important Examples

Several infinite-type surfaces have standard names, which makes them easy to identify; these are as follows:

- *The Loch Ness monster surface*: the infinite-genus surface with exactly one end (which is necessarily non-planar).
- *Jacob's ladder surface*: the infinite-genus surface with exactly two ends, both non-planar.
- *The Cantor tree surface*: the planar surface whose space of ends is a Cantor space. Hence, this surface is homeomorphic to a sphere minus a Cantor set.
- *The blooming Cantor tree surface*: the infinite-genus surface whose space of ends is a Cantor space, and such that every end is non-planar.
- *The flute surface*: the planar surface whose space of ends has a unique accumulation point. Hence, this surface is homeomorphic to $\mathbb{C} \setminus \mathbb{Z}$ (and the end space is homeomorphic to $\{0\} \cup \{\frac{1}{n} : n \in \mathbb{N}\}$, viewed as a subset of $\mathbb{R}$).

The Loch Ness monster surface, Jacob's ladder surface, and the blooming Cantor tree surface are shown in Fig. 12.2; the Cantor tree surface can be seen in Fig. 12.4.

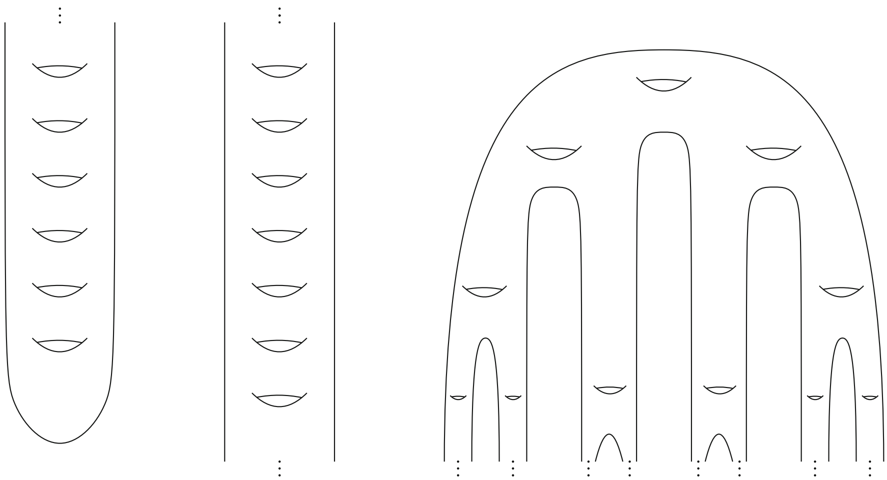


Fig. 12.2 From left to right: Loch Ness monster surface, Jacob's ladder surface, and the blooming Cantor tree surface

To the authors' knowledge, the first two of these names were introduced by Phillips–Sullivan [88], the second two by Ghys [51], and the last by Basmajian [12]. It is worth noting that in [51], Ghys shows that a generic non-compact leaf of 2-dimensional lamination of a metric space is either the plane, the cylinder, or one of the first four surfaces above.

12.2.2 Arcs and Curves

By an *arc* on S we mean the homotopy class of a properly embedded copy of $\mathbb{R}$. Abusing notation, we will not distinguish between arcs and their representatives. Two arcs are *disjoint* if they have disjoint representatives; otherwise we say that they intersect. The *intersection number*, denoted $i(\cdot, \cdot)$, between two arcs is the minimum (possibly infinite) number of points of intersection between representatives.

By a *curve* on S we mean the homotopy class of a simple closed curve on S which does not bound a disk, a punctured disk, or an annulus whose other boundary component is contained in ∂S. As was the case with arcs, we will use the same notation for curves and their representatives. We say that a curve α is *non-separating* if $S \setminus \alpha$ is connected; otherwise we say that α is *separating*. Again, we may talk about when two curves are disjoint or intersect, and define their intersection number as we did with arcs and use the same notation. Note, however, that the intersection number between two curves is necessarily a finite number.

A *multicurve* is a set of pairwise-distinct and pairwise-disjoint curves. A *pants decomposition* is a multicurve P that is maximal with respect to inclusion, and such that any compact set on S is intersected by only finitely many elements of P. As such, the interior of every connected component of the complement of P in S is homeomorphic to a sphere with three points removed, commonly referred to as a *pair of pants*.

12.2.3 Mapping Class Group

Consider the group $\mathrm{Homeo}(S, \partial S)$ of homeomorphisms of S that restrict to the identity on the boundary of S, equipped with the compact-open topology, and the subgroup $\mathrm{Homeo}^+(S, \partial S)$ consisting of those elements that preserve orientation. Let $\mathrm{Homeo}_0(S, \partial S)$ denote the path component of the identity in $\mathrm{Homeo}(S, \partial S)$, and note that $\mathrm{Homeo}_0(S, \partial S) \subset \mathrm{Homeo}^+(S, \partial S)$. The *extended mapping class group* is

$$\mathrm{Map}^{\pm}(S) := \mathrm{Homeo}(S, \partial S)/\, \mathrm{Homeo}_0(S, \partial S),$$

and the *mapping class group* is the subgroup

$$\mathrm{Map}(S) := \mathrm{Homeo}^{+}(S, \partial S)/\,\mathrm{Homeo}_0(S, \partial S).$$

The extended mapping class group becomes a topological group with the quotient topology coming from the compact-open topology on Homeo($S, \partial S$). Combining [39, Theorem 6.4] and [44, Theorem 1], we see that the elements of Map(S) are exactly the isotopy classes of orientation-preserving homeomorphisms of S (see the appendix in [103] for a more detailed discussion).

(Note that is not clear or obvious that the mapping class group is Hausdorff, since—a priori—path components are not closed subsets. Being Hausdorff is a condition that is often required in the definition of topological group. We will deal with this in Sect. 12.4.)

12.2.4 Several Natural Subgroups

Throughout the survey, several natural subgroups of mapping class groups will appear: we provide their definition here.

12.2.4.1 Pure Mapping Class Group

Observe that every homeomorphism of S induces a type-preserving homeomorphism of its space of ends. In other words, there is a natural map

$$\mathrm{Homeo}^{+}(S, \partial S) \to \mathrm{Homeo}(\mathrm{Ends}(S), \mathrm{Ends}_{np}(S)), \tag{12.1}$$

where the latter group denotes the subgroup of the homeomorphism group of Ends(S) whose elements preserve $\mathrm{Ends}_{np}(S)$ setwise. One checks this is a continuous homomorphism when $\mathrm{Homeo}(\mathrm{Ends}(S), \mathrm{Ends}_{np}(S))$ is equipped with the (subgroup topology coming from the) compact-open topology.

Richards's proof of the classification of surfaces can readily be adapted to establish the surjectivity of the homomorphism given in (12.1). As an isotopy fixes every end of a surface, the homomorphism (12.1) factors through Map(S) yielding a surjective homomorphism

$$\mathrm{Map}(S) \to \mathrm{Homeo}(\mathrm{Ends}(S), \mathrm{Ends}_{np}(S)). \tag{12.2}$$

The *pure mapping class group*, written PMap(S), is the kernel of the above homomorphism. In particular, we have a short exact sequence

$$1 \to \mathrm{PMap}(S) \to \mathrm{Map}(S) \to \mathrm{Homeo}(\mathrm{Ends}(S), \mathrm{Ends}_{np}(S)) \to 1 \tag{12.3}$$

It is worth noting that by Stone's representation theorem, there is a one-to-one correspondence (or, technically, a contravariant functor) between closed subsets of the Cantor set and countable Boolean algebras. There is a large amount of literature about automorphism groups of boolean algebras, which can be translated to homeomorphism groups of end spaces of surfaces (and vice versa).

We also note that, by the definition of the mapping class groups, $\mathrm{Map}(S) = \mathrm{PMap}(S)$ if and only if either $|\mathrm{Ends}(S)| \leq 1$ or $|\mathrm{Ends}(S)| = 2$ and S has exactly one planar end.

12.2.4.2 Compactly Supported Mapping Class Group

An element of $\mathrm{Map}(S)$ is *compactly supported* if it has a representative homeomorphism that is the identity outside of a compact subset. The *compactly supported mapping class group*, denoted $\mathrm{Map}_c(S)$, is the subgroup of $\mathrm{Map}(S)$ consisting of the compactly supported elements. Observe that, in fact, $\mathrm{Map}_c(S) < \mathrm{PMap}(S)$.

We say a compact subsurface X of a surface S is *essential* if no component of $S \smallsetminus X$ is a disk or annulus. If X is an essential compact subsurface of S, then $\mathrm{Map}(X) < \mathrm{Map}_c(S)$. Note that for any two essential compact subsurfaces X and Y of S, we have $\mathrm{Map}(X) < \mathrm{Map}(Y)$ whenever $X \subset Y$. Moreover, the union of all compact subsurfaces of S is equal to S; hence, we have:

Proposition 12.2.4 *For any surface S,*

$$\mathrm{Map}_c(S) = \varinjlim \mathrm{Map}(X),$$

where the direct limit is taken over all essential compact subsurfaces X of S, ordered by inclusion.

12.2.4.3 Torelli Group

Observe that every element of $\mathrm{Map}(S)$ acts on the homology group $H_1(S, \mathbb{Z})$ by automorphisms. In other words, there is a homomorphism

$$\mathrm{Map}(S) \to \mathrm{Aut}(H_1(S, \mathbb{Z})). \tag{12.4}$$

We remark that if S is a finite-type surface of genus g and with at most one puncture, then $\mathrm{Aut}(H_1(S, \mathbb{Z}))$ is isomorphic to the symplectic group $\mathrm{Sp}(2g, \mathbb{Z})$, although this is not true in general. The *Torelli group* $\mathcal{I}(S)$ is the kernel of the homomorphism (12.4); in other words, it is the subgroup of $\mathrm{Map}(S)$ whose elements act trivially on homology. Observe that $\mathcal{I}(S)$ is a subgroup of $\mathrm{PMap}(S)$.

12.2.5 Modular Groups

Naturally associated to a Riemann surface is the subgroup QC(X) of Homeo$^+(X)$ consisting of the quasi-conformal homeomorphisms. The image of QC(X) in Map(X), denoted Mod(X), is commonly referred to as either the *Teichmüller modular group* of X or the *quasi-conformal mapping class group* of X. In the case that X is of finite topological type, Mod(X) and Map(X) agree and are routinely interchanged in the literature; however, this fails to be the case for infinite-type surfaces.

In the infinite-type setting, unlike mapping class groups, modular groups have a long history of being studied, especially from the theory of Riemann surfaces and Teichmüller theory. As such, discussing the modular group would be a survey in-of-itself and we will make no further mention of it. But, we note that there are surely many interesting questions and problems related to how Mod(X) sits as a subgroup of Map(S), where X is a Riemann surface homeomorphic to an infinite-type surface S.

12.3 Two Important Results

In this section we present two results that underpin a large number of the topics discussed in latter sections. Throughout this section, every surface is assumed to have empty boundary.

12.3.1 Alexander Method

As mentioned in the introduction, Map(S) inherits a natural topology when viewed as a quotient of Homeo$^+(S)$, equipped with the compact-open topology. It is standard to require that a topological group be Hausdorff, and so it is not immediately obvious that Map(S) in this topology is in fact a topological group. However, we can use the extension of Alexander's method to infinite-type surfaces given in [61]. Here, we state the corollary we require:

Theorem 12.3.1 ([61, Corollary 1.2]) *Let S be an infinite-type surface. If $f \in$ Homeo$^+(S)$ fixes the isotopy class of every simple closed curve, then f is isotopic to the identity.*

Theorem 12.3.1 can used to separate the identity from any other element in Map(S) by an open set and, for topological groups, this is enough to guarantee the group is Hausdorff; hence, Map(S) is a topological group.

12.3.2 Automorphisms of the Curve Graph

The *curve graph* $\mathcal{C}(S)$ of S is the simplicial graph whose vertex set is the set of curves on S, and where two vertices are adjacent in $\mathcal{C}(S)$ if and only if the corresponding curves on S are disjoint. From now on we will not distinguish between vertices of $\mathcal{C}(S)$ and the curves they represent.

Observe that $\mathrm{Map}^{\pm}(S)$ acts on $\mathcal{C}(S)$ by simplicial automorphisms. In fact, the combined work of Ivanov [63], Korkmaz [71], and Luo [74] shows that, with the exception of the twice-holed torus, there are no other automorphisms of $\mathcal{C}(S)$ when S is of finite type. In the infinite-type setting, the analogous result was proved independently by Hernández–Morales–Valdez [60] and Bavard–Dowdall–Rafi [17]:

Theorem 12.3.2 *If S is an infinite-type surface, then the group of simplicial automorphisms of $\mathcal{C}(S)$ is naturally isomorphic to* $\mathrm{Map}^{\pm}(S)$.

Note that, in particular, Theorem 12.3.1 is required to show that the action of $\mathrm{Map}(S)$ on $\mathcal{C}(S)$ has no kernel.

12.4 Topological Aspects

We will see in this section that big mapping class groups are interesting topological groups—a divergence from the finite-type setting. This offers exciting new connections for mapping class groups, some of which we explore below.

It follows from the Alexander method for finite-type surfaces (see [43, Proposition 2.8]) that $\mathrm{Map}(S)$ is discrete when S is of finite-type. However, this is far from true for big mapping class groups: to see this, let S be an infinite-type surface and let $\{c_n\}_{n\in\mathbb{N}}$ be a sequence of simple closed curves such that, for every compact subset K of S, there is an integer N such that $K \cap c_n = \emptyset$ for all $n > N$. If T_n is the Dehn twist about c_n, then the sequence $\{T_n\}_{n\in\mathbb{N}}$ limits to the identity in $\mathrm{Map}(S)$.

12.4.1 The Permutation Topology

In order to investigate the topology of $\mathrm{Map}(S)$ in more depth, it is convenient to have a more combinatorial description of its topology.

Let Γ be a simplicial graph with a countable set of vertices, and let $\mathrm{Aut}(\Gamma)$ be the group of simplicial automorphisms of Γ. Given a subset A of Γ, let

$$U(A) := \{g \in \mathrm{Aut}(\Gamma) \mid g(a) = a \text{ for all } a \in A\}.$$

Then $\mathrm{Aut}(\Gamma)$ may be endowed with a natural topology, called the *permutation topology*, defined by declaring the $\mathrm{Aut}(\Gamma)$-translates of $U(A)$, for every finite subset

A of Γ, a basis for the topology. Equivalently, the permutation topology is the coarsest topology in which, for every $v \in \mathcal{C}(S)$, the function $\omega_v\colon \operatorname{Aut}(\Gamma) \to \Gamma$ defined by $\omega_v(g) = g(v)$ is continuous.

With respect to the permutation topology, Aut(Γ) becomes a second-countable (and in particular, separable) topological group. Moreover, it is a standard exercise in descriptive set theory texts to show that Aut(Γ) supports a complete metric (which—usually—fails to be Aut(Γ)-invariant).

In particular, Aut(Γ) is an example of a *Polish group*, that is, a separable and completely metrizable group. Polish groups are a well-studied class of groups and we will make use of their theory.

For an infinite-type surface S with empty boundary, let $\Gamma = \mathcal{C}(S)$, then, by Theorem 12.3.2, we can identify $\operatorname{Map}^{\pm}(S)$ with Aut(Γ) and equip $\operatorname{Map}^{\pm}(S)$ with the associated permutation topology. It is an exercise in definitions and the Alexander method to show that this permutation topology agrees with the compact-open topology. Recall that a G_δ subset of a topological space is a subset that can be written as the intersection of countably many open sets (note that in a metrizable space, every closed set is a G_δ subset). As a consequence of this discussion, we have:

Proposition 12.4.1 *Let S be a infinite-type surface, possibly with non-empty boundary. Then,* $\operatorname{Map}^{\pm}(S)$ *and all its G_δ-subgroups, including* Map(S) *and* PMap(S)*, are Polish.*

Note that, unlike the preceding discussion, Proposition 12.4.1 does not require S to have empty boundary: this is because the mapping class group of a bordered surface can be embedded in a borderless surface as a closed subgroup.

12.4.2 Basic Properties

Now that we have an understanding of the topology of mapping class groups, we can investigate their basic properties. First, note that the sets in the basis defined above for Map(S) are in fact clopen and hence mapping class groups are zero-dimensional.

Now, let S be of infinite type. Observe that if $A \subset \mathcal{C}(S)$ and $c \in \mathcal{C}(S)$ such that $c \cap a = \varnothing$ for all $a \in A$, then the sequence $\{T_c^n\}_{n\in\mathbb{N}}$ has no limit point and is contained in $U(A)$; in particular, again by homogeneity, we can conclude that every compact subset of Map(S) is nowhere dense. This also establishes the weaker fact that Map(S) fails to be locally compact. Moreover, as a Polish space cannot be the countable union of nowhere dense subsets, we can conclude that Map(S) is not compactly generated.[1] Lastly, the Alexandrov–Urysohn Theorem

[1]There are two standard meanings for compactly generated, one algebraic and one topological. For clarity, we are referring to the algebraic setting: specifically, we mean that if a set $\mathcal{S}$ generates Map(S), as a group, then $\mathcal{S}$ cannot be compact.

(see [66, Theorem 7.7]) establishes $\mathbb{N}^{\mathbb{N}}$ as the unique space, up to homeomorphism, that is non-empty, Polish, zero-dimensional, and in which every compact subset has non-empty interior; hence, $\mathrm{Map}(S)$ is homeomorphic to $\mathbb{N}^{\mathbb{N}}$. We record these observations in the following theorem:

Theorem 12.4.2 *For every infinite-type surface S,*

(1) $\mathrm{Map}(S)$ *is not locally compact,*
(2) $\mathrm{Map}(S)$ *is not compactly generated,*
(3) $\mathrm{Map}(S)$ *is homeomorphic to the Baire space* $\mathbb{N}^{\mathbb{N}}$ *(which in turn is homeomorphic to* $\mathbb{R} \smallsetminus \mathbb{Q}$*).*

Theorem 12.4.2 establishes big mapping class groups as large topological groups. It is often the case that the tools developed for studying finitely-generated groups have natural analogs in the setting of locally-compact compactly-generated topological groups. The failure of big mapping class groups to fall into this category will generally complicate matters, but simultaneously offers big mapping class groups as potential fertile ground for applying the rapidly developing and exciting theory and tools of non-locally-compact topological groups. We will see this below when we discuss the geometry of mapping class groups.

12.4.3 Topological Generation

Since big mapping class groups are separable, they are necessarily topologically generated by a countable set, that is, there exists a countable set that generates a dense subgroup. The goal of this subsection is to produce such a topological generating set whose elements are relatively simple. Recall that for a connected finite-type surface S, its pure mapping class group $\mathrm{PMap}(S)$ is generated by—a finite set of—Dehn twists. In order to generate the full mapping class group, it is necessary to add half-twists, which correspond to transpositions in the symmetric group isomorphic to $\mathrm{Map}(S)/\,\mathrm{PMap}(S)$.

In the infinite-type setting, Eq. (12.3) tells us that $\mathrm{Map}(S)/\,\mathrm{PMap}(S)$ is isomorphic to $\mathrm{Homeo}(\mathrm{Ends}(S), \mathrm{Ends}_{np}(S))$, so in order to understand topological generating sets for $\mathrm{Map}(S)$, we would also have to do so for the latter homeomorphism groups; this will take us too far afield and so we will focus on generating $\mathrm{PMap}(S)$.

Using the fact that the mapping class group of a compact surface is generated by Dehn twists, we see that the group $\mathrm{Map}_c(S)$ consisting of compactly supported mapping classes is generated by Dehn twists. It is natural to ask if the closure of this group is all of $\mathrm{PMap}(S)$. The next result, proved by Patel and the second author in [87], shows that this is true only in certain cases:

Theorem 12.4.3 ([87]) *The set of Dehn twists topologically generate* $\mathrm{PMap}(S)$ *if and only if S has at most one non-planar end.*

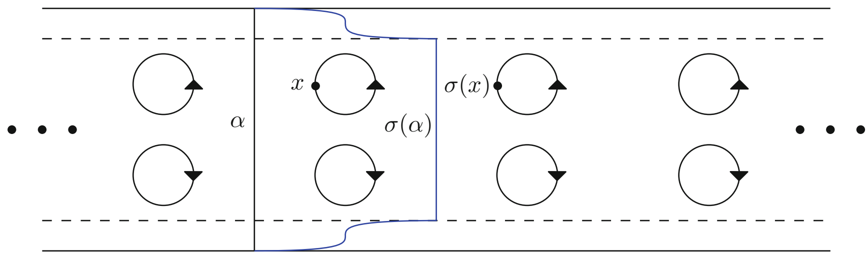


Fig. 12.3 The circles are identified vertically to obtain Σ

The only impediment to Dehn twists topologically generating is the existence of a homeomorphism $f: S \to S$ and a separating curve γ non-trivial in homology such that $f(\gamma) \cap \gamma = \varnothing$. As it turns out, this can only be done—while fixing the ends—if there are at least two non-planar ends. Let us give an example of such a homeomorphism, known as a *handle shift*, which was introduced in [87].

For $n \in \mathbb{Z}$, let $B_n^\pm$ be the open Euclidean disks of radius 1 in $\mathbb{R}^2$ centered at $(n, \pm 2)$, respectively. Let Σ be the (infinite-genus) surface obtained from $\mathbb{R} \times [-4, 4]$ by, for each $n \in \mathbb{Z}$, removing $B_n^\pm$ and identifying ∂B_n^+ and ∂B_n^- via an orientation-reversing homeomorphism. Up to isotopy, there is a unique homeomorphism $\sigma: \Sigma \to \Sigma$ determined by requiring

1. $\sigma((x, y)) = (x + 1, y)$ for all $(x, y) \in \Sigma$ with $|y| \leq 3$, and
2. $\sigma((x, \pm 4)) = (x, \pm 4)$ for all $x \in \mathbb{R}$.

See Fig. 12.3 to see the behavior of σ on a vertical arc. Now, for an infinite-genus surface S, we say a homeomorphism $h: S \to S$ is a *handle shift* if there exists a proper embedding $\iota: \Sigma \to S$ such that

$$h = \begin{cases} \iota \circ \sigma \circ \iota^{-1}(x) & x \in \iota(\Sigma) \\ x & \text{otherwise} \end{cases}$$

We will also refer to a mapping class containing a handle shift as a handle shift itself. Identifying Σ with its image under ι, we say that h is *supported* on Σ. Since the embedding ι is required to be proper, there is an induced map $\iota_\infty: \text{Ends}(\Sigma) \to \text{Ends}(S)$. It follows that h has an *attracting* and a *repelling* end, which we label h^+ and h^- respectively, and that satisfy

$$\lim_{n \to \pm\infty} h^n(x) = h^\pm$$

for every x in the interior of Σ (the limit is formally taken in the Freudenthal compactification of S). Note that if h_1 and h_2 are isotopic handle shifts, then $h_1^\pm = h_2^\pm$; therefore, we can talk about the attracting and repelling ends of a mapping class associated to a handle shift.

Let h be a handle shift supported on Σ in an infinite-genus surface S with at least two non-planar ends and such that $h^+ \neq h^-$. Now observe that if we take a separating curve γ that is non-trivial in homology and such that $\gamma \cap \Sigma$ is connected and isotopic to a vertical arc, then γ is non-trivial in homology, γ is not homotopic to $h(\gamma)$, and $i(\gamma, h(\gamma)) = 0$. As described in [87], these conditions guarantee that h is not a limit of compactly supported mapping classes.

It was shown in [87] that the set of Dehn twists together with the set of handle shifts topologically generate $\mathrm{PMap}(S)$. But, the set of handle shifts is uncountable and we want a countable dense subset. As a corollary of a—much stronger—result in [11], we can reduce to a countable collection:

Theorem 12.4.4 ([11]) *If S is an infinite-genus surface with at least two non-planar ends, then there exists a countable set consisting of Dehn twists and handle shifts topologically generating* $\mathrm{PMap}(S)$.

The handle shifts obtained from [11] will pairwise commute; however, for a weaker, but direct version, it would suffice to choose a countable dense subset $\{(e_n^+, e_n^-)\}_{n\in\mathbb{N}}$ in $\mathrm{Ends}_{np}(S) \times \mathrm{Ends}_{np}(S)$ and handle shifts $h_n \in \mathrm{PMap}(S)$ such that $h_n^{\pm} = e_n^{\pm}$. It can be checked that these handle shifts along with Dehn twists will topologically generate $\mathrm{PMap}(S)$.

Adapting an argument presented in [43, Theorem 7.16] showing that the mapping class group of a finite-type surface is generated by torsion elements, Afton–Freedman–Lanier–Yin [1] observed:

Theorem 12.4.5 ([1]) *If S is an infinite-genus surface, then* $\mathrm{PMap}(S)$ *is topologically generated by handle shifts.*

12.4.3.1 Torelli Group

As noted previously, $\mathcal{I}(S)$ is contained in $\mathrm{PMap}(S)$; moreover, handle shifts act non-trivially on homology and hence $\mathcal{I}(S)$ contains no handle shifts. This is enough to imply that $\mathcal{I}(S)$ is contained in the closure of $\mathrm{Map}_c(S)$ (this follows from Corollary 12.5.18 below). Letting $\mathcal{I}_c(S)$ denote the intersection $\mathcal{I}(S) \cap \mathrm{Map}_c(S)$, it is natural to ask if the closure of $\mathcal{I}_c(S)$ is all of $\mathcal{I}(S)$. The answer is yes:

Theorem 12.4.6 ([10]) *If S is an infinite-type surface, then $\mathcal{I}_c(S)$ is dense in $\mathcal{I}(S)$.*

Combining results of Birman [23], Powell [90] and an argument due to Justin Malestein, the above theorem implies the following (see [10] for details and definitions):

Theorem 12.4.7 ([10]) *Let S be any surface of infinite type. Then $\mathcal{I}(S)$ is topologically generated by separating twists and bounding-pair maps.*

12.4.4 Coarse Boundedness

Before we begin, we note that all the general theory about Polish groups discussed here is developed in Rosendal's forthcoming book [99].

The theories of finitely-generated groups and locally-compact compactly-generated topological groups have many analogies, especially from the viewpoint of geometric group theory. This is naturally due to compactness being a natural generalization of finiteness; however, as noticed by Rosendal, there is a weaker condition on topological groups that allows one to still capture many of the key aspects of the theory of locally-compact compactly-generated groups.

The key observations is to note that a compact subset of a (pseudo-)metric space always has finite diameter; it turns out this is the property to focus on. In a Polish group G, a subset A of G is *coarsely bounded*, or *CB* for short, if it has finite diameter in every continuous pseudo-metric on G (in fact, it is sufficient to only consider left-invariant continuous pseudo-metrics). A Polish group is *coarsely bounded*, or *CB*, if it is coarsely bounded as a subset; it is *locally coarsely bounded*, or *locally CB*, if there exists a coarsely-bounded open neighborhood of the identity; it is *CB generated* if there exists a coarsely bounded set algebraically generating the group.

One should naturally think of CB as a generalization of compact, locally CB as a generalization of locally compact, and CB generated as a generalization of compactly generated. Conveniently, every CB generated Polish group is locally CB [99, Theorem 2.30] (note: it is not the case that every compactly-generated group is locally compact, e.g. $(\mathbb{Q}, +)$ is compactly generated but not locally compact).

From the point of view of this survey, the main result of the theory of CB-generated Polish groups is that, up to quasi-isometry, they have a well-defined metric. In particular, CB-generated Polish groups have a well-defined geometry and they can be studied through the lens of geometric group theory. Let us now describe this result.

A left-invariant continuous pseudo-metric d is *maximal* if for any other left-invariant continuous pseudo-metric d' there exits constants $K, L \geq 0$ such that $d' < K \cdot d + L$. In particular, up to quasi-isometry, if a maximal pseudo-metric exists, then it is unique. Before stating the theorem, a subset of a Polish space is *analytic* if it is the continuous image of a Polish space. Now, combining pieces of Theorem 1.2, Proposition 2.52, Theorem 2.53, and Example 2.54 from [99], we have:

Theorem 12.4.8 ([99]) *Let G be a CB-generated Polish group. Then:*

(1) G admits a left-invariant continuous maximal metric d.
(2) G has an analytic symmetric coarsely-bounded generated set; moreover, G equipped with the word metric associated to any such generating set is quasi-isometric to (G, d).

Note that the metric topology associated to a word metric is always discrete and hence cannot be continuous on a non-discrete topological group. However, the above theorem tells us that (non-continuous) word metrics capture the geometry of the group.

In recent work, Mann–Rafi [77] classify the CB, locally CB, and CB-generated mapping class groups. The most general version of their result is a bit technical to state, so we will state a specific case that captures the main flavor. It is a classical result of Mazurkiewicz and Sierpinski [80] that every countable compact Hausdorff topological space is homeomorphic to an ordinal space of the form $\omega^\alpha n + 1$, where α is a countable ordinal, n is a natural number, and ω is the first infinite ordinal.

Theorem 12.4.9 ([77]) *Let S be an infinite-type surface so that either every end of S is planar or every end of S is non-planar. If the end space of S is countable and homeomorphic to $\omega^\alpha n + 1$, then*

(1) $\mathrm{Map}(S)$ *is CB if and only if* $n = 1$.
(2) If $n \geq 2$ and α is a successor ordinal, then $\mathrm{Map}(S)$ *is CB generated, but not CB.*
(3) If $n \geq 2$ and α is a limit ordinal, then $\mathrm{Map}(S)$ *is locally CB, but not CB generated.*

The full statement of Mann–Rafi's theorem involves generalizing the trichotomy above to uncountable end spaces; they do this by introducing a partial order on the ends. We encourage the interested reader to see their paper for details; we believe the various cases described will be essential for researchers interested in proving results about all big mapping class groups.

For examples, the mapping class group of the Loch Ness monster surface is CB as is the mapping class group of the flute surface. Also, though it does not fit into the countable version of the Mann–Rafi theorem given above, the mapping class group of the Cantor tree surface is CB. For $n \in \mathbb{N}$, let Ω_n denote the infinite-genus surface with n ends, all of which are non-planar. If $n \geq 2$, then $\mathrm{Map}(\Omega_n)$ is CB generated, but not CB; in particular, $\mathrm{Map}(\Omega_n)$ is not quasi-isometric to $\mathrm{Map}(\Omega_1)$ if $n \geq 2$. Therefore, we ask:

Question 12.4.10 *Are* $\mathrm{Map}(\Omega_n)$ *and* $\mathrm{Map}(\Omega_m)$ *quasi-isometric if and only if $n = m$?*

As a complementary question, we propose:

Question 12.4.11 *Are there computable quasi-isometry invariants of CB-generated big mapping class groups (e.g. geometric rank)?*

12.4.5 Automatic Continuity

A topological group G has the *automatic continuity property* if every abstract group homomorphism from G to a separable topological group is continuous. There is a beautiful history to studying automatic continuity given in [98]; however, we only discuss several relevant examples (and non-examples).

For a non-example, consider the following: the real line $\mathbb{R}$ and the real plane $\mathbb{R}^2$, each equipped with the standard Euclidean topology and the group operation of (vector) addition, are isomorphic as groups. To see this, observe that both $\mathbb{R}$ and $\mathbb{R}^2$ are infinite-dimensional vector spaces over the rationals $\mathbb{Q}$ with bases of cardinality $2^{\aleph_0}$ and hence they are isomorphic. However, $\mathbb{R}$ and $\mathbb{R}^2$ are not homeomorphic and hence this group isomorphism cannot be continuous.

For examples, none of which are trivial, the homeomorphism group of the Cantor set [67] as well as the homeomorphism group of any closed manifold [75, 97] has the automatic continuity property. The automatic continuity property for homeomorphism groups (and some diffeomorphism groups) has been key to recent developments in approaches to the dimension growth question of Ghys [50] regarding actions of infinite groups on compact manifolds (e.g. Chen–Mann [33], Hurtado [62]). The application of automatic continuity in understanding the rigidity of homeomorphism groups of compact manifolds motivates us to ask about automatic continuity in mapping class groups, where there are also open rigidity questions (see Sect. 12.5).

Question 12.4.12 *Classify the surfaces S for which the groups* $\mathrm{Homeo}(S)$ *and/or* $\mathrm{Map}(S)$ *have the automatic continuity property.*

Recently, building on her previous work [75], Mann proved that the homeomorphism group of any manifold that can be realized as the interior of a compact manifold with boundary has the automatic continuity property [76]. In the same article, Mann gave the first examples of infinite-type surfaces (e.g. the sphere minus a Cantor set) whose homeomorphism groups have the automatic continuity property. Mann's result actually shows these groups have a stronger property (they are Steinhaus), which passes to quotients and hence yields:

Theorem 12.4.13 ([76, Corollary 2.1]) *Let S be an infinite-type surface of finite genus whose space of ends is of the form $C \sqcup F$, where C is a Cantor space and F is a finite discrete space. Then,* $\mathrm{Map}(S)$ *has the automatic continuity property.*

In [76, Example 2.3], Mann also gives an example of an infinite-type surface whose homeomorphism group and mapping class group do not have the automatic continuity property.

All the arguments establishing automatic continuity for the homeomorphism groups mentioned above rely on the same core technique, which unfortunately does not readily extend to non-compact surfaces with infinite-genus nor finite-genus with non-perfect end space.

12.5 Algebraic Aspects

12.5.1 Algebraic Rigidity

In this subsection, all surfaces are assumed to have empty boundary. A classical result of Ivanov [64] asserts that, with several well-understood exceptions, every automorphism of the mapping class group of a finite-type surface S is induced by a homeomorphism of S. Ivanov gave a simplified proof of this result using the curve complex in [63]; however, in this case, he assumes the underlying surface has genus at least two. This simplified proof was adapted to the remaining cases by Korkmaz [71] and Luo [74] independently. In the infinite-type setting, the analogous result was established by Bavard–Dowdall–Rafi [17]; namely, one has:

Theorem 12.5.1 ([17]) *For any infinite-type surface S,*

$$\mathrm{Aut}(\mathrm{Map}(S)) \cong \mathrm{Map}^{\pm}(S).$$

The idea of the proof of Theorem 12.5.1 is similar in spirit to that of Ivanov, adapted to the context of infinite-type surfaces. First, the authors prove that an element of $\mathrm{Map}(S)$ is supported on a finite-type subsurface of S if and only if its conjugacy class is countable, and from this they obtain an algebraic characterization of Dehn twists, similar to Ivanov's original one, which is preserved by automorphisms. As a consequence, any given automorphism of $\mathrm{Map}(S)$ induces a simplicial automorphism of the curve complex $\mathcal{C}(S)$ which in turn, by Theorem 12.3.2, is induced by an element of $\mathrm{Map}^{\pm}(S)$. At this point, the mapping class obtained this way coincides with the original automorphism on every Dehn twist, from which one quickly deduces that they are equal.

12.5.1.1 Injective and Surjective Homomorphisms

Ivanov's theorem gave rise to a large number of stronger rigidity results about mapping class groups. For instance, a result of Ivanov–McCarthy [65] asserts that mapping class groups of surfaces of genus at least three are *co-Hopfian*, that is, every injective endomorphism is an automorphism. Hence, every injective endomorphism is induced by a homeomorphism of the underlying surface. The analog in the infinite-type setting is not known:

Question 12.5.2 *Are mapping class groups of infinite-type surfaces co-Hopfian?*

One of the main hurdles in this direction is that, for infinite-type surfaces, simplicial injections of the curve complex into itself need not come from mapping classes, in stark contrast to the case of finite-type surfaces (see [58] for the strongest result of this type). An example of this, for surfaces of infinite genus, may be found in [59, Lemma 5.3]. We now present another instance of this phenomenon, which can be easily generalized to other punctured surfaces:

Example (Non-surjective Simplicial Injections Between Curve Graphs) Let S be the flute surface. As such, we may realize S as the surface obtained by removing from $\mathbb{S}^2$ a convergent sequence together with its limit point.

Fix a hyperbolic structure on S, and realize every simple closed curve on S by its unique geodesic representative. Since there are only countably many simple closed curves on S, we may pick a point p in the complement of the union of all the simple closed geodesics. Therefore we obtain a map $h : \mathcal{C}(S) \to \mathcal{C}(S \setminus \{p\})$ which is easily seen to be injective, since two curves that are disjoint on S remain disjoint after puncturing. Finally, observe that $S \setminus \{p\}$ is homeomorphic to S, but that the map h is not induced by a homeomorphism, as it is not surjective. This finishes the example.

With respect to surjective homomorphisms, a group is *Hopfian* if every surjective endomorphism is an automorphism. It is an exercise to show that every finitely-generated residually-finite group is Hopfian; hence, mapping class groups of finite-type surfaces are Hopfian. It is therefore natural to ask if big mapping class groups are Hopfian. But, we quickly find a counterexample:

Example (Non-Hopfian Mapping Class Group) Let E be a closed subset of the Cantor set such that the set E' of accumulation points of E satisfies $E' \neq E$ and E' is homeomorphic to E. For example, the ordinal space $\omega^\omega + 1$ has this property. Embed E into the 2-sphere $\mathbb{S}^2$. We then have that the embedding $\mathbb{S}^2 \smallsetminus E \hookrightarrow \mathbb{S}^2 \smallsetminus E'$ induces a forgetful homomorphism $\mathrm{Map}(\mathbb{S}^2 \smallsetminus E) \to \mathrm{Map}(\mathbb{S}^2 \smallsetminus E')$ that is surjective, but not injective. Now, $\mathbb{S}^2 \smallsetminus E$ is homeomorphic to $\mathbb{S}^2 \smallsetminus E'$ and hence we see there exists a surjective endomorphism of $\mathrm{Map}(\mathbb{S}^2 \smallsetminus E)$ that fails to be an automorphism. Note that the forgetful map exists only because $E \smallsetminus E'$—the set of isolated points of E—is invariant under the action of $\mathrm{Map}(\mathbb{S}^2 \smallsetminus E)$.

Question 12.5.3 *If a surjective endomorphism of a mapping class group fails to be an automorphism, is it necessarily a forgetful homomorphism?*

12.5.1.2 General Homomorphisms

A result of Souto and the first author [7] describes all non-trivial homomorphisms $\mathrm{PMap}(S) \to \mathrm{PMap}(S')$, where the genus of S is at least six and the genus of S' is less than twice the genus of S, showing that they arise as combinations of *subsurface inclusions*, *forgetting punctures*, and *deleting boundary components*. A homomorphism between mapping class groups that comes from a *manipulation at the level of the underlying surfaces* is called *geometric*.

Other than Theorem 12.5.1, there are no results of this kind in the context of infinite-type surfaces. In fact, as a consequence of Theorem 12.5.15 below, if S has at least two non-planar ends then there are non-geometric endomorphisms of $\mathrm{PMap}(S)$. However, all these examples factor through the (non-trivial) abelianization of $\mathrm{PMap}(S)$. An ambitious question is to ask if this is the only way to produce non-geometric endomorphisms:

Question 12.5.4 *Let S be a surface of infinite type with no boundary. Does every non-geometric endomorphism of* $\mathrm{PMap}(S)$ *factor through its abelianization?*

A much more humble question to which we do not know the answer (although we expect it to be negative) is:

Question 12.5.5 *Let S be Jacob's ladder surface and let S' be the Loch Ness monster. Are there any homomorphisms* $\mathrm{PMap}(S) \to \mathrm{PMap}(S')$ *with non-abelian image?*

12.5.1.3 Rigidity of Subgroups

In fact, the aforementioned result of Ivanov [63] applies to injections between finite-index subgroups of mapping class groups. In other words, it asserts that the *abstract commensurator* $\mathrm{Comm}(\mathrm{Map}(S))$ of $\mathrm{Map}(S)$ is equal to $\mathrm{Map}^{\pm}(S)$, provided the genus of S is large enough. For infinite-type surfaces, the analog is due to Bavard–Dowdall–Rafi [17] (the proof is the same as for Theorem 12.5.1):

Theorem 12.5.6 ([17]) *For any infinite-type surface S,*

$$\mathrm{Comm}(\mathrm{Map}(S)) \cong \mathrm{Map}^{\pm}(S).$$

In [10], it is shown that $\mathcal{I}(S)$ is also algebraically rigid; more concretely:

Theorem 12.5.7 ([10]) *For any infinite-type surface S,*

$$\mathrm{Aut}(\mathcal{I}(S)) \cong \mathrm{Comm}(\mathcal{I}(S)) \cong \mathrm{Map}^{\pm}(S).$$

The equivalent statement for finite-type surfaces was proved by Farb–Ivanov [42] for automorphisms, and by Brendle–Margalit [26] for commensurations.

We remark that it is not known whether $\mathcal{I}(S)$ has any finite-index subgroups at all; hence we ask:

Question 12.5.8 *Does $\mathcal{I}(S)$ have any proper finite-index subgroups?*

Note that if the answer to the above question were negative, then $\mathrm{Comm}(\mathcal{I}(S))$ would be equal to $\mathrm{Aut}(\mathcal{I}(S))$ *a priori*.

Finally, we should mention a recent theorem of Brendle–Margalit [27] (for closed surfaces) and McLeay [82] (for surfaces with punctures) which vastly generalizes the theorems above, proving that every normal subgroup which contains elements of *sufficiently small support* has the extended mapping class group as its automorphism and abstract commensurator group. In the setting of infinite-type surfaces one expects fewer necessary conditions, as the following result of McLeay [83] shows:

Theorem 12.5.9 ([83]) *Let S be the Cantor tree surface. If N is any normal subgroup of* $\mathrm{Map}(S)$*, then*

$$\mathrm{Aut}(N) \cong \mathrm{Map}^{\pm}(S).$$

Though not directly a rigidity result, we finish this subsection by recalling a result of Lanier–Loving [72] that fits with the discussion:

Theorem 12.5.10 ([72]) *If S is an infinite-type surface, then every normal subgroup has trivial center.*

12.5.2 Abelianization

A classical result of Powell [90], building up on previous work of Mumford [86] and Birman [23], shows that the abelianization of the mapping class group of a closed surface of genus at least three is trivial. Moreover, the lantern relation can be used to establish the same result for all finite-type surfaces:

Theorem 12.5.11 (See [43, Theorem 5.2]) *Let S be a finite-type surface of genus at least* 3. *Then* $\mathrm{PMap}(S)$ *has trivial abelianization.*

By Proposition 12.2.4, $\mathrm{Map}_c(S)$ is a direct limit of finite-type mapping class groups, and hence:

Corollary 12.5.12 *Let S be a surface of genus at least 3. Then* $\mathrm{Map}_c(S)$ *has trivial abelianization.*

We would like to promote the above corollary to a statement about the pure mapping class group, and here is one instance where automatic continuity is incredibly useful. Indeed, a result of Dudley [38] asserts that if G is a Polish group, then any homomorphism $G \to \mathbb{Z}$ is continuous. Combining this with Corollary 12.5.12, we have:

Theorem 12.5.13 *Let S be a surface of genus at least 3. Then, every homomorphism*

$$\overline{\mathrm{Map}_c(S)} \to \mathbb{Z}$$

is trivial. In other words,

$$H^1(\overline{\mathrm{Map}_c(S)}, \mathbb{Z}) = \{1\}.$$

In light of Theorem 12.4.3 above, this has the following consequence:

Corollary 12.5.14 *Let S be a surface with at most one non-planar end. Then* $H^1(\mathrm{PMap}(S), \mathbb{Z}) = \{1\}$.

However, in [11] it was shown that the situation for general infinite-type surfaces is rather different. Namely, one has:

Theorem 12.5.15 ([11]) *Let S be a surface of genus at least two, and let $\hat{S}$ denote the result of filling every planar end of S. Then*

$$H^1(\mathrm{PMap}(S), \mathbb{Z}) \cong H_1^{sep}(\hat{S}, \mathbb{Z}),$$

where the latter group is the subgroup of $H_1(\hat{S}, \mathbb{Z})$ generated by homology classes with separating representatives.

In particular, $H^1(\mathrm{PMap}(S), \mathbb{Z})$ is not trivial as soon as S has at least two non-planar ends. A natural problem is:

Problem 12.5.16 *Compute the low-dimensional (co-)homology groups of* $\mathrm{Map}(S)$ *and* $\mathrm{PMap}(S)$.

In his original blogpost, Calegari [30] showed that the mapping class group of the Cantor tree surface is uniformly perfect, which implies that both H_1 and H^1 are trivial (with integer coefficients). Recently, Calegari–Chen have computed the second homology; we record both results below:

Theorem 12.5.17 ([30, 31]) *Let Γ denote the mapping class group of the Cantor tree surface. Then $H^1(\Gamma, \mathbb{Z})$, $H_1(\Gamma, \mathbb{Z})$ and $H^2(\Gamma, \mathbb{Z})$ are trivial, and $H_2(\Gamma, \mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}$.*

The following structural result about pure mapping class groups provides the core piece in the proof of Theorem 12.5.15; compare with Theorem 12.4.4 above:

Theorem 12.5.18 *For any surface S, we have*

$$\mathrm{PMap}(S) = \overline{\mathrm{Map}_c(S)} \rtimes \prod_{s \in \mathcal{S}} \langle h_s \rangle,$$

where the rightmost group is a direct product of cyclic groups generated by pairwise-commuting handle shifts h_s, where s ranges over a free basis of $H_1^{sep}(\hat{S}, \mathbb{Z})$.

Theorem 12.5.15 leaves out some low-genus cases, which were subsequently settled by Domat–Plummer [37]. More concretely, they proved the following result for genus-one surfaces:

Theorem 12.5.19 ([37]) *Let S be an infinite-type surface of genus one. Then*

$$H^1(\mathrm{PMap}(S), \mathbb{Z}) = 0.$$

For an infinite-type surface S of genus-zero the situation is different, for in this case there is a surjective homomorphism $\mathrm{PMap}(S) \to \mathbb{F}_2$, the free group on two generators, since the pure mapping class group of a four-times punctured sphere is isomorphic to $\mathbb{F}_2$. Nevertheless, Domat–Plummer prove:

Theorem 12.5.20 ([37]) *Let S be an infinite-type surface of genus zero. Then $H^1(\mathrm{PMap}(S), \mathbb{Z})$ contains uncountably many classes which do not come from forgetful maps to spheres with finitely many punctures.*

12.5.3 Quantifying Rigidity

In Sect. 12.5.1, we saw that automorphisms of mapping class groups are geometric. In fact, something stronger is true: outside several low-complexity cases, given two surfaces S_1 and S_2 any isomorphism $\mathrm{Map}(S_1) \to \mathrm{Map}(S_2)$ (or $\mathrm{PMap}(S_1) \to \mathrm{PMap}(S_2)$) is induced by a homeomorphism $S_1 \to S_2$ (this is shown in [17] in the infinite-type setting and can be deduced in the finite-type setting from [64, 71, 74]). In particular, in the finite-type setting, using the virtual cohomological dimension [56] and algebraic rank [21] of $\mathrm{Map}(S)$, it is possible to determine the topology of S from algebraic invariants of $\mathrm{Map}(S)$. Given that rigidity holds in big mapping class groups, it should be possible to do the same:

Question 12.5.21 *Is there a list of algebraic invariants of* $\mathrm{Map}(S)$ *that determine the topology of S?*

Let us provide some examples connecting algebraic invariants of $\mathrm{Map}(S)$ and the topology of S. First, we have the following corollary of Theorem 12.5.15:

Corollary 12.5.22 ([11]) *The algebraic rank of $H^1(\mathrm{PMap}(S), \mathbb{Z})$ is:*

- *0 if and only if S has at most one non-planar end.*
- *$n \in \mathbb{N}$ if and only if S has $n + 1$ non-planar ends.*
- *infinite if and only if S has infinitely many non-planar ends.*

Next, recall that a group is *residually finite* if and only if the intersection of all its normal subgroups is the identity.

Theorem 12.5.23 ([87]) *Let S be any surface.*

- $\mathrm{PMap}(S)$ *is residually finite if and only if S has finite genus.*
- $\mathrm{Map}(S)$ *is residually finite if and only if S is of finite type.*

Now, it follows from the work of Bavard–Walker [16] that if S has an isolated planar end then $\mathrm{PMap}(S)$ is circularly orderable (though not equivalent, the reader can read this as "acts faithfully on the circle"). Moreover, by forthcoming work of Aougab, Patel, and the second author [5], every finite group can be realized as a subgroup of $\mathrm{PMap}(S)$ whenever S has infinite-genus and no planar ends. Combining these facts, with the two results mentioned in this subsection and the fact that $\mathrm{Aut}(\mathrm{PMap}(S)) \cong \mathrm{Map}^{\pm}(S)$ when S is of infinite-type [17], we are able to give a complete answer to Question 12.5.21 for a countably infinite family of surfaces:

Theorem 12.5.24 ([5]) *For $n \in \mathbb{N}$, let Ω_n denote the n-ended infinite-genus surface with no planar ends and let $G = \mathrm{PMap}(S)$ for some surface S. The surface S is homeomorphic to Ω_n if and only if G satisfies each of the following properties:*

(1) G is not residually finite,
(2) G is not circularly orderable,
(3) $H^1(G, \mathbb{Z})$ has rank $n - 1$, and
(4) G is finite index in $\mathrm{Aut}(G)$.

12.5.4 Homology Representation

As mentioned in Sect. 12.5.1, there is a homomorphism

$$\rho_S : \mathrm{Map}(S) \to \mathrm{Aut}(H_1(S, \mathbb{Z})),$$

given by the action of mapping classes on the homology of the surface. For finite-type surfaces with at most one puncture or boundary component, the *algebraic intersection* pairing of homology classes is a symplectic form, and one shows that the homomorphism

$$\mathrm{Map}(S) \to \mathrm{Sp}(2g, \mathbb{Z}),$$

where g is the genus of S, is surjective; see [43, Section 6] for details.

The homology representation for infinite-type surfaces has been studied by Fanoni, Hensel, and the second author [41]. In the infinite-type setting, there is only one surface with at most one end, namely the Loch Ness monster surface; in this case, it turns out an analogous result holds:

Theorem 12.5.25 ([41]) *Let S be the Loch Ness monster surface. Then the image of the homology representation is the subgroup of $\mathrm{Aut}(H_1(S, \mathbb{Z}))$ consisting of those elements which preserve the algebraic intersection form. In other words,*

$$\mathrm{Im}(\phi_S) = \mathrm{Sp}(\mathbb{N}, \mathbb{Z}).$$

For surfaces with more than one end (or boundary component), preserving algebraic intersection is not enough to characterize the image of ρ_S in $H_1(S, \mathbb{Z})$ (this is true in both the finite-type and infinite-type settings). In the same article [41], the authors give a characterization of the image of ρ_S for an arbitrary surface S in terms of preserving a filtration of the first homology. The full statement is a bit technical, so we refer the interested reader directly to [41].

12.5.5 Nielsen Realization

Kerckhoff's *Nielsen Realization Theorem* [68] asserts that every finite subgroup of the mapping class group of a finite-type surface S of negative Euler characteristic lifts to $\mathrm{Homeo}(S)$; moreover, it may be realized as a subgroup of the isometry group of some hyperbolic metric on S.

In the context of big mapping class groups, the analogous statement has been obtained by Afton–Calegari–Chen–Lyman [2]:

Theorem 12.5.26 *Let S be a surface of infinite type. Then every finite subgroup of* $\mathrm{Map}(S)$ *lifts to* $\mathrm{Homeo}^+(S)$. *Moreover, every finite group can be realized as a group of isometries of some hyperbolic metric on S.*

We should also note that there is analog of Nielsen realization in the setting of analytically-infinite Riemann surfaces due to Markovic [78]. A *hyperbolic Riemann surface* is a complex 1-manifold whose universal cover is isomorphic to the unit disk.

Theorem 12.5.27 ([78]) *Let S be an infinite-type surface and let G be a subgroup of* $\mathrm{Map}(S)$. *If there exists a hyperbolic Riemann surface X homeomorphic to S and a constant $K > 1$ such that every element of G can be realized by a K-quasi-conformal homeomorphism $X \to X$, then there is a hyperbolic Riemann surface Y such that Y is quasi-conformally equivalent to X and $G <$* $\mathrm{Isom}(Y)$.

12.5.6 The Relation with Thompson Groups

Thompson's groups F, T and V constitute prominent examples of discrete subgroups of $\mathrm{Homeo}(C)$, the homeomorphism group of the Cantor set. Among many other features, they are infinite groups of type F_∞, and which have simple commutator subgroup (in fact, V itself is simple). We now briefly review the construction of these groups, referring the reader to the standard reference [32] for a thorough treatment of Thompson's groups.

12.5.6.1 Thompson's Groups

Let $\mathcal{T}$ be a rooted binary tree, noting that its space of ends of $\mathcal{T}$ is homeomorphic to the Cantor set C. The tree $\mathcal{T}$ has a natural left-to-right orientation once we fix a realization of $\mathcal{T}$ as a subset of the hyperbolic plane. With respect to this orientation, given a subtree of $\mathcal{T}$ with n leaves, we may order its set of leaves using the numbers $1, \ldots, n$, so that the numbers increase from left to right.

Let τ, τ' be subtrees of $\mathcal{T}$ with the same number of leaves, and such that both contain the root of $\mathcal{T}$. If σ is a bijection between the sets of leaves of τ and τ', then the triple (τ, τ', σ) extends in a natural way to a homeomorphism of C. Of course, the same homeomorphism may be induced by different such triples (obtained by *expanding* and *contracting* a given finite subtree), and Thompson's group V is the group of equivalence classes of such triples. In turn, Thompson's group T (resp. F) corresponds to the case when the bijection σ is a cycle (resp. the identity).

12.5.6.2 Asymptotic Mapping Class Groups

We now explain the relation between Thompson's groups and big mapping class groups. To this end, let S denote either the Cantor tree surface or the blooming Cantor tree surface. In these particular cases, the exact sequence (12.3) reads

$$1 \to \mathrm{PMap}(S) \to \mathrm{Map}(S) \to \mathrm{Homeo}(C) \to 1. \tag{12.5}$$

Over the last two decades, numerous authors have given geometric constructions of finitely-presented subgroups H of $\mathrm{Map}(S)$ for which the sequence (12.5) restricts to

$$1 \to \mathrm{Map}_c(S) \to H \to G \to 1, \tag{12.6}$$

where G is one of Thompson's groups F, T or V (or their commutator subgroups).

To the best of our knowledge, the first step in this direction was the paper of Greenberg–Sergiescu [52], whose objective was to construct an acyclic extension of F', the commutator subgroup of F, by the braid group B_∞ on infinitely many strands. This was later generalized simultaneously by Brin [28] and Dehornoy [35] to the construction of an extension of V by B_∞, the so-called *braided Thompson groups*. Funar–Kapoudjian [45, 46], and later Funar and the first author [6], constructed finitely-generated (and often finitely-presented) extensions of V by a direct limit of mapping class groups of compact surfaces. Part of the motivation [46] is to construct a finitely-presented group whose homology agrees with the *stable homology* of pure mapping class groups, after a seminal result of Harer [55].

A common feature of all of the above constructions is that they may be expressed in terms of groups of homeomorphisms of an infinite-type surface which *eventually* preserve some topological data; these are the *asymptotic mapping class groups* introduced by Funar–Kapoudjian in [45]. We now briefly recall their definition in the simpler case of a surface of genus zero.

12.5.6.3 The Case of the Cantor Tree Surface

Let S be the Cantor tree surface, that is, a sphere with a Cantor set removed. Fix, once and for all, a pants decomposition P of S and a set A of pairwise-disjoint, properly-embedded arcs on S such that $S \setminus A$ has exactly two connected components $\nu^\pm$, and each connected component of $S \setminus P$ is intersected by exactly three elements of A; see Fig. 12.4. The triple (P, A, ν^+) is called a *rigid structure* on S.

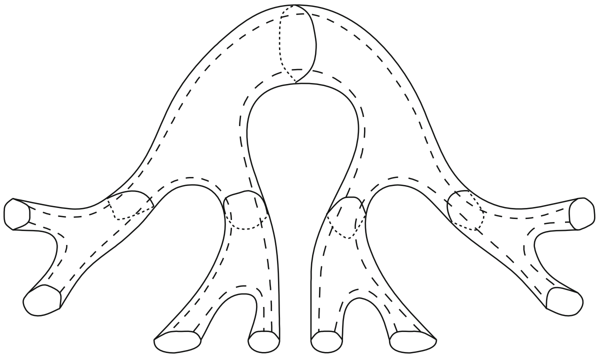


Fig. 12.4 The rigid structure on S

We say that a homeomorphism $f : S \to S$ is *asymptotically rigid* if there exists a compact subsurface $X \subset S$ with $\partial X \subset P$, such that $\partial f(X) \subset P$ and the restriction homeomorphism

$$f : S \setminus X \to S \setminus f(X)$$

setwise preserves (the relevant part of) the rigid structure. The group $\mathcal{B}$ is then defined as the subgroup of Map(S) whose elements have an asymptotically rigid homeomorphism. In their paper [45], Funar and Kapoudjian showed that the restriction of the sequence (12.6) yields

$$1 \to \mathrm{Map}_c(S) \to \mathcal{B} \to V \to 1, \tag{12.7}$$

As such, $\mathcal{B}$ contains the mapping class group of every compact surface of genus zero with non-empty boundary. In light of this, the main result of [45] is rather striking:

Theorem 12.5.28 ([45]) *The group $\mathcal{B}$ is finitely presented.*

Moreover, they observed:

Proposition 12.5.29 *The short exact sequence* (12.7) *splits over Thompson's group T. As a consequence, $\mathcal{B}$ is not linear and does not have Kazhdan's Property (T).*

We remark that a well-known question about finite-type mapping class groups asks whether they are linear or have Kazhdan's Property (T).

12.5.6.4 Other Compact Surfaces with a Cantor Set Removed

The construction of asymptotic mapping class groups makes sense for arbitrary surfaces. In fact, as commented in [45], the group constructed by Brin [28] and Dehornoy [35] are asymptotic mapping class groups of a closed disc with a Cantor set removed, and as such embeds as a subgroup of $\mathcal{B}$. In addition, Funar and the first author [6] generalized the construction of $\mathcal{B}$ to the surface Σ_g obtained by removing a Cantor set from a closed surface S_g of finite genus $g \geq 1$. Roughly speaking, a *rigid structure* on Σ_g is determined by a simple closed curve $\alpha \subset \Sigma_g$ that cuts off a once-punctured surface of genus g, together with a rigid structure for the planar component of Σ_g. One then defines the notion of an asymptotically rigid homeomorphism in an analogous way, and constructs the *asymptotic mapping class group* $\mathcal{B}_g$ as the subgroup of $\mathrm{Map}(\Sigma_g)$ whose elements have an asymptotically rigid representative. In this case, the restriction of the short exact sequence (12.5) to the group $\mathcal{B}_g$ reads

$$1 \to \mathrm{Map}_c(\Sigma_g) \to \mathcal{B}_g \to V \to 1; \tag{12.8}$$

in particular, $\mathcal{B}_g$ contains the mapping class group of every compact surface of genus at most g and with non-empty boundary. The following is one of the main results of [6]:

Theorem 12.5.30 ([6]) *For every $g \geq 1$, the group $\mathcal{B}_g$ is finitely presented. In addition, it is not linear and does not have Kazhdan's Property (T).*

In addition, in [6] the authors explore the structure of the groups $\mathcal{B}_g$ in connection with mapping class groups of finite-type surfaces. For instance, every automorphism of $\mathcal{B}_g$ is induced by a homeomorphism of Σ_g (compare with Theorems 12.5.1 and 12.5.7).

12.5.6.5 The Case of the Blooming Cantor Tree

In [46], Funar and Kapoudjian constructed an asymptotic mapping class group $\mathcal{B}_\infty$ for the blooming Cantor tree, which we denote by Σ_∞. In a similar fashion, the short exact sequence (12.5), when restricted to $\mathcal{B}_\infty$, yields:

$$1 \to \mathrm{Map}_c(\Sigma_\infty) \to \mathcal{B}_\infty \to V \to 1. \tag{12.9}$$

The following is the main result of [46]:

Theorem 12.5.31 ([46]) *The group $\mathcal{B}_\infty$ is finitely generated. Moreover, its rational cohomology coincides with the stable rational cohomology of the mapping class group.*

Note that, while asymptotic mapping class groups of finite genus are finitely presented, the group $\mathcal{B}_\infty$ is only known to be finitely generated. In light of this, we ask:

Question 12.5.32 *Determine whether the asymptotic mapping class groups $\mathcal{B}_n$, for $n \in \mathbb{N} \cup \{\infty\}$, satisfy stronger finiteness properties. Are they F_∞?*

A positive answer to the above question, in the case of $n = 0$, is conjectured in [45, p. 967]. The question of whether $\mathcal{B}_\infty$ is finitely presented appears in [48].

12.5.6.6 A Dense Asymptotic Mapping Class Group

In addition, in [6] the authors considered a subgroup $\mathcal{H}_g$ with $\mathcal{B}_g < \mathcal{H}_g < \mathrm{Map}(\Sigma_g)$. The definition of $\mathcal{H}_g$ is similar to that of $\mathcal{B}_g$, without the requirement that its elements preserve the connected component ν^+ appearing in the definition of rigid structure. In short, the difference between $\mathcal{B}_g$ and $\mathcal{H}_g$ is that the latter contains *half-twists* about separating curves cutting off a disk minus a Cantor set. For this reason, the group $\mathcal{H}_g$ is referred to as the *group of half-twists*.

A large part of the motivation for considering $\mathcal{H}_g$ comes from the study of smooth mapping class groups, as explained in [47]. Indeed, put a differentiable structure on the closed surface S_g of genus g, and realize C as the the middle-third Cantor set on a smoothly embedded interval on S_g. Let $\mathrm{Mod}^s(S_g, C)$ denote the *smooth mapping class group* of the pair (S_g, C), namely the group of isotopy classes of smooth diffeomorphisms of S_g preserving globally the Cantor set C. The following is a recent result of Funar and Neretin [47]:

Theorem 12.5.33 ([47], Cor. 2) *For every $g \geq 0$, we have $\mathcal{H}_g \cong \mathrm{Mod}^s(S_g, C)$.*

Using the same techniques as with $\mathcal{B}_g$, Funar and the first author [6] proved:

Theorem 12.5.34 ([6]) *For every $g \geq 1$, the group $\mathcal{H}_g$ is finitely presented. In addition, it is not linear and does not have Kazhdan's Property (T).*

However, a nice extra feature of the group $\mathcal{H}_g$ is the following result, which should be compared with Theorem 12.4.3:

Theorem 12.5.35 *For every $g \geq 0$ the group $\mathcal{H}_g$ is dense in* $\mathrm{Map}(\Sigma_g)$.

Finally, the restriction to $\mathcal{H}_g$ of the sequence (12.7) reads

$$1 \to \mathrm{Map}_c(\Sigma_g) \to \mathcal{H}_g \to V_2[\mathbb{Z}_2] \to 1, \tag{12.10}$$

where $V_2[\mathbb{Z}_2]$ is the *Higman–Thompson group* $V_2[\mathbb{Z}_2]$ [22]. A surprising result of Bleak–Donoven–Jonušas [22] establishes that V and $V_2[\mathbb{Z}_2]$ are conjugate as subgroups of $\mathrm{Homeo}(C)$ through an explicit homeomorphism of C (a *cellular automaton*). An obvious questions then is:

Question 12.5.36 *Are the groups $\mathcal{B}_g$ and $\mathcal{H}_g$ isomorphic?*

It would be surprising if the question above had a positive answer, since isomorphisms between (sufficiently rich) subgroups of mapping class groups tend to come from surface homeomorphisms.

We end this section with the following vague question:

Question 12.5.37 *Are there other geometrically-defined subgroups of* $\mathrm{Map}(\Sigma_g)$ *which surject to other interesting classes of subgroups of* $\mathrm{Homeo}(C)$*, such as Higman-Thompson groups, Neretin groups, etc.?*

12.6 Geometric Aspects

Mapping class groups of finite-type surfaces have been successfully studied through their action on various combinatorial complexes, notably the curve graph; a first instance of this is Ivanov's Rigidity Theorem mentioned in Sect. 12.5.1. Moreover, it turns out that the geometric structure of $\mathcal{C}(S)$, equipped with its natural path metric, sheds intense light on the algebraic and geometric structure of $\mathrm{Map}(S)$. In this direction, the following is a seminal theorem of Masur–Minsky [79]:

Theorem 12.6.1 ([79]) *Let* S *be a finite-type surface. If* $\mathcal{C}(S)$ *is connected, then it is hyperbolic (in the sense of Gromov).*

A number of authors have proved analogous results for other combinatorial complexes associated to surfaces, such as the disk graph [80], the non-separating curve graph $\mathrm{NonSep}(S)$ [54, 80], the arc graph $\mathcal{A}(S)$ [57], etc. In fact, a surprising phenomenon is that the hyperbolicity constant in Theorem 12.6.1, as well as the those of other complexes, turn out to be independent of the underlying surface; we say that the corresponding family of complexes are *uniformly hyperbolic*. The following theorem is a combination of the results of [4, 25, 34, 57, 93]:

Theorem 12.6.2 *Let* S *be a finite-type surface.*

(1) ([57]) $\mathcal{A}(S)$ *is uniformly hyperbolic.*
(2) ([4, 25, 34, 57]) $\mathcal{C}(S)$ *is uniformly hyperbolic*
(3) ([93]) For fixed g*, the graph* $\mathrm{NonSep}(S_{g,n}))$ *is hyperbolic with respect to a constant which does not depend on* n*.*

The above result may be regarded as a curiosity at first, but it happens to be of central importance in the study of big mapping class groups, as we will explain next.

12.6.1 Complexes for Infinite-Type Surfaces

As in the finite-type case, one may be tempted to use interesting geometric properties of analogous combinatorial models, built from arcs and/or curves, in order to study mapping class groups. This initial surge of enthusiasm is thwarted by the following immediate observation; before we state it, we recall that, for an infinite-type surface S, the arc graph $\mathcal{A}(S)$ is defined to be the simplicial graph whose vertices are properly embedded arcs on S which join two (not necessarily distinct) planar ends of S, and where adjacency corresponds to disjointness.

Fact 12.6.3 *Let S be a surface of infinite type. Then $\mathcal{C}(S)$ has diameter two. Furthermore, if S has infinitely many planar ends, then $\mathcal{A}(S)$ also has diameter two.*

However, as mentioned in the introduction, in [30] Calegari proposed studying $\mathrm{Map}(\mathbb{R}^2 \setminus C)$ via its action on a certain subgraph of $\mathcal{A}(\mathbb{R}^2 \setminus C)$; observe that, by the above, $\mathcal{A}(\mathbb{R}^2 \setminus C)$ itself has diameter two. Calegari's idea was to consider the subgraph $\mathcal{A}_\infty$ of $\mathcal{A}(\mathbb{R}^2 \setminus C)$ consisting of arcs with at most one endpoint in C (hence, necessarily one end of an arc in $\mathcal{A}_\infty$ is contained in the unique isolated planar end of $\mathbb{R}^2 \setminus C$). The next result was proved by Juliette Bavard [13, 14] proving a conjecture posed by Calegari:

Theorem 12.6.4 *$\mathcal{A}_\infty$ is a Gromov-hyperbolic space of infinite diameter.*

Based on this result, and with a lot of extra work, she also proved that $\mathrm{Map}(\mathbb{R}^2 \setminus C)$ has an infinite-dimensional space of quasi-morphisms. This is in stark contrast to $\mathrm{Map}(\mathbb{S}^2 \setminus C)$, which Calegari shows admits no quasi-morphisms (and even stronger, we know $\mathrm{Map}(\mathbb{S}^2 \setminus C)$ is CB [77]). We note that the automorphism group of $\mathcal{A}_\infty$ and related graphs are computed in [101] and shown to be the extended mapping class group.

The above theorem may be regarded as part of a more general phenomenon, which we now explain. In order to do so, we need the following terminology due to Schleimer [102]. Given a graph $\mathfrak{X}(S)$ built from arcs and/or curves on S, say that a subsurface $Y \subset S$ is a *witness* for $\mathfrak{X}(S)$ if every vertex of $\mathfrak{X}$ intersects Y non-trivially. For instance, the only non-trivial witness for $\mathcal{C}(S)$ is S itself while, in the case of $\mathcal{A}(S)$, any subsurface $Y \subset S$ which contains every puncture of S is a witness.

The following theorem is a reformulated version of [8, Theorem 1] (see also [36, Section 6] for another formulation). In an intuitive way, it encapsulates the idea of *taking a limit of a family of uniformly hyperbolic spaces*:

Theorem 12.6.5 *Let $\mathfrak{X}(S)$ be a connected* $\mathrm{Map}(S)$*-invariant graph, whose vertices are defined by finite sets of arcs or curves on S, and where edges correspond to bounded intersection number. Given a subsurface $Y \subset S$, define $\mathfrak{X}(Y)$ to be the full subgraph of $\mathfrak{X}(S)$ spanned by those vertices which are entirely contained in Y and equip $\mathfrak{X}(Y)$ with the induced path metric. Suppose that:*

(1) For every triangle T in $\mathfrak{X}(S)$ there exists a finite-type witness Y such that T is contained in $\mathfrak{X}(Y)$ and $\mathfrak{X}(Y)$ is connected;
(2) There exists constants $\delta, K, C > 0$ such that for every finite-type witness Y of S with $\mathfrak{X}(Y)$ connected, the following conditions are satisfied:

(a) $\mathfrak{X}(Y)$ is a δ-hyperbolic graph of infinite diameter.
(b) The inclusion map $\mathfrak{X}(Y) \hookrightarrow \mathfrak{X}(S)$ is a (K, C)-quasi-isometric embedding.

Then $\mathfrak{X}(S)$ is hyperbolic and has infinite diameter.

Given a finite set P of isolated planar ends of S, denote by $\mathcal{A}(S; P)$ the subgraph of $\mathcal{A}(S)$ spanned by those arcs which have at least one endpoint in P; observe that every subsurface of S which contains P is a witness for $\mathcal{A}(S; P)$. The above result and the uniform hyperbolicity presented in Theorem 12.6.2 are used to prove the following:

Theorem 12.6.6 *Let S be an infinite-type surface.*

(1) ([8, 9, 13]) Let P be a non-empty finite set of isolated punctures of S. Then, $\mathcal{A}(S; P)$ is hyperbolic.
(2) ([93]) If S has finite genus at least 2, then the graph NonSep(S) *is hyperbolic.*

Remark There is a subtlety about Theorem 12.6.6 which is worth mentioning at this point; see also [8, Theorem 1]. Let P, Q be two finite sets of isolated punctures of S, with $P \cap Q = \emptyset$, and consider the subgraph $\mathcal{A}(S; P, Q)$ of $\mathcal{A}(S)$ which have one endpoint in P *and* one endpoint in Q. Then $\mathcal{A}(S; P, Q)$ is not hyperbolic.

Indeed, this is a manifestation of Schleimer's *Disjoint Witness Property* [102], which asserts that if a graph or curves/arcs has two disjoint witnesses of infinite diameter then it is not hyperbolic, for one may use subsurface projections to construct a quasi-isometrically embedded copy of $\mathbb{Z}^2$ inside the graph.

Finally, observe that the graph $\mathcal{A}(S; P, Q)$ contains two disjoint witnesses, since one can take two finite-type surfaces, one containing P and the other containing Q. This finishes the remark.

These different phenomena were clarified in subsequent work of Durham, Fanoni and the second author [36]. The motivation of their work was to find actions of big mapping class group not relying on isolated planar ends. Before explaining their result, we need some definitions.

Let $\mathcal{Q}$ be a collection of pairwise-disjoint closed subsets of Ends(S). Every separating curve on S partitions Ends(S); let Sep$_2(S, \mathcal{Q})$ denote the subgraph of $\mathcal{C}(S)$ consisting of separating curves on S that partition $\mathcal{Q}$ into two sets, each of cardinality at least 2 (there is a slight modification if $|\mathcal{Q}| = 4$, see [36] for details).

Theorem 12.6.7 ([36]) *Let S be an infinite-type surface. Let $\mathcal{Q}$ be a collection of pairwise-disjoint closed subsets of* Ends(S) *such that, for every $\omega \in \mathcal{Q}$ and every $f \in$* Map(S)*, there exists $\omega' \in \mathcal{Q}$ with $f(\omega) = \omega'$. Then,* Sep$_2(S, \mathcal{Q})$ *is hyperbolic, infinite diameter,* Map(S)*-invariant, and there are infinitely many mapping classes which act with positive translation length on* Sep$_2(S, \mathcal{Q})$.

For example, if $S = \Omega_n$ (the n-ended infinite-genus surface with no planar ends) with $n \geq 4$, then $\mathcal{Q} =$ Ends(S) satisfies the hypothesis of the above theorem.

We note that in the days this survey was being finalized, Fanoni–Ghaswala–McLeay [40] constructed new examples of hyperbolic infinite-diameter graphs that admit actions of big mapping class groups with unbounded orbits. We direct the reader to their article for details.

Klarreich [70] showed that the Gromov boundary of the curve graph is Map(S)-equivariantly homeomorphic to the space of *ending laminations* on the surface; see also [53] for a different argument, and Pho-On's thesis [89] for an effective proof of this using the *unicorn* machinery of [57]. In unpublished work, Schleimer proved that the boundary of the arc graph is naturally identified with the space of all ending laminations supported on witnesses of S; this is also carried out in an effective manner in Pho-On's thesis [89].

In light of these results, we ask the following natural question:

Question 12.6.8 *Describe the Gromov boundary of the various hyperbolic complexes associated to an infinite-type surface S, ideally in terms of laminations/foliations on S.*

For the case when the surface is $\mathbb{R}^2 \setminus C$, the Gromov boundary of the relative arc graph $\mathcal{A}_\infty$ of Theorem 12.6.4 is described by Bavard-Walker [16] in terms of rays on the surface. Rasmussen [94] has recently reproved a result of Hamenstädt computing the Gromov boundary of the graph of non-separating curves and points out that his techniques can be extended to the infinite-type setting; however, the issue is a lack of understanding of laminations on infinite-type surfaces. We should note at this point that Šarić [100] recently developed the theory of train tracks for infinite-type surfaces, which should aid in investigating laminations.

The natural motivation for understanding the Gromov boundary is to gain insight into a potential classification of big mapping classes akin to that of the Nielsen–Thurston classification. We should note that there is much research in this direction for quasi-conformal mapping class groups and their action on Teichmüller space.

12.6.2 Weak Proper Discontinuity and Acylindricity

Let G be a group acting by isometries on a hyperbolic metric space (X, d). We say that the action is *acylindrical* if, for every $D \geq 0$, there exists $R \geq 0$ such that, for every $x, y \in X$ with $d(x, y) \geq D$, the cardinality of the set

$$\{g \in G \mid d(x, gx), d(y, gy) \leq R\}$$

is finite. To exclude uninteresting pathologies, we restrict our attention to actions where there are infinitely many points on the Gromov boundary of X that are accumulation points of G-orbits; call such an action *non-elementary*. We say that a group is *acylindrically hyperbolic* if it admits a non-elementary acylindrical action on some Gromov-hyperbolic space.

A result of Bowditch [24] asserts that, if S has finite type, the action of $\mathrm{Map}(S)$ on $\mathcal{C}(S)$ is acylindrical. Bavard–Genevois [15] proved that the analogous statement does not hold for infinite-type surfaces:

Theorem 12.6.9 ([15]) *If S has infinite type, then* $\mathrm{Map}(S)$ *is not acylindrically hyperbolic.*

Prior to the notion of acylindricity, Bestvina–Fujiwara [19] introduced the concept of *weak proper discontinutity* (WPD, for short), and used it to show that if a group has an interesting WPD action then it has an infinite-dimensional space of quasimorphisms; equivalently, its second bounded cohomology group is infinite-dimensional. We briefly recall these ideas. Let G be a group acting on a Gromov-hyperbolic metric space (X, d), and $g \in G$ be a loxodromic element. We say that g is a *WPD element* if, for every $x \in X$ and every $R \geq 0$, there is $N \in \mathbb{N}$ such that the set

$$\{h \in G \mid d(x, h(x)) \leq R, d(g^N(x), hg^N(x) \leq R\}$$

is finite. Bestvina–Fujiwara [19] proved that, for a finite-type surface S, any pseudo-Anosov element of $\mathrm{Map}(S)$ is WPD with respect to the natural action on the curve complex. This notion was further weakened by Bromberg–Bestvina–Fujiwara [20] to that of a *WWPD action*: suppose again G acts on a hyperbolic space X, and let g be a loxodromic element of G with fixed points $\eta^{\pm}$ on the Gromov boundary ∂X of X. We say that g is a *WWPD* element if, for every sequence $\{h_n\}_{n \in \mathbb{N}}$ of elements of G, with $h_n(\eta^+) \to \eta^+$ and $h_n(\eta)^- \to \eta^-$, there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, one has

$$h_n(g^+) = g^+ \qquad \text{and} \qquad h_n(g^-) = g^-.$$

The existence of WWPD elements of big mapping class groups has been recently studied by Rasmussen [92]. Let S be an infinite-type surface with at least one isolated puncture p, and let $\mathcal{A}(S, p)$ be the relative arc graph of S based at p. Rasmussen proved:

Theorem 12.6.10 ([92]) *An element $g \in \mathrm{Map}(S)$ is WWPD with respect to the action of $\mathrm{Map}(S)$ on $\mathcal{A}(S, p)$ if and only if there exists a finite-type g-invariant subsurface $Y \subset S$, with $p \in Y$, such that the restriction of g to Y is pseudo-Anosov.*

As a consequence, he deduces that a class of subgroups of $\mathrm{Map}(S)$ have infinite-dimensional second bounded cohomology.

We finish with mentioning a very recent construction of Morales–Valdez [85], in which they construct examples of mapping classes which act loxodromically on $\mathcal{A}_\infty$ and do not preserve any finite-type subsurface.

Acknowledgments We would like to thank Athanase Papadopoulos for inviting us to write this survey.

This chapter is heavily influenced by the AIM workshop "Surfaces of infinite type" (29 April–3 May 2019). We would like to thank the organizers of the workshop, J. Bavard, A. Randecker, P. Patel, and J. Tao; we are also grateful to AIM for its hospitality and financial support.

J.A. was supported by grants RYC-2013-13008 and PGC2018-101179-B-I00. He acknowledges financial support from the Spanish Ministry of Science and Innovation, through the "Severo Ochoa Programme for Centres of Excellence in R&D" (SEV-2015-0554) and from the Spanish National Research Council, through the "Ayuda extraordinaria a Centros de Excelencia Severo Ochoa" (20205CEX001).

N.G.V. was partially supported by PSC-CUNY grant 62571-00 50.

Finally, J.A. is indebted to his baby daughter Berta for improving his ability at typing with one hand.

References

1. S. Afton, S. Freedman, J. Lanier, L. Yin, Generators, relations, and homomorphisms of big mapping class groups (in preparation)
2. S. Afton, D. Calegari, L. Chen, R.A. Lyman, Nieslen realization for infinite-type surfaces in *To appear in Proceedings of the AMS*
3. L.V. Ahlfors, L. Sario. *Riemann Surfaces*. Princeton Mathematical Series, vol. 26 (Princeton University Press, Princeton, NJ, 1960)
4. T. Aougab, Uniform hyperbolicity of the graphs of curves. Geom. Topol. **17**(5), 2855–2875 (2013)
5. T. Aougab, P. Patel, N.G. Vlamis, Isometry groups of infinite-genus hyperbolic surfaces. Preprint. arXiv:2007.01982
6. J. Aramayona, L. Funar, Asymptotic mapping class groups of closed surfaces punctured along Cantor sets. Mosc. Math. J. Preprint. arXiv:1701.08132
7. J. Aramayona, J. Souto, Homomorphisms between mapping class groups. Geom. Topol. **16**(4), 2285–2341 (2012)
8. J. Aramayona, F. Valdez, On the geometry of graphs associated to infinite-type surfaces. Math. Z. **289**(1–2), 309–322 (2018)
9. J. Aramayona, A. Fossas, H. Parlier, Arc and curve graphs for infinite-type surfaces. Proc. Am. Math. Soc. **145**(11), 4995–5006 (2017)
10. J. Aramayona, T. Ghaswala, A. M^{c}Leay, A. E. Kent, J. Tao, R. Winarski, Big Torelli groups: generation and commensurations. Groups Geom. Dyn. **13**(4), 1373–1399 (2019)
11. J. Aramayona, P. Patel, N. Vlamis, The first integral cohomology of big mapping class groups. Int. Math Res. Not. (in press). arXiv:1711.03132
12. A. Basmajian, Generalizing the hyperbolic collar lemma. Bull. Am. Math. Soc. **27**(1), 154–158 (1992)
13. J. Bavard, Hyperbolicité du graphe des rayons et quasi-morphismes sur un gros groupe modulaire. Geom. Topol. **20**(1), 491–535 (2016)
14. J. Bavard, Gromov-hyperbolicity of the ray graph and quasimorphisms on a big mapping class group (2018). Preprint. arXiv:1802.02715
15. J. Bavard, A. Genevois, Big mapping class groups are not acylindrically hyperbolic. Math. Slovaca **68**(1), 71–76 (2018)
16. J. Bavard, A. Walker, The Gromov boundary of the ray graph. Trans. Am. Math. Soc. **370**(11), 7647–7678 (2018)
17. J. Bavard, S. Dowdall, K. Rafi, Isomorphisms between big mapping class groups. Int. Math. Res. Not. **2020**(10), 3084–3099 (2020)
18. L. Bers, Universal Teichmüller space. *Analytic Methods in Mathematical Physics (Sympos., Indiana Univ., Bloomington, Ind., 1968)* (Gordon and Breach, New York, 1970), pp. 65–83
19. M. Bestvina, K. Fujiwara, Bounded cohomology of subgroups of mapping class groups. Geom. Topol. **6**, 69–89 (2002)
20. M. Bestvina, K. Bromberg, K. Fujiwara, Constructing group actions on quasi-trees and applications to mapping class groups. Publ. Math. Inst. Hautes Études Sci. **122**, 1–64 (2015)
21. J. Birman, A. Lubotzky, J. McCarthy, Abelian and solvable subgroups of the mapping class groups. Duke Math. J. **50**(4), 1107–1120 (1983)
22. C. Bleak, C. Donoven, J. Jonušas, Some isomorphism results for Thompson-like groups $V_n(G)$. Israel J. Math. **222**(1), 1–19 (2017)
23. J. Birman, On Siegel's modular group. Math. Ann. **191**, 59–68 (1971)

24. B.H. Bowditch, Tight geodesics in the curve complex. Invent. Math. **171**(2), 281–300 (2008)
25. B.H. Bowditch, Uniform hyperbolicity of the curve graphs. Pac. J. Math. **269**(2), 269–280 (2014)
26. T. Brendle, D. Margalit, Commensurations of the Johnson kernel. Geom. Topol. **8**, 1361–1384 (2004)
27. T. Brendle, D. Margalit, Normal subgroups of mapping class groups and the metaconjecture of Ivanov. J. Am. Math. Soc. **32**(4), 1009–1070 (2019)
28. M.G. Brin, The algebra of strand splitting. I. A braided version of Thompson's group V. J. Group Theory **10**(6), 757–788 (2007)
29. D. Calegari, Big mapping class groups and complex dynamics. Preprint. Available at https://math.uchicago.edu/~dannyc/courses/dynamics_2019/big_mcg_dynamics.pdf
30. D. Calegari, Mapping class groups and dynamics. Blogpost available at https://lamington.wordpress.com/2009/06/22/big-mapping-class-groups-and-dynamics/
31. D. Calegari, L. Chen, Big mapping class groups and rigidity of the simple circle. Preprint. arXiv:1907.07903
32. J.W. Cannon, W.J. Floyd, W.J. Parry, Introductory notes on Richard Thompson's groups. Enseign. Math. (2) **42**(3–4), 215–256 (1996)
33. L. Chen, K. Mann, Structure theorems for actions of homeomorphism groups. Preprint. arXiv:1902.05117
34. M. Clay, K. Rafi, S. Schleimer, Uniform hyperbolicity of the curve graph via surgery sequences. Algebr. Geom. Topol. **14**(6), 3325–3344 (2014)
35. P. Dehornoy, The group of parenthesized braids. Adv. Math. **205**(2), 354–409 (2006)
36. M. Durham, F. Fanoni, N. Vlamis, Graphs of curves on infinite-type surfaces with mapping class group actions. Ann. Inst. Fourier **68**(6), 2581–2612 (2018)
37. G. Domat, P. Plummer, First cohomology of pure mapping class groups of big genus one and zero surfaces. New York J. Math. **26**, 322–333 (2020)
38. R.M. Dudley, Continuity of homomorphisms. Duke Math. J. **28**, 587–594 (1961)
39. D.B.A. Epstein, Curves on 2-manifolds and isotopies. Acta Math. **115**, 83–107 (1966)
40. F. Fanoni, T. Ghaswala, A. McLeay, Homeomorphic subsurfaces and the arcs that survive (2020). Preprint. arXiv:2003.04750
41. F. Fanoni, S. Hensel, N. Vlamis, Big mapping class groups acting on homology. Preprint. arXiv:1905.12509
42. B. Farb, N. V. Ivanov, The Torelli geometry and its applications. Math. Res. Lett. **12**(3), 293–301 (2005)
43. B. Farb, D. Margalit, *A Primer on Mapping Class Groups*. Princeton Mathematical Series, vol. 49 (Princeton University Press, Princeton, NJ, 2012)
44. R.H. Fox, On topologies for function spaces. Bull. Am. Math. Soc. **51**, 429–432 (1945)
45. L. Funar, C. Kapoudjian, On a universal mapping class group of genus zero. Geom. Funct. Anal. **14**(5), 965–1012 (2004)
46. L. Funar, C. Kapoudjian, An infinite genus mapping class group and stable cohomology. Commun. Math. Phys. **287**(3), 784–804 (2009)
47. L. Funar, Y. Neretin, Diffeomorphism groups of tame Cantor sets and Thompson-like groups. Compos. Math. **154**(5), 1066–1110 (2018)
48. L. Funar, C. Kapoudjian, V. Sergiescu, Asymptotically rigid mapping class groups and Thompson's groups. *Handbook of Teichmüller theory. Volume III*. IRMA Lectures in Mathematics and Theoretical Physics, vol. 17 (European Mathematical Society, Zürich, 2012), pp. 595–664
49. D. Gauld, *Non-metrisable Manifolds*, vol. 206 (Springer, Singapore, 2014)
50. É. Ghys, Prolongements des difféomorphismes de la sphère. L'Enseign. Math. **37**, 45–59 (1991)
51. É. Ghys, Topologie des feuilles génériques. Ann. Math. **141**, 387–422 (1995)
52. P. Greenberg, V. Sergiescu, An acyclic extension of the braid group. Comment. Math. Helv. **66**(1), 109–138 (1991)

53. U. Hamenstädt, Train tracks and the Gromov boundary of the complex of curves, in *Spaces of Kleinian Groups*. London Mathematical Society Lecture Note Series, vol. 329 (Cambridge University Press, Cambridge, 2006), pp. 187–207
54. U. Hamenstädt, Hyperbolicity of the graph of nonseparating multicurves. Algebr. Geom. Topol. **14**(3), 1759–1778 (2014)
55. J.L. Harer, Stability of the homology of the mapping class groups of orientable surfaces. Ann. Math. (2) **121**(2), 215–249 (1985)
56. J.L. Harer, The virtual cohomological dimension of the mapping class group of an orientable surface. Invent. Math. **84**(1), 157–176 (1986)
57. S. Hensel, P. Przytycki, R. Webb, 1-slim triangles and uniform hyperbolicity for arc graphs and curve graphs. J. Eur. Math. Soc. **17**(4), 755–762 (2015)
58. J. Hernández Hernández, Edge-preserving maps of curve graphs. Topol. Appl. **246**, 83–105 (2018)
59. J. Hernández Hernández, F. Valdez, Automorphism groups of simplicial complexes of infinite-type surfaces. Publ. Mat. **61**(1), 51–82 (2017)
60. J. Hernández Hernández, I. Morales, F. Valdez, Isomorphisms between curve graphs of infinite-type surfaces are geometric. Rocky Mountain J. Math. **48**(6), 1887–1904 (2018)
61. J. Hernández Hernández, I. Morales, F. Valdez, The Alexander method for infinite-type surfaces. Michigan Math. J. **68**(4), 743–753 (2019)
62. S. Hurtado, Continuity of discrete homomorphisms of diffeomorphism groups. Geom. Top. **19**(4), 2117–2154 (2015)
63. N.V. Ivanov, Automorphism of complexes of curves and of Teichmüller spaces. Int. Math. Res. Not. **14**, 651–666 (1997)
64. N.V. Ivanov, Automorphisms of Teichmüller modular groups, in *Topology and Geometry—Rohlin Seminar*. Lecture Notes in Mathematics, vol. 1346 (Springer, Berlin, 1988), pp. 199–270
65. N.V. Ivanov, J.D. McCarthy, On injective homomorphisms between TeichmÏler modular groups. I. Invent. Math. **135**(2), 425–486 (1999)
66. A. Kechris, *Classical Descriptive Set Theory*, vol. 156 (Springer Science & Business Media, Berlin/Heidelberg, 2012)
67. A. Kechris, C. Rosendal, Turbulence, amalgamation, and generic automorphisms of homogeneous structures. Proc. Lond. Math. Soc. (3), **94**(2), 2302–2350 (2007)
68. S.P. Kerckhoff, The Nielsen realization problem. Ann. Math. (2) **117**(2), 235–265 (1983)
69. B. Kerékjártó, *Vorlesungen über Topologie*. I (Springer, Berlin, 1923)
70. E. Klarreich, The boundary at infinity of the curve complex and the relative Teichmüller space. Preprint. https://arxiv.org/abs/1803.10339
71. M. Korkmaz, Automorphisms of complexes of curves on punctured spheres and on punctured tori. Topol. Appl. **95**(2), 85–111 (1999)
72. J. Lanier, M. Loving, Centers of subgroups of big mapping class groups and the Tits alternative. Glas. Mat. **55**(1), 85–91 (2020)
73. L. Liu, A. Papadopoulos, Some metrics on Teichmüller spaces of surfaces of infinite type. Trans. Am. Math. Soc. **363**(8), 4109–4134 (2011)
74. F. Luo, Automorphisms of the complex of curves. Topology **39**(2), 283–298 (2000)
75. K. Mann, Automatic continuity for homeomorphism groups and applications. Geom. Topol. **20**(5), 3033–3056 (2016)
76. K. Mann, Automatic continuity for some homeomorphism groups and mapping class groups of non-compact manifolds. Preprint, arXiv:2003.01173
77. K. Mann, K. Rafi, Large scale geometry of big mapping class groups. Preprint, arXiv:1912.10914
78. V. Markovic, Quasisymmetric groups. J. Am. Math. Soc. **19**, 673–715 (2006)
79. H.A. Masur, Y.N. Minsky, Geometry of the complex of curves. I. Hyperbolicity. Invent. Math. **138**(1), 103–149 (1999)
80. H.A. Masur, S. Schleimer, The geometry of the disk complex. J. Am. Math. Soc. **26**(1), 1–62 (2013)

81. K. Matsuzaki, Infinite-dimensional Teichmüller spaces and modular groups, in *Handbook of Teichmüller theory. Vol. IV*. IRMA Lectures in Mathematics and Theoretical Physics, vol. 19 (European Mathematical Society, Zürich, 2014), pp. 681–716
82. A. McLeay, Geometric normal subgroups in mapping class groups of punctured surfaces. New York J. Math. **25**, 839–888 (2019)
83. A. McLeay, Normal subgroups of big mapping class groups. Preprint. Available from https://amcleayblog.wordpress.com/research/
84. E.Y. Miller, The homology of the mapping class group. J. Differ. Geom. **24**(1), 1–14 (1986)
85. I. Morales, F. Valdez, Loxodromic elements in big mapping class groups via the Hooper-Thurston-Veech construction. Preprint
86. D. Mumford, Abelian quotients of the Teichmüller modular group. J. Analyse Math. **18**, 227–244 (1967)
87. P. Patel, N. Vlamis, Algebraic and topological properties of big mapping class groups. Algebr. Geom. Topol. **18**(7), 4109–4142 (2018)
88. A. Phillips, D. Sullivan, Geometry of leaves. Topology **20**(2), 209–218 (1981)
89. W. Pho-On, Infinite unicorn paths and Gromov boundaries. Groups Geom. Dyn. **11**(1), 353–370 (2017)
90. J. Powell, Two theorems on the mapping class group of a surface. Proc. Am. Math. Soc. **68**(3), 347–350 (1978)
91. A. Randecker, Wild translation surfaces and infinite genus. Algebr. Geom. Topol. **18**(5), 2661–2699 (2018)
92. A. Rasmussen, WWPD elements of big mapping class groups. Preprint. https://arxiv.org/abs/1909.06680
93. A. Rasmussen, Uniform hyperbolicity of the graphs of nonseparating curves via bicorn curves. Proc. Am. Math. Soc. **148**(6), 2345–2357 (2020)
94. A.J. Rasmussen, Geometry of the graphs of nonseparating curves: covers and boundaries. Geom. Dedicata 1–34 (2020)
95. M. Reichbach, The power of topological types of some classes of 0-dimensional sets. Proc. Am. Math. Soc. **13**, 17–23 (1962)
96. I. Richards, On the classification of noncompact surfaces. Trans. Am. Math. Soc. **106**, 259–269 (1963)
97. C. Rosendal, Automatic continuity in homeomorphism groups of compact 2-manifolds. Israel J. Math. **166**, 349–367 (2008)
98. C. Rosendal, Automatic continuity of group homomorphisms. Bull. Symb. Log. **15**(2), 184–214 (2009)
99. C. Rosendal, Coarse geometry of topological groups. Preprint of manuscript available at http://homepages.math.uic.edu/~rosendal/PapersWebsite/Coarse-Geometry-Book23.pdf
100. D. Šarić, Train tracks and measured laminations on infinite surfaces (2019). Preprint. arXiv:1902.03437
101. A. Schaffer-Cohen, Automorphisms of the loop and arc graph of an infinite-type surface (2019). Preprint. arXiv:1912.06774
102. S. Schleimer, Notes on the curve complex. Manuscript available from the author's website http://homepages.warwick.ac.uk/~masgar/math.html
103. N. Vlamis, Notes on the topology of big mapping class groups (discussions from an AIM workshop). Manuscript available at https://math.nickvlamis.com/research

Chapter 13
Teichmüller Theory, Thurston Theory, Extremal Length Geometry and Complex Analysis

Hideki Miyachi

Abstract The aim of this chapter is to report on a recent progress of the author's research on Complex analysis on Teichmüller space based on Thurston's theory on surface topology. The main goal is to give a characterization of the pluriharmonic measures and the Poisson kernel (in the sense of Demailly) on the Bers slices via Extremal length geometry.

Keywords Teichmüller space · The Teichmüller distance · Extremal length · Bers slice · Pluricomplex Green function · Pluriharmonic measure

2010 Mathematics Subject Classification 30F60, 30F40, 30F25, 32G15, 31B05, 31B10, 32U05, 32U35

13.1 Introduction

This chapter is devoted to reporting on a recent progress of the author's research on Complex analysis in Teichmüller theory via Extremal length geometry and Thurston's theory with referring a relation in Fig. 13.1 and a dictionary in Fig. 13.2 as research guidelines.

This work is supported by JSPS KAKENHI Grant Numbers 16K05202, partially, 16H03933, 17H02843.

H. Miyachi (✉)
School of Mathematics and Physics, College of Science and Engineering, Kanazawa University, Kanazawa, Japan
e-mail: miyachi@se.kanazawa-u.ac.jp

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_13

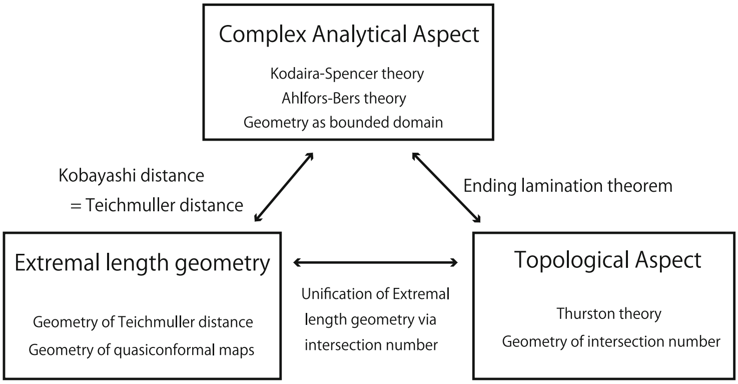


Fig. 13.1 Relations among several aspects in Teichmüller theory

Teichmüller theory	Function theory, Pluripotential theory
Bers slice	hyperconvex domain polynomially convex domain
holomorphic family	equivariant holomorphic map
Teichmüller distance	Kobayashi distance pluricomplex Green function
log-Extremal length	horofunction Busemann cocycle
ratio of Extremal lengths	Kobayashi distance Poisson kernel
Thurston measure	invariant measure of geodesic flow pluriharmonic measure

Fig. 13.2 A dictionary

13.1.1 Background

The complex structure of Teichmüller space was established by Teichmüller [77] (see [78] for an English translation. See also [1, 4]). From Teichmüller's original definition of Teichmüller space, the complex structure includes the definition of the holomorphic families of Riemann surfaces. Teichmüller space as a complex manifold is the universal classifying space of holomorphic fiber spaces of Riemann surfaces (cf. [71]. See also [2, 11]). The infinitesimal complex structure is described by the first cohomology group of the sheaf of holomorphic vector fields (cf. [41]), and by equivalent classes of infinitesimal Beltrami differentials (infinitesimal quasiconformal deformations) (cf. [75]).

Bers [9] realized Teichmüller space complex analytically as a bounded domain in a complex Euclidean space. Hence, Teichmüller space admits a canonical boundary, the so-called *Bers boundary*, defined from the complex analytical view point (cf. [10]), though the boundary does depend on the choice of the base point which is used in the definition of the embedding (cf. [37]). The Bers realization is also comprehended as a part of the deformation space of Kleinian surface groups.

Each marked Riemann surface corresponds to a quasifuchsian group, and the Bers boundary consists of structually unstable Kleinian groups with a (unique) simply connected invariant domain (with respect to allowable representations in the sense of [10]).

From the topological aspect, Teichmüller space is often seen as the orbifold universal covering space of the moduli space of Riemann surfaces. This point of view is useful in the study the mapping class group. For instance, the rational cohomology group of the moduli space is isomorphic to that of the mapping class group (e.g. [30]).

Teichmüller spaces are also used to construct the *boundary at infinity* in the description of degenerations of conformal and hyperbolic structures on surfaces. Actually, Thurston introduced the completion of weighted simple closed curves, the so-called *measured foliation (lamination)* space. This completion yields a boundary at infinity called the *Thurston boundary* and a compactification called the *Thurston compactification*. In Thurston's picture, most of the geometric invariants are interpreted as (extensions of) the geometric intersection number with simple closed curves. With the behavior of the intersection number, projective classes of measured foliations are guides for the description of the degeneration of divergent sequences in Teichmüller space (cf. [25, Exposé 8]).

In his classification of three-dimensional manifolds, Thurston also developed the theory of Kleinian groups. In his program, Thurston poses a famous conjecture, the ending lamination conjecture (cf. [82, Problem 11 in §6]). This was settled by Brock, Canary and Minsky [19, 53] after a breakthrough in the study of the complex of curves by Masur and Minsky [49, 50]. Together with the double limit theorem due to Thurston [83], the ending lamination theorem provides not only a parametrization of most of the Bers boundary by the space of projective classes of measured foliations, but it also a connection between the complex analytical aspect and the topological aspect in Teichmüller theory (cf. [19]. See also Sect. 13.4.5).

Extremal length geometry on Teichmüller space (named after [28]) is the geometry on Teichmüller space studied with the extremal length functions. After the Kerckhoff formula for the Teichmüller distance, Extremal length geometry on Teichmüller space stands for the geometry of Teichmüller space with respect to the Teichmüller distance (cf. [36]). Gardiner and Masur [28] defined a canonical compactification of Teichmüller space in the setting of Extremal length geometry by applying Thurston's compactifying procedure. Extremal length geometry is also thought of as the geometry of the (generalized) intersection number associated with extremal length, and merged with Thurston's framework (cf. [59]). Namely, it is developed as Thurston's theory with extremal length as a background. Royden showed that the Teichmüller distance coincides with the Kobayashi distance under the canonical complex stucture [70]. Royden's observation is a connection between the complex analytical aspect and the Extremal length geometry in Teichmüller theory.

13.1.2 Aim of This Chapter

The main topic in this chapter is to give the *Poisson integral formula* for holomorphic functions and pluriharmonic functions on the Bers compactification (Corollary 13.7.1).

A basic problem behind our research is "*what are holomorphic functions on Teichmüller spaces?*", while each holomorphic mapping *into* Teichmüller space admits a geometric interpretation via Teichmüller's original definition of the complex structure of Teichmüller space. With the help of Extremal length geometry, our Poisson integral formula is expected to strengthen the connection between the topological aspect and the complex analytical aspect in Teichmüller theory, and to encourage the development of a framework of *Complex analysis on Teichmüller space with Thurston's theory*.

In this chapter, for simplicity, we deal with the Teichmüller space of a closed orientable surface of genus g. All the results are valid for finite-dimensional Teichmüller spaces.

13.2 Teichmüller Theory

13.2.1 Teichmüller Space

Let Σ_g be a closed orientable surface of genus $g \geq 2$. A *marked Riemann surface* (X, f) of genus g is a pair of a compact Riemann surface X and an orientation-preserving homeomorphism $f : \Sigma_g \to X$. Two marked Riemann surfaces (X_1, f_1) and (X_2, f_2) are *Teichmüller-equivalent* if there is a biholomorphism $h : X_1 \to X_2$ such that $h \circ f_1$ is homotopic to f_2. The *Teichmüller space* $\mathcal{T}_g$ of Riemann surfaces of genus g is the totality of Teichmüller-equivalence classes of marked Riemann surfaces of genus g. For a comprehensive introduction, we refer the reader to the book [32] by Imayoshi and Taniguchi.

For $K \geq 1$, a *K-quasiconformal mapping* $f : D_1 \to D_2$ between domains D_1, $D_2 \subset \mathbb{C}$ is an orientation preserving homeomorphism whose first distributional partial derivatives are locally in L^2 and satisfies $|f_{\overline{z}}| \leq k|f_z|$ almost everywhere on D_1 where $k = (K-1)/(K+1)$. The infimum $K(f)$ of such K is called the *maximal dilatation* of f. Quasiconformal mappings between Riemann surfaces are canonically defined.

For $x_i = (X_i, f_i) \in \mathcal{T}_g$ $(i = 1, 2)$, we define the *Teichmüller distance* d_T by

$$d_T(x_1, x_2) = \frac{1}{2} \log \inf_h K(h)$$

where h runs over all quasiconformal mappings $h : X_1 \to X_2$ which are homotopic to $f_2 \circ f_1^{-1}$. The Teichmüller distance is complete and it makes $\mathcal{T}_g$ a uniquely geodesic metric space (cf. [76, 79] for an English translation. See also [3]). The mapping class group $\mathrm{MCG}(\Sigma_g)$ of Σ_g acts isometrically on $\mathcal{T}_g$ by

$$\omega(x) = (X, f \circ \omega^{-1}) \tag{13.2.1}$$

for $x = (X, f) \in \mathcal{T}_g$.

13.2.2 Complex Structure

Let X be a closed Riemann surface of genus g. Let $L^\infty(X)$ be the complex Banach space of complex-valued measurable $(-1, 1)$-forms $\mu = \mu(z)(d\overline{z}/dz)$ with the essential supremum norm

$$\|\mu\|_\infty = \operatorname*{ess.sup}_{p\in X} |\mu(p)| < \infty.$$

Let $B(X)$ be the unit ball in $L^\infty(X)$. An element in $B(X)$ is called a *Beltrami differential* on X. For any $\mu \in B(X)$, there is a quasiconformal mapping $f^\mu\colon X \to f^\mu(X)$ satisfying the Beltrami differential equation $\overline{\partial} f^\mu = \mu \partial f^\mu$ on X. Fix an orientation-preserving homeomorphism $h\colon \Sigma_g \to X$. We define a natural projection, called the *Bers projection* with basepoint $x_0 = (X, h) \in \mathcal{T}_g$ by

$$\Phi\colon B(X) \ni \mu \to (f^\mu(X), f^\mu \circ h) \in \mathcal{T}_g$$

which sends the origin $0 \in B(X)$ to $x_0 = (X, h) \in \mathcal{T}_g$. Teichmüller space $\mathcal{T}_g$ admits a complex structure such that the Bers projection is a holomorphic submersion (in fact, the sections of the submersion define holomorphic local charts). This complex structure is unique in the sense that it is independent of the choice of the basepoint $x_0 \in \mathcal{T}_g$ (e.g. [32, §6.2.3]).

A *holomorphic quadratic differential* $q = q(z)dz^2$ on X is a section of the square of the canonical line bundle on X. Let $\mathcal{Q}_X$ be the complex Banach space of holomorphic quadratic differentials q on X with L^1-norm

$$\|q\| = \int_X |q(z)| \frac{\sqrt{-1}}{2} dz \wedge d\overline{z} < \infty.$$

By the Riemann–Roch theorem, the space $\mathcal{Q}_X$ is linearly isomorphic to $\mathbb{C}^{3g-3}$ over $\mathbb{C}$. For $x = (X, f) \in \mathcal{T}_g$, we set $\mathcal{Q}_x = \mathcal{Q}_X$. The complex vector bundle $\mathcal{Q}_g = \cup_{x\in\mathcal{T}_g} \mathcal{Q}_x$ is a holomorphic vector bundle over $\mathcal{T}_g$. A natural pairing

$$B(X) \times \mathcal{Q}_X \ni (\mu, q) \mapsto \langle \mu, q\rangle = \int_X \mu(z) q(z) \frac{\sqrt{-1}}{2} dz \wedge d\overline{z}$$

gives the identification between the holomorphic cotangent space $T^*_x\mathcal{T}_g$ with $\mathcal{Q}_x$. The holomorphic tangent space is described as

$$T_x\mathcal{T}_g \cong L^\infty(X)/\{\mu \in L^\infty(X) \mid \langle \mu, q\rangle = 0,\ \forall q \in \mathcal{Q}_x\}$$

by Teichmüller's lemma (cf. [75, 80] for an English translation. See also [8, 27]). The action (13.2.1) of the mapping class group is a holomorphic action on the Teichmüller space. Conversely, any holomorphic action on Teichmüller space comes from that of the mapping class group (cf. [70]. See also [51, 59]).

13.2.3 Toy Model: The Case of Tori

Let Σ_1 be a (topological) torus. Let $\{[A], [B]\}$ be a pair of generators of the homology group of Σ_1 with $[A] \cdot [B] = 1$ (the algebraic intersection number). The pair $(\Sigma_1, \{[A], [B]\})$ is a marked torus.

The deformation space (Teichmüller space) $\mathcal{T}_1$ of marked tori is identified with the unit disk $\mathbb{D}$ in the following way: Any $\xi \in \mathbb{D}$ corresponds to a marked torus X_ξ which is the quotient space of $\mathbb{C}$ by the marked lattice with ordered pair of generators $\{1, \tau\}$ where $\tau = \tau(\xi) = \sqrt{-1}(1+\xi)/(1-\xi)$. For instance, the marked square torus corresponds to the origin $\xi = 0 \in \mathbb{D}$.

Under the identification $\mathcal{T}_1 \cong \mathbb{D}$, the Teichmüller distance coincides with the hyperbolic distance on $\mathbb{D}$ of curvature -4 (cf. Sect. 13.5.3).

13.3 Thurston's Theory on Surface Topology

13.3.1 Measured Foliations

Let $\mathcal{S} = \mathcal{S}(\Sigma_g)$ be the totality of homotopy classes of non-trivial simple closed curves on Σ_g. The set of formal products $\mathcal{WS} = \mathcal{WS}(\Sigma_g) = \{t\alpha \mid t \geq 0, \alpha \in \mathcal{S}\}$ is called the set of *weighted simple closed curves*. The closure $\mathcal{MF}$ of the embedding

$$\mathcal{WS} \ni t\alpha \mapsto [\beta \to t\, i(\alpha, \beta)] \in \mathbb{R}^{\mathcal{S}}_{\geq 0}$$

is called the space of *measured foliations*, where the function space $\mathbb{R}^{\mathcal{S}}_{\geq 0}$ is topologized with the pointwise convergence topology. For $F \in \mathcal{MF}$, the value $F(\alpha)$ at $\alpha \in \mathcal{S}$ is said to be the *intesection number* between F and α, and denoted by $i(F, \alpha)$.

Notice that measured foliations are originally defined as equivalence classes F of pairs $(\mathcal{F}, \mu)$ of singular foliations and transverse measures to the foliations. For $\alpha \in \mathcal{S}$, the infimum of the integration of the transverse measure along simple closed curves in the homotopy class α defines the intersection number $i(F, \alpha)$ (cf. [25, 81]).

The mapping class group $\mathrm{MCG}(\Sigma_g)$ acts on $\mathcal{MF}$ by

$$i(\omega(F), \alpha) = i(F, \omega^{-1}(\alpha))$$

for $F \in \mathcal{MF}$, $\alpha \in \mathcal{S}$ and $\omega \in \mathrm{MCG}(\Sigma_g)$.

The spaces $\mathcal{MF} = \mathcal{MF}(\Sigma_g)$ and $\mathbb{R}_{\geq 0}^{\mathcal{S}}$ admit actions of the group of positive numbers $\mathbb{R}_{>0}$ by multiplication. The projective space $\mathcal{PMF} = (\mathcal{MF} - \{0\})/\mathbb{R}_{>0}$ is called the space of *projective measured foliations*. Set $\mathcal{PR} = (\mathbb{R}_{\geq 0}^{\mathcal{S}} - \{0\})/\mathbb{R}_{>0}$. The intersection number $i(t\alpha, s\beta) = tsi(\alpha, \beta)$ on $\mathcal{WS} \times \mathcal{WS}$ extends continuously on $\mathcal{MF} \times \mathcal{MF}$ [16, 25, 81]. In this chapter, we say that a measured foliation $F \in \mathcal{MF} - \{0\}$ is *uniquely ergodic* if $i(F, G) = 0$ ($G \in \mathcal{MF} - \{0\}$) implies $G = tF$ for some $t > 0$.

Any measured foliation F is described as an equivalence class of a family of differential forms in a generalized sense as follows. Fix a differential structure on Σ_g. We consider a family $\psi = \{\psi_i\}_i = \{(\psi_i, U_i)\}_i$ with the following conditions.[1] There exist $k_1, \cdots, k_n \in \mathbb{N}$ and $x_1, \cdots, x_n \in \Sigma_g$ such that

(1) $\{U_i\}_i$ is an open cover of $\Sigma_g - \{x_1, \cdots, x_n\}$ and ψ_i is a non-vanishing C^∞ real-valued closed 1-form on each U_i;
(2) $\psi_i = \pm\psi_j$ on $U_i \cap U_j$;
(3) at each x_i, there is a local chart $(u, v)\colon V \to \mathbb{R}^2$ such that for $z = u + \sqrt{-1}v$, $\psi_i = \mathrm{Re}(z^{k_i/2}dz)$ on $V \cap U_i$ for some branch of $z^{k_i/2}$ in $U_i \cap V$.

By definition, the absolute value $|\psi| = |\psi_i|$ on U_i of $\psi = \{\psi_i\}_i$ is well-defined on Σ_g. Each singularity is a zero of the form $|\psi|$. The intersection number between $\psi = \{\psi_i\}_i$ and $\alpha \in \mathcal{S}$ is defined by

$$i(|\psi|, \alpha) = \inf_{\alpha' \sim \alpha} \int_{\alpha'} |\psi|.$$

(cf. [31, Chapter I, §1]). Two such families $\psi_1 = \{\psi_i^1\}_i$ and $\psi_2 = \{\psi_j^2\}_i$ are *measure equivalent* if $|\psi_1| = |\psi_2|$ as functions on $\mathcal{S}$. A representative of a measured foliation is said to be *generic* if all its singularities are represented by a differential form with simple zeros. Any measured foliation admits a generic representative.

The integral curves of the projective vector field on Σ_g representing the kernel of a differential form define a foliation with singularities on Σ_g, and the absolute value of the differential form defines the transverse measure to the foliation. Under this correspondence, measured foliations are described as measure equivalence classes of pairs consisting of foliations with singularities and transverse measures to the foliations.

[1] In the literature (for instance, [28, 31]), ψ itself is sometimes called a measured foliation. However, in our context, measured foliations are equivalence classes. Hence, we do not call this a measured foliation here to avoid confusion.

13.3.2 Measured Laminations

Suppose $g \geq 2$ and fix a hyperbolic structure on Σ_g. A *geodesic lamination* is a closed subset of Σ_g consisting of disjoint simple complete geodesics. A *transverse measure* of a geodesic lamination is an assignment of a Radon measure to each transverse arc to the geodesic lamination. A *measured lamination* $L = (|L|, \mu)$ is a pair consisting of a geodesic lamination $|L|$, which is called the *support* of L, and a transverse measure μ to the geodesic lamination [22, 81]. Any measured lamination L is assumed to have *full support* in the sense that for any transverse arc I to $|L|$, the support of the Radon measure assigned to I coincides with $|L| \cap I$. We denote by $\mathcal{ML} = \mathcal{ML}(\Sigma_g)$ the totality of measured laminations on Σ_g. Any weighted simple closed curve $t\alpha \in \mathcal{WS}$ is associated with a measured lamination whose support is the geodesic representative of α and the transverse measure is t times the counting measure for the intersection with the geodesic representative.

The *intersection number* between a measured lamination $L = (L, \mu)$ and $\alpha \in \mathcal{S}$ is defined by

$$i(L, \alpha) = \inf_{\alpha' \sim \alpha} \int_{\alpha'} d\mu.$$

There is a natural bijection between $\mathcal{MF}$ and $\mathcal{ML}$ such that $F \in \mathcal{MF}$ corresponds to $L \in \mathcal{ML}$ if and only if

$$i(F, \alpha) = i(L, \alpha) \tag{13.3.1}$$

for all $\alpha \in \mathcal{S}$ (cf. [68, §1.7]). The space of measured laminations is also topologized with the pointwise convergence of the intersection number, and the bijective correspondence $\mathcal{MF} \cong \mathcal{ML}$ becomes a homeomorphism under this topology. The space $\mathcal{PML}$ of projective measured laminations is defined in the same way, and is homeomorphic to $\mathcal{PMF}$. Thurston showed that $\mathcal{MF}$ (and hence $\mathcal{ML}$) is homeomorphic to $\mathbb{R}^{6g-6}$ when $g \geq 2$ and to $\mathbb{R}^2$ when $g = 1$, and $\mathcal{PMF}$ (and hence $\mathcal{PML}$) is homeomorphic to the sphere of dimension $6g-7$ if $g \geq 2$ and to the circle if $g = 1$ (cf. [25, 81]). The space $\mathcal{ML}$ (and hence $\mathcal{MF}$) admits a canonical piecewise (integral) linear structure inherited from the transverse measures or the intersection number (cf. [81, Proposition 9.5.8]. See also [17, Part II] and [69, §1]).

13.3.3 Thurston's Measure

The space of measured foliations (and hence measured laminations) admits a canonical ergodic measure, called the *Thurston measure*, equivariant under the action of Mapping class group. The following definion of the measure is due to Masur [48, §4].

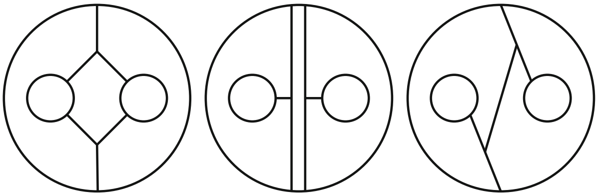


Fig. 13.3 Three patterns of singularities in pairs of pants

Let $\{\gamma_i\}_{i=1}^{3g-3}$ be a pants decomposition of Σ_g such that no γ_i occurs twice on the boundary of a pair of pants. Let $F \in \mathcal{MF}$. A pair of pants D defined from $\{\gamma_i\}_{i=1}^{3g-3}$ is said to be of *Case I* for F if the intersection number of F with each boundary curve of D is less than the sum of the other two, of *Case II* if the intersection number of F with one boundary curve of D is larger than the sum of the other two, and of *Case III* if the intersection number of F with one boundary component is equal to the sum of the other two. Case III is a limiting case of cases I and II (cf. Fig. 13.3. See also [48, Figure in p. 177]).

A subset $U \subset \mathcal{MF}$ is said to be of *constant type* if all the γ_i are transverse for any $F \in U$ and if $F_1, F_2 \in U$ and D is any pair of paths with respect to $\{\gamma_i\}_i$, then D is either of Case I for both F_1 and F_2 or Case II for both. $\mathcal{MF}$ is a disjoint union of domains of constant type together with lower-dimensional sets where some γ_i is a leaf or some pair of pants is of Case III. Such a U is an open set on $\mathcal{MF}$ because the intersection number $\mathcal{MF} \ni F \mapsto i(F, \gamma_i)$ is continuous for $i = 1, \cdots, 3g-3$.

Let $U \subset \mathcal{MF}$ be a domain of constant type. For $F \in U$, let $\Pi\colon \tilde{\Sigma} \to \Sigma_g$ be the orientation double covering ramified over the singularities of F such that the pullback $\Pi^* F$ is given by a closed 1-form $\tilde{\psi}_F$ on $\tilde{\Sigma}$. The singularities of the normal forms (in the sense of [25, Exposé 6]) of measured foliations in U with respect to $\{\gamma_i\}_{i=1}^{3g-3}$ are all simple and contained in pairs of pants. The combinatorial pattern of singularities and singular leaves on each pair of pants does not change when the measured foliations vary in U (cf. Fig. 13.3). Hence, up to isotopy, we can take a common double branched covering $\Pi\colon \tilde{\Sigma} \to \Sigma_g$ for all $F \in U$.

Fix a basis $\alpha_1, \cdots, \alpha_{6g-6}$ for the odd homology group $H_1(\tilde{\Sigma}, \mathbb{Z})^-$ with respect to the action of the canonical involution $\tilde{\Sigma} \to \tilde{\Sigma}$ of the covering $\Pi\colon \tilde{\Sigma} \to \Sigma_g$. The cohomology class of the 1-form $\tilde{\psi}_F$ on $\tilde{\Sigma}$ defined from $F \in U$ is odd with respect to the involution, and is recognized as an element in the dual of the odd homology group $H_1(\tilde{\Sigma}, \mathbb{Z})^-$. Then, for $A \subset U$, the *Thurston measure* $\mu_{Th}(A)$ of A is defined by the Lebesgue measure of the image

$$\{(\psi_F(\alpha_1), \cdots, \psi_F(\alpha_{6g-6})) \mid F \in A\} \subset \mathbb{R}^{6g-6}.$$

We define μ_{Th} to be zero on the complement of all such U and thus μ_{Th} is defined on $\mathcal{MF}$. The Thurston measure μ_{Th} is defined independently of the choice of the basis $\{\alpha_1, \cdots, \alpha_{6g-6}\}$ and it is invariant under the action of mapping class group (cf. [48, Lemmas 4.2 and 4.3]). The Thurston measure defined here coincides with the volume form of the Thurston symplectic form up to constant multiple (e.g. [54, 68]).

13.3.4 Toy Model: The Case of Tori

We shall frequently use the notions given in Sect. 13.2.3. For $p/q \in \hat{\mathbb{Q}} \cup \{\infty\}$, the *$p/q$-curve* $\gamma_{p/q}$ on Σ_1 is, in our convention, an (unoriented) simple closed curve representing $\pm(p[A]-q[B])$. The geometric intersection number between the p/q-curve and the r/s-curve is equal to $|ps - rq|$ (cf. [25, Exposé 1]). After identifying $(\Sigma_1, \{[A], [B]\})$ and a marked square torus X_0, the measured foliation associated with the p/q-curve is defined by a 1-form $\psi_{p/q} = qdu + pdv$ (where $u + iv$ is a standard Euclidean coordinate on $X_0 = \mathbb{C}/\langle z + 1, z + \sqrt{-1}\rangle$). Consider a mapping

$$\mathcal{WS} \ni t\gamma_{p/q} \mapsto [tq, tp] = \left[t\psi_{p/q}(A), t\psi_{p/q}(B)\right] \in \mathbb{R}^2/\mathbb{Z}_2 \tag{13.3.2}$$

where $\mathbb{Z}_2$ is the group of automorphisms of $\mathbb{R}^2$ generated by the π-rotation with respect to the origin. By taking the closure of the image, we get the identification $\mathcal{MF} \cong \mathbb{R}^2/\mathbb{Z}_2 \cong \mathbb{R}^2$. For $\mathfrak{a} \in \mathcal{MF} \cong \mathbb{R}^2/\mathbb{Z}_2$, we denote by $F_{\mathfrak{a}}$ the measured foliation corresponding to $\mathfrak{a}$. Geometrically, the measured foliation $F_{[a,b]}$ is associated with the flow defined by the differential form $\psi_{[a,b]} = adu + bdv$ and the corresponding foliation consists of lines in the direction $(-b, a)$. The measured foliation $F_{[a,b]}$ is characterized by the equation

$$i(F_{[a,b]}, F_{[q,p]}) = |ap - bq|$$

for $p/q \in \hat{\mathbb{Q}}$. From (13.3.2), the identification between $\mathcal{MF}$ and $\mathbb{R}^2/\mathbb{Z}_2$ is concretely obtained by

$$\mathcal{MF} \ni F_{[a,b]} \mapsto [a, b] = \left[\psi_{F_{[a,b]}}(A), \psi_{F_{[a,b]}}(B)\right] \in \mathbb{R}^2/\mathbb{Z}_2. \tag{13.3.3}$$

The Thurston measure μ_{Th} on $\mathcal{MF}$ is a Borel measure on $\mathcal{MF}$ which is the pullback of the Lebesgue measure on $\mathbb{R}^2/\mathbb{Z}_2$ via the mapping (13.3.3).

The space $\mathcal{PMF} = \mathcal{PMF}(\Sigma_1) = (\mathcal{MF} - \{0\})/\mathbb{R}_{>0}$ of projective measured foliations is identified with the circle $\mathbb{S}^1 = \{|\xi| = 1\} \subset \mathbb{C}$ via the projection

$$\mathcal{MF} \ni [a, b] \mapsto e^{\sqrt{-1}\Theta([a,b])} = \tau^{-1}(b/a) = (b - \sqrt{-1}a)/(b + \sqrt{-1}a) \in \mathbb{S}^1,$$

where the function Θ satisfies $0 < \Theta([a, b]) \leq 2\pi$ for all $[a, b] \in \mathcal{PMF}$ and $\Theta([0, 1]) = 2\pi$.

13.4 Thurston's Theory on Kleinian Surface Groups

In this section, we recall Thurston's picture of the theory of Kleinian surface groups. We also discuss the Bers slice with sophisticated results in the theory of Kleinian surface groups.

13.4.1 Kleinian Groups

A discrete subgroup Γ in $\mathrm{PSL}_2(\mathbb{C}) \cong \mathrm{Isom}^+(\mathbb{H}^3)$ is called a *Kleinian group*. The *limit set* $\Lambda_\Gamma \subset \hat{\mathbb{C}}$ of a Kleinian group Γ is, by definition, the set of accumulation points of the orbit of Γ of a point in $\mathbb{H}^3$. A Kleinian group Γ is called *non-elementary* if Λ_Γ contains at least 3 points, *elementary* otherwise. Any Kleinian group in this chapter is assumed to be non-elementary unless otherwise specified. The complement Ω_Γ of Λ_Γ in $\hat{\mathbb{C}}$ is called the *region of discontinuity* of Γ. A Kleinian group Γ is called a *Fuchsian group* if the limit set Λ_Γ of Γ is a round circle in $\hat{\mathbb{C}}$.

13.4.2 Quasiconformal Deformations

A measurable function μ on $\hat{\mathbb{C}}$ is said to be *invariant* under a Kleinian group Γ if $\mu \circ \gamma(\overline{\gamma'}/\gamma') = \mu$ on $\hat{\mathbb{C}}$ for all $\gamma \in \Gamma$. Let $L^\infty(\Gamma)$ be the complex Banach space of invariant bounded measurable functions under Γ with the essential supremum norm. Let $B(\Gamma)$ be the unit ball in $L^\infty(\Gamma)$. For any $\mu \in B(\Gamma)$, there is a quasiconformal mapping w^μ (which is unique up to pre-composing with Möbius transformations) such that $\overline{\partial} w^\mu = \mu \partial w^\mu$. Then, the conjugation $\Gamma^\mu = w^\mu \Gamma (w^\mu)^{-1}$ also becomes a Kleinian group. We call Γ^μ a *quasiconformal deformation* of Γ. A *quasifuchsian group* is, by definition, a quasiconformal deformation of a Fuchsian group.

13.4.3 Classification of Marked Kleinian Surface Groups

A *Kleinian surface group* is, by defintion, a Kleinian group which is isomorphic to the fundamental group of a compact surface (a type-preserving condition is needed if the surface has boundary. For instance, see [53]). A Kleinian surface group is said to be *marked* if it is assigned a homomorphism from the fundamental group of a fixed compact surface.

Let $\rho\colon \pi_1(\Sigma_g) \to \mathrm{PSL}(2, \mathbb{C})$ be a faithful discrete representation. An *accidental parabolic transformation* (APT) of ρ is an element $\gamma \in \pi_1(\Sigma_g)$ or the image $\rho(\gamma)$ such that $\rho(\gamma)$ is parabolic (cf. [10]).

13.4.3.1 End Invariants

Bonahon–Thurston's tameness theorem asserts that the representation ρ is induced from a homeomorphism from the quotient manifold $\mathcal{N}_\rho = \mathbb{H}^3/\mathrm{Im}(\rho)$ to the product manifold $\Sigma_g \times \mathbb{R}$ (cf. [16, 81]).

Suppose that ρ admits no APT. The quotient manifold $\mathcal{N}_\rho$ has two *ends* e_+ and e_- (in the sense of Freudenthal) corresponding to $\Sigma_g \times \{t > 0\}$ and $\Sigma_g \times \{t < 0\}$. An end e_s ($s = \pm$) is said to be *geometrically infinite* if any neighborhood of the end e_s intersects a closed geodesic. An end which is not geometrically infinite is called *geometrically finite*.

Any geometrically infinite end of $\mathcal{N}_\rho$ is *simply degenerate* in the sense that there is a sequence $\{\gamma_n\}_n$ of simple closed curves on $\Sigma_g = \Sigma_g \times \{0\} \subset \Sigma_g \times \mathbb{R} \cong \mathcal{N}_\rho$ such that the geodesic representative of γ_n in $\mathcal{N}_\rho$ exits the end e_s (cf. [16, 81]). Fix a hyperbolic structure on Σ_g and consider the sequence of geodesics $\{\gamma_n^*\}_n$ representing such a $\{\gamma_n\}_n$. The *ending lamination* of the geometrically infinite end e_s is a geodesic lamination which is the support of the accumulation points of $\{\gamma_n^*\}_n$ in $\mathcal{PML}$. The ending lamination is well-defined and has a minimal and filling property saying that any leaf of $|L|$ is dense and any measured lamination L whose suppost is the ending lamination of some faithful discrete representation without APT satisfies that $i(L, \alpha) \neq 0$ for all $\alpha \in \mathcal{S}$ (cf. [81, Proposition 9.3.8]).

When the end e_s is geometrically finite, there is a component Ω_s of $\Omega_{\mathrm{Im}(\rho)}$ such that the union $\mathcal{N}_\rho \cup (\Omega_s/\mathrm{Im}(\rho))$ defines a compactification of the end e_s. The quotient surface $\Omega_s/\mathrm{Im}(\rho)$ has a conformal structure inherited from $\hat{\mathbb{C}}$ (the orientation is inherited from the product structure of $\mathcal{N}_\rho \cong \Sigma_g \times \mathbb{R}$), and a marking induced from ρ. Therefore, we can associate to a geometrically finite end e_s a point in $\mathcal{T}_g$, which is called the *Teichmüller end invariant* of the geometrically finite end e_s. Teichmüller end invariants and ending laminations for the ends of $\mathcal{N}_\rho$ are well-defined for the representation ρ, and called the *end invariants* of ρ.

We could also define end invariants for general faithful-discrete representations of $\pi_1(S)$. However, we omit this definition because we do not need the general situation in this chapter (e.g. [53, §2.2]).

13.4.3.2 Ending Lamination Theorem

The *ending lamination theorem* asserts that faithful discrete $\mathrm{PSL}(2, \mathbb{C})$-representions of $\pi_1(S)$ are classified by the end invatiants [19, 53]. Furthermore, with Thurston's double limit theorem [83], any Kleinian surface group is obtained as the algebraic limit of a sequence of marked quasifuchsian groups. This is known as the *density theorem* (cf. [19]. See also [18, 65, 66] for the general case).

13.4.4 Bers Slice

Let Γ be a Fuchsian group isomorphic to $\pi_1(\Sigma_g)$ and satisfying $\Lambda_\Gamma = \hat{\mathbb{R}} = \mathbb{R}\cup\{\infty\}$. Let $A_2(\mathbb{H}^*, \Gamma)$ be the space of holomorphic functions φ on the lower-half plane $\mathbb{H}^*$ which satisfy $\varphi(\gamma(z))(\gamma'(z))^2 = \varphi(z)$ for all $z \in \mathbb{H}^*$ and $\gamma \in \Gamma$ and the function $z \mapsto \mathrm{Im}(z)^2|\varphi(z)|$ is bounded on $\mathbb{H}^*$.

Let $X_0 = \mathbb{H}/\Gamma$ and set $\pi : \mathbb{H} \to X_0$ to be the universal covering projection. Fix a marking $f_0 \colon \Sigma_g \to X_0$ and set $x_0 = (X_0, f_0) \in \mathcal{T}_g$. For $\mu \in B(X_0)$, by the measurable Riemann mapping theorem [5], there is a unique quasiconformal mapping w^μ on $\mathbb{C}$ satisfying

$$\frac{\overline{\partial} w^\mu}{\partial w^\mu} = \begin{cases} \pi^*(\mu) & \text{on } \mathbb{H} \\ 0 & \text{on } \mathbb{C} - \mathbb{H} \end{cases}$$

and $w^\mu(0) = w^\mu(1) - 1 = 0$. Then, a complex analytic mapping

$$B(X_0) \ni \mu \mapsto Sch(w^\mu \mid_{\mathbb{H}^*}) \in A_2(\mathbb{H}^*, \Gamma)$$

descends to a complex analytic embedding $\mathcal{T}_g \to A_2(\mathbb{H}^*, \Gamma)$, where $Sch(h)$ stands for the Schwarzian derivative of a holomorphic function h (cf. [9]). The embedding is called the *Bers embedding*, and the image $\mathcal{T}_{x_0}^B$ is called the *Bers slice* (e.g. [32, §6.1]).

In the theory of Kleinian groups, the Bers slice is described as follows. For $\varphi \in A_2(\mathbb{H}^*, \Gamma)$, there is a locally univalent holomorphic mapping W_φ on $\mathbb{H}^*$ such that the Schwarzian derivative of W_φ is equal to φ. The *holonomy representation* ρ_φ is a homomorphism $\Gamma \to \mathrm{PSL}(2, \mathbb{C})$ satisfying $\rho_\varphi(\gamma) \circ W_\varphi = W_\varphi \circ \gamma$ on $\mathbb{H}^*$ for all $\gamma \in \Gamma$. By fixing an identification $\pi_1(\Sigma_g) \cong \Gamma$ (in other words, we fix a marking $\Sigma_g \to X$), we also think of ρ_φ as a $\mathrm{PSL}(2, \mathbb{C})$-representation of $\pi_1(\Sigma_g)$. The local univalent holomorphic mapping W_φ is regarded as the developing mapping for the (marked) projective structure with holonomy ρ_φ (e.g. [21, 73]).

When $\mu \in B(X_0)$ satisfies $Sch(w^\mu \mid_{\mathbb{H}^*}) = \varphi$, W_φ coincides with the composition of a Möbius transformation and the restriction $w^\mu \mid_{\mathbb{H}^*}$. In fact, the Bers slice $\mathcal{T}_{x_0}^B$ is equal to the set of $\varphi \in A_2(\mathbb{H}^*, \Gamma)$ such that W_φ admits a quasiconformal extension on $\hat{\mathbb{C}}$. The closure of $\mathcal{T}_{x_0}^B$ in $A_2(\mathbb{H}^*, \Gamma)$ is known as the *Bers compactification* of $\mathcal{T}_g$. The boundary $\partial\mathcal{T}_{x_0}^B$ is called the *Bers boundary* of $\mathcal{T}_g$ (cf. [10]). Each point of the Bers compactification defines a marked Kleinian surface group $\Gamma_\varphi = \mathrm{Im}(\rho_\varphi)$ with isomorphism $\rho_\varphi \colon \pi_1(\Sigma_g) \cong \Gamma \to \Gamma_\varphi$. To the negative end of the marked Kleinian surface group ρ_φ of any φ in the Bers compactification of $\mathcal{T}_{x_0}^B$, the Teichmüller end invariant $x_0 = (X_0, f_0) \in \mathcal{T}_g$ is assigned. From the ending lamination theorem and the double limit theorem, any $\varphi \in A_2(\mathbb{H}^*, \Gamma)$ with univalent W_φ is in the Bers compactification of $\mathcal{T}_{x_0}^B$ (see also [20] for a different approach).

Kerckhoff and Thurston [37] observed that the Bers compactification is dependent on the choice of the base point in the sense that for another $x_1 \in \mathcal{T}_g$, the natural biholomorphic mapping $\mathcal{T}_{x_0}^B \to \mathcal{T}_{x_1}^B$ induced by the identity mapping on $\mathcal{T}_g$ does not extend homeomorphically to the Bers closures. On the other hand, as we will mention in Sect. 13.4.5, the natural biholomorphic mapping extends homeomorphically "almost everywhere" to the Bers boundaries (cf. [67]. See also [12]).

13.4.5 Structure of the Bers Boundary

In this section, we recall a topological parametrization of a part of the Bers boundary via the ending lamination theorem.

13.4.5.1 Complex of Curves and the Gromov Boundary

The *complex of curves* $\mathcal{C}_g$ is a simplicial complex where any k-simplex is an unordered sequence $[\alpha_0, \cdots, \alpha_k]$ of homotopy classes of simple closed curves on Σ_g such that $i(\alpha_i, \alpha_j) = 0$ and $\alpha_i \neq \alpha_j$ for $0 \leq i < j \leq k$. After identifying each k-simplex with the Euclidean standard k-simplex, the complex of curves becomes a metric space. Masur and Minsky [49] showed that the complex of curves is Gromov hyperbolic.

The *Gromov boundary* $\partial\mathcal{C}_g$ of $\mathcal{C}_g$ is described as follows. A measured foliation (lamination) F is said to be *minimal* if $i(F, \alpha) \neq 0$ for all $\alpha \in \mathcal{S}$. Two measured foliations (laminations) F_1 and F_2 are said to be *topologically equivalent* if the underlying foliations are equivalent with respect to isotopy and Whitehead moves. Klarreich [38] showed that the Gromov boundary $\partial\mathcal{C}_g$ is identified with the set of topological equivalence classes of minimal foliations (see also [29]).

13.4.5.2 Boundary Groups Without APTs

Let $x_0 \in \mathcal{T}_g$. Let $\partial^{APT}\mathcal{T}_{x_0}^B$ be the subset of the Bers boundary which consists of boundary groups with APT. Let $\partial^{min}\mathcal{T}_{x_0}^B = \partial\mathcal{T}_{x_0}^B - \partial^{APT}\mathcal{T}_{x_0}^B$. By virtue of the ending lamination theorem and the Thurston double limit theorem, we have the following homeomorphism

$$\Phi\colon \partial\mathcal{C}_g \to \partial^{min}\mathcal{T}_{x_0}^B \tag{13.4.1}$$

which assigns $[F] \in \partial\mathcal{C}_g$ to the boundary group whose ending lamination is the support of the measured lamination corresponding to F under the identification (13.3.1) (cf. [44]). The boundary $\partial\mathcal{C}_g$ contains a subset $\partial^{ue}\mathcal{C}_g$ consisting of topological equivalence classes of minimal uniquely ergodic measured foliations. Let $\partial^{ue}\mathcal{T}_{x_0}^B$ be the image of $\partial^{ue}\mathcal{C}_g$ under the identification (13.4.1).

13.4.6 The Case of Once-Punctured Tori

Minsky [52] solved the ending lamination conjecture for once-punctured torus groups before solving the general case. Applying the ending lamination theorem, he showed that the Bers slice for once-punctured tori is homeomorphic to the Thurston compactification of the Teichmüller space of once-puncture tori. Namely, the boundary groups are parametrized by $\mathcal{PMF}$. It is well-known that any measured foliation on the once-punctured torus is either rational or minimal uniquely ergodic (e.g. [14, 35]).

13.5 Extremal Length and Thurston Measures on $\mathcal{PMF}$

13.5.1 Hubbard–Masur Differentials and Extremal Length

Let $x = (X, f) \in \mathcal{T}_g$. For $q \in \mathcal{Q}_x$, the *vertical foliation* $v(q) \in \mathcal{MF}$ of q is defined by

$$i(v(q), \alpha) = \inf_{\alpha' \sim f(\alpha)} \int_{\alpha'} |\mathrm{Re}\sqrt{q}|$$

for $\alpha \in \mathcal{S}$. Hubbard and Masur [31] observe that the mapping

$$\mathcal{Q}_x \ni q \mapsto v(q) \in \mathcal{MF}$$

is a homeomorphism. For $F \in \mathcal{MF}$, the *Hubbard–Masur differential* $q_{F,x} \in \mathcal{Q}_x$ is defined by the relation $v(q_{F,x}) = F$. The *extremal length* of $F \in \mathcal{MF}$ on $x \in \mathcal{T}_g$ is defined by the L^1-norm

$$\mathrm{Ext}_x(F) = \|q_{F,x}\|.$$

We can also consider the *extremal length* for measured laminations by the correspondence (13.3.1). The extremal length satisfies that $\mathrm{Ext}_x(tF) = t^2\mathrm{Ext}_x(F)$ for $F \in \mathcal{MF}$ and $t \geq 0$, and that $\mathcal{T}_g \times \mathcal{MF} \ni (x, F) \mapsto \mathrm{Ext}_x(F)$ is continuous.

Kerckhoff [36] observed that for any $x, y \in \mathcal{T}_g$,

$$d_T(x, y) = \frac{1}{2} \log \sup_{\alpha \in \mathcal{S}} \frac{\mathrm{Ext}_x(\alpha)}{\mathrm{Ext}_y(\alpha)}. \tag{13.5.1}$$

The formula (13.5.1) is called the *Kerckhoff formula* of the Teichmüller distance.

13.5.2 Thurston Measures on $\mathcal{PMF}$

Fix a point $x \in \mathcal{T}_g$. After [48] (see also [6, §2.3]), we define a measure ν^x_{Th} on $\mathcal{PMF}$, which we also call the *Thurston measure on* $\mathcal{PMF}$ *with base point* x here, by

$$\nu^x_{Th}(E) = \frac{1}{C_{HM}}\mu_{Th}(\{F \in \mathcal{MF} \mid [F] \in E, \mathrm{Ext}_x(F) \leq 1\}), \tag{13.5.2}$$

where the bracket $[F]$ means the projective class of $F \in \mathcal{MF}$ and C_{HM} is the *Hubbard–Masur constant* discussed in [6, §2.3] and [26, §5.7]. The multiplication of the reciprocal of the Hubbard–Masur constant in the right-hand side of (13.5.2) means that the Thurston measure ν^x_{Th} defined here is normalized in such a way that the total mass of $\mathcal{PMF}$ is equal to 1 for each $x \in \mathcal{T}_g$.

13.5.3 Toy Model: The Case of Tori

The marking determines the p/q-curve on each X_ξ. The Hubbard–Masur differential on X_ξ with respect to $F_{[a,b]} \in \mathcal{MF}$ is

$$q_{[a,b],\xi} = -\left(\frac{-b + a\overline{\tau(\xi)}}{\mathrm{Im}\tau(\xi)}\right)^2 dz^2.$$

(cf, Sect. 13.2.3). The extremal length of $F_{[a,b]}$ on X_ξ is

$$\mathrm{Ext}_{X_\xi}(F_{[a,b]}) = \|q_{[a,b],\xi}\| = \frac{|-b + a\overline{\tau(\xi)}|^2}{\mathrm{Im}\tau(\xi)} = (a^2+b^2)\frac{|\xi - e^{\sqrt{-1}\Theta([a,b])}|^2}{1-|\xi|^2}. \tag{13.5.3}$$

The Kerckhoff formula asserts that the Teichmüller distance $d_{\mathcal{T}_1}$ on $\mathcal{T}_1$ is

$$\begin{aligned} d_{\mathcal{T}_1}(X_\xi, X_\eta) &= \frac{1}{2}\log \sup_{p/q \in \hat{\mathbb{Q}}} \frac{\mathrm{Ext}_{X_\xi}(F_{[-p,q]})}{\mathrm{Ext}_{X_\eta}(F_{[-p,q]})} \\ &= \frac{1}{2}\log \max_{0\leq\Theta\leq 2\pi} \frac{|\xi - e^{\sqrt{-1}\Theta}|^2}{1-|\xi|^2}\frac{1-|\eta|^2}{|\eta - e^{\sqrt{-1}\Theta}|^2} \\ &= \frac{1}{2}\log\frac{|1-\overline{\xi}\eta| + |\xi-\eta|}{|1-\overline{\xi}\eta| - |\xi-\eta|}, \end{aligned}$$

which coincides with the hyperbolic distance on $\mathbb{D}$ with curvature -4 (cf. [7]).

We calculate the Thurston measure ν^{ξ}_{Th} for $\xi \in \mathbb{D} \cong \mathcal{T}_1$. Let B_ξ be the unit ball with respect to the extremal length $\mathrm{Ext}_{X_\xi}(\cdot)$ in $\mathcal{MF} \cong \mathbb{R}^2/\mathbb{Z}_2$. Take $\Theta_\xi \in (0, 2\pi]$ with $e^{\sqrt{-1}\Theta_\xi} = \xi/|\xi|$. From (13.5.3), B_ξ is the unit disk if $\xi = 0$, and the ellipse whose major axis is a segment of length $2(1+|\xi|)/(1-|\xi|)$ of direction Θ_ξ, and whose minor axis is a segment of length $2(1-|\xi|)/(1+|\xi|)$ of direction $\Theta_\xi + \pi/2$, otherwise. The identification $\mathcal{PMF} \cong \mathbb{S}^1 \to \partial B_\xi \subset \mathcal{MF} \cong \mathbb{R}^2/\mathbb{Z}_2$ is given by

$$\mathbb{S}^1 \ni e^{\sqrt{-1}\Theta} \mapsto \left[\frac{\sqrt{1-|\xi|^2}}{|\xi - e^{\sqrt{-1}\Theta}|}\sin\frac{\Theta}{2}, -\frac{\sqrt{1-|\xi|^2}}{|\xi - e^{\sqrt{-1}\Theta}|}\cos\frac{\Theta}{2}\right].$$

Hence, the Thurston measure ν^{ξ}_{Th} for X_ξ is

$$\nu^{\xi}_{Th}(E) = \frac{1}{2\pi}\int_E \frac{1-|\xi|^2}{|\xi - e^{\sqrt{-1}\Theta}|^2} d\Theta \tag{13.5.4}$$

for $E \subset \mathcal{PMF} \cong \mathbb{S}^1$, from the normalization of the Thurston measures.

13.6 Thurston Theory with Extremal Length

13.6.1 Gardiner–Masur Compactfication

The *Gardiner–Masur compactification* is defined by Gardinar and Masur in [28] as the closure of the image of the embedding

$$\mathcal{T}_g \ni x \mapsto [\mathcal{S} \ni \alpha \mapsto \mathrm{Ext}_x(\alpha)^{1/2}] \in \mathcal{PR} = (\mathbb{R}^{\mathcal{S}}_{\geq 0} - \{0\})/\mathbb{R}_{>0}. \tag{13.6.1}$$

The boundary $\partial_{GM}\mathcal{T}_g$ of the compactification is called the *Gardiner–Masur boundary*. The Gardiner–Masur boundary contains the space $\mathcal{PMF}$ of projective measured foliations (cf. [28, Theorem 7.1]) as a proper subset.

In [56], the author observed that when fixing $x_0 \in \mathcal{T}_g$, for any $p \in \mathcal{T}_g \cup \partial_{GM}\mathcal{T}_g$, there is a unique continuous function $\mathcal{E}_p$ on $\mathcal{MF}$ such that the restriction of $\mathcal{E}_p$ to $\mathcal{S}$ represents p in the closure of the image of (13.6.1), $\mathcal{E}_{p_n} \to \mathcal{E}_p$ when $p_n \to p$ in $\mathcal{T}_g \cup \partial_{GM}\mathcal{T}_g$, and $\max_{\mathrm{Ext}_{x_0}(F)=1} \mathcal{E}_p(F) = 1$. For instance,

$$\mathcal{E}_x(F) = e^{-d_T(x_0,x)}\mathrm{Ext}_x(F)^{1/2} \tag{13.6.2}$$

$$\mathcal{E}_{[G]}(F) = \frac{i(F,G)}{\mathrm{Ext}_{x_0}(G)^{1/2}} \tag{13.6.3}$$

for $x \in \mathcal{T}_g$ and $[G] \in \mathcal{PMF} \subset \partial_{GM}\mathcal{T}_g$ (see also [59, §3]).

13.6.2 Thurston's Theory with Extremal Length

Let $\tilde{\partial}_{GM}$ be the preimage of the projection $\mathbb{R}^{\mathcal{S}}_{\geq 0} \to \mathcal{PR}$ of the Gardiner–Masur boundary. Notice from Gardiner and Masur's observation that $\mathcal{MF} \subset \tilde{\partial}_{GM}$. The extermal length function $\mathrm{Ext}_x(\cdot)$ on $\mathcal{MF}$ extends continuously on $\tilde{\partial}_{GM}$, and the *intersection number* $i(\,\cdot\,,\,\cdot\,)$ on $\mathcal{MF} \times \mathcal{MF}$ also extends continuously on $\tilde{\partial}_{GM} \times \tilde{\partial}_{GM}$ (cf. [59, Theorems 1 and 3]). Using the intersection number on $\tilde{\partial}_{GM}$, the formula (13.6.3) is extended as

$$\mathcal{E}_p(F) = \frac{i(F, \mathfrak{a})}{\mathrm{Ext}_{x_0}(\mathfrak{a})^{1/2}}$$

where $\mathfrak{a} \in \tilde{\partial}_{GM} - \{0\}$ is a representative of p. The *Gromov product with base point* $x_0 \in \mathcal{T}_g$

$$\langle x \mid y \rangle_{x_0} = \frac{1}{2}(d_T(x_0, x) + d_T(x_0, y) - d_T(x, y))$$

is recognized as the log-intersection number between marked Riemann surfaces in the sense that

$$\exp(-2\langle x_n \mid y_n \rangle_{x_0}) \to \frac{i(\mathfrak{a}, \mathfrak{b})}{\mathrm{Ext}_{x_0}(\mathfrak{a})^{1/2}\mathrm{Ext}_{x_0}(\mathfrak{b})^{1/2}}$$

as $n \to \infty$ for $x \in \mathcal{T}_g$, when the sequences $\{x_n\}_n$ and $\{y_n\}_n$ in $\mathcal{T}_g$ converge to the projective classes of $\mathfrak{a}$ and $\mathfrak{b} \in \tilde{\partial}_{GM} - \{0\}$, respectively (cf. [59, Corollary 1]).

13.6.3 Toy Model: The Case of Tori

The *Gardiner–Masur compactification* of $\mathcal{T}_1$ was already discussed in [55]. Indeed, the Gardiner–Masur compactification coincides with the Thurston compactification.[2] For the convenience of the readers, we discuss briefly this coincidence. Consider a function

$$\begin{aligned}\mathcal{E}_\xi(F_{[a,b]}) &= e^{-d_{\mathcal{T}(\Sigma_1)}(0,\xi)}\mathrm{Ext}_{X_\xi}(F_{[a,b]})^{1/2} \\ &= \sqrt{a^2+b^2}\,\frac{|\xi - e^{\sqrt{-1}\Theta([a,b])}|}{1+|\xi|}\end{aligned}$$

[2]We need to define with care the Thurston compactification of the Teichmüller space of flat tori, since there is no hyperbolic structure on the torus. Indeed, we adopt here flat structures instead of hyperbolic structures for defining the Thurston compactification.

for $[a, b] \in \mathcal{MF} \cong \mathbb{R}^2/\mathbb{Z}_2$ (cf. (13.6.2)). Let $\{\xi_n\}_n$ be a divergent sequence in $\mathcal{T}_1 \cong \mathbb{D}$. By taking a subsequence if necessary, we may assume that ξ_n converges to $e^{\sqrt{-1}\Theta} = \tau^{-1}(b/a)$ with $\Theta = \Theta([a, b])$. Then,

$$\mathcal{E}_{\xi_n}(F_{[q,p]}) \to \frac{|ap - bq|}{\sqrt{a^2 + b^2}} = \frac{i\left(F_{[a,b]}, F_{[q,p]}\right)}{\mathrm{Ext}_{X_0}(F_{[a,b]})^{1/2}}$$

as $n \to \infty$ for any $p/q \in \hat{\mathbb{Q}} \cong \mathcal{S}$. Thus, the sequence $\{\mathcal{E}_{\xi_n}\}_n$ is thought of as converging to the function $i(\,\cdot\,, F_{[a,b]})/\mathrm{Ext}_{X_0}(F_{[a,b]})^{1/2}$ on the space of functions on $\mathcal{S}$. Since the Gardiner–Masur boundary contains $\mathcal{PMF}$, this boundary is identified with $\mathcal{PMF}(\Sigma_1)$. Furthermore, the identification $\mathcal{T}_1 \cong \mathbb{D}$ extends homeomorphically between the Gardinar-Masur compactification of $\mathcal{T}_1$ and the closed unit disk $\overline{\mathbb{D}}$.

The Gromov product of the Teichmüller distance satisfies

$$\exp(-2\langle \xi \mid \eta \rangle_0) = \left(\frac{1 - |\xi|}{1 + |\xi|} \frac{1 - |\eta|}{1 + |\eta|} \frac{|1 - \overline{\eta}\xi| + |\xi - \eta|}{|1 - \overline{\eta}\xi| - |\xi - \eta|} \right)^{1/2}$$

for $\xi, \eta \in \mathcal{T}_1 \cong \mathbb{D}$. Hence, if ξ and η tend to the projective classes $[F_{[a,b]}]$ and $[F_{[c,d]}]$ of $F_{[a,b]}$, $F_{[c,d]} \in \mathcal{MF} - \{0\} \cong (\mathbb{R}^2 - \{0\})/\mathbb{Z}_2$, we have

$$\exp(-2\langle \xi \mid \eta \rangle_0) \to \frac{|ad - bc|}{\sqrt{a^2 + b^2}\sqrt{c^2 + d^c}} = \frac{i(F_{[a,b]}, F_{[c,d]})}{\mathrm{Ext}_{X_0}(F_{[a,b]})^{1/2}\mathrm{Ext}_{X_0}(F_{[c,d]})^{1/2}}.$$

13.7 Complex Analysis on Teichmüller Space

13.7.1 Complex Analysis

Plurisubharmonic functions are fundamental functions in the theory of several complex variables. An upper-semicontinuous function u on a domain Ω in $\mathbb{C}^N$ which is not identically $-\infty$ is said to be *plurisubharmonic* if for all $a \in \Omega$ and $v \in \mathbb{C}^N$, $\lambda \mapsto u(a + \lambda v)$ is subharmonic or identically $-\infty$ on every component of the set $\{\lambda \in \mathbb{C} \mid a + \lambda v \in \Omega\}$ [40, §2.9]. Let u be a C^2 function on a domain Ω in $\mathbb{C}^N$. Let $z = (z_i)_{i=1}^N \in \Omega$ and $v = (v_i)_{i=1}^N \in \mathbb{C}^N = T_z\Omega$. The *Levi form* is defined by

$$L(u)(z)[v, \overline{v}] = \sum_{i,j=1}^{N} \frac{\partial^2 u}{\partial z_i \partial \overline{z}_j}(z) v_i \overline{v_j}$$

(cf. [40, (1.4.2)]). A C^2-function u on Ω is said to be plurisubharmonic if and only if $L(u)(z)[v, \overline{v}] \geq 0$ for any $z \in \Omega$ and $v \in \mathbb{C}^N$. We can define plurisubharmonic functions on complex manifolds in the usual manner.

A domain Ω in $\mathbb{C}^N$ is said to be *hyperconvex* (in the sense of Stehlé [74]) if there is an upper-bounded continuous plurisubharmonic exhaustion on Ω. Demailly [23] showed that for any bounded hyperconvex domain Ω in $\mathbb{C}^N$ and $w \in \Omega$, there is a unique continuous plurisubharmonic function $u_{\Omega,w} = u_\Omega(\cdot, w)$ on $\overline{\Omega}$ with values in $[-\infty, 0]$ such that

(1) $u_{\Omega,w} \mid_{\partial\Omega} = 0$;
(2) the current $(dd^c u_\Omega)^N$ is the Dirac measure $(2\pi)^N \delta_w$ supported at w; and
(3) $u_{\Omega,w}(z, w) = \log \|z - w\| + O(1)$ as $z \to w$.

Indeed, the function $u_{\Omega,w}$ is characterized as

$$u_{\Omega,w}(z, w) = \sup_v v(z)$$

where the supremum runs over all non-positive plurisubharmonic functions v satisfying $v(z) \leq \log \|z - w\| + O(1)$ as $z \to w$. The function u_Ω is called the *pluricomplex Green function* on Ω (see [39]).

Demailly [23] also found the pluriharmonic measures for bounded hyperconvex domains. For a bounded hyperconvex domain Ω in $\mathbb{C}^N$, the *pluriharmonic measure* is a Radon measure supported on $\partial\Omega$ defined as the limit of the family of the Monge–Ampère measures associated to the pluricomplex Green function (cf. [23, (5.2) Définition]). The pluriharmonic measure $\boldsymbol{\omega}_z$ of $z \in \Omega$ satisfies the Lelong–Jensen formula

$$V(z) = \int_{\partial\Omega} V(\zeta) d\boldsymbol{\omega}_z(\zeta) - \frac{1}{(2\pi)^N} \int_\Omega (dd^c V) \wedge |u_{\Omega,z}| (dd^c u_{\Omega,w})^{N-1} \tag{13.7.1}$$

for any continuous function V on $\overline{\Omega}$ which is plurisubharmonic on Ω (cf. [23, (5.1) Théorème]). In particular, since $dd^c V = 0$ on Ω when V is holomorphic (or pluriharmonic) on a neighborhood of $\overline{\Omega}$, the *Poisson integral formula*

$$V(z) = \int_{\partial\Omega} V(\zeta) d\boldsymbol{\omega}_z(\zeta) \tag{13.7.2}$$

holds from (13.7.1) for a holomorphic (or pluriharmonic) function V on Ω which is continuous on $\overline{\Omega}$.

13.7.2 The Complex Structure on Teichmüller Space Revisited

As discussed in Sect. 13.4.4, the Bers slice is a realization of Teichmüller space $\mathcal{T}_g$ as a bounded domain in a complex Euclidean space. The Bers slice, as a bounded domain, has rich properties in the complex analytical aspect. For instance, Shiga [72] showed that the Bers slice is polynomially convex (see also [24]). Krushkal [42] observed that the Teichmüller space $\mathcal{T}_g$ is hyperconvex (see also [61, 84]).

13.7.3 Complex Analysis with Extremal Length

Extremal length functions have nice complex analytic properties. Indeed, the negatively reciprocal of the extremal length function

$$\mathcal{T}_g \ni x \mapsto -\frac{1}{\mathrm{Ext}_x(F)} \tag{13.7.3}$$

is plurisubharmonic on $\mathcal{T}_g$ for all $F \in \mathcal{MF} - \{0\}$ (cf. [61, Theorem 5.3]). We have a concrete formula of the Levi form of the extremal length function of a generic measured foliation (cf. [61, Theorem 5.1]). The plurisubharmonicity of the functions (13.7.3) implies that the extremal length functions themselves are plurisubharmonic, and the logarithm of any polynomial of extremal length functions with positive coefficients is also plurisubharmonic (cf. [61, Corollary 1.1]). Especially, the Teichmüller distance function

$$\mathcal{T}_g \ni x \mapsto d_T(x, y) = \frac{1}{2} \sup_{\alpha \in \mathcal{S}} (\log \mathrm{Ext}_x(\alpha) - \log \mathrm{Ext}_y(\alpha))$$

is plurisubharmonic for fixed $y \in \mathcal{T}_g$ (cf. [61]. See also [40, Theorem 2.6.1]). The plurisubharmonicity of the extremal length functions was first proved by Liu and Su [47]. The plurisubharmonicity of the Teichmüller distance function was first observed by Krushkal [43]. Our discussions are different from them in any case.

13.7.4 Pluricomplex Green Function

Krushkal [43] observed that the pluricomplex Green function on $\mathcal{T}_g$ satisfies the following remarkable formula, the so-called *Krushkal formula*

$$u_{\mathcal{T}_g}(x, y) = \log \tanh d_T(x, y) \tag{13.7.4}$$

for $x, y \in \mathcal{T}_g$, where $u_{\mathcal{T}_g}(x, y) = u_{\mathcal{T}_g, y}(x, y)$ is the pluricomplex Green function on $\mathcal{T}_g$ with pole at y discussed in the previous section (see also [63] for a different proof). A concrete formula for the Levi form of the Teichmüller distance function and the pluricomplex Green function are given in [63] on an open dense subset of $\mathcal{T}_g$, and the Levi forms are described in terms of the complexified Thurston symplectic form (see [26] for related results). The pluricomplex Green function is expected to be a connector between the topological aspect and the complex analytical aspect in Teichmüller theory (cf. Sect. 13.8.3).

13.7.5 Pluriharmonic Measures

We give a review of the recent progress on Demailly's pluriharmonic measures on Teichmüller spaces.

Fix $x_0 \in \mathcal{T}_g$. The pluriharmonic measures are supported on the Bers boundary $\partial\mathcal{T}^B_{x_0}$. The Kerckhoff–Thuston theorem asserts that the Bers compactification is dependent on the choice of the base point x_0 (cf. [37]). Hence, we denote by $\boldsymbol{\omega}^{x_0}_x$ the pluriharmonic measure of $x \in \mathcal{T}_g$ on $\mathcal{T}^B_{x_0}$.

For $\varphi \in \partial^{ue}\mathcal{T}^B_{x_0}$, let $F_\varphi \in \mathcal{MF}$ be the representative of the measured foliation whose support corresponds to the ending lamination of the marked Kleinian surface group Γ_φ (cf. Sect. 13.4.5.2). We define a function $\mathcal{T}_g \times \mathcal{T}_g \times \partial\mathcal{T}^B_{x_0}$ by

$$\mathbb{P}(x, y, \varphi) = \begin{cases} \left(\dfrac{\mathrm{Ext}_x(F_\varphi)}{\mathrm{Ext}_y(F_\varphi)}\right)^{3g-3} & (\varphi \in \partial^{ue}\mathcal{T}^B_{x_0}) \\ 1 & (\text{otherwise}). \end{cases}$$

Theorem 13.7.1 (Poisson Kernel) *For any $x, y \in \mathcal{T}_g$, we have*

$$d\boldsymbol{\omega}^{x_0}_y = \mathbb{P}(x, y, \cdot\,)d\boldsymbol{\omega}^{x_0}_x \tag{13.7.5}$$

almost everywhere on $\partial\mathcal{T}^B_{x_0}$ with respect to the pluriharmonic measure.

Demailly observed that for any $x, y \in \mathcal{T}_g$, $\boldsymbol{\omega}^{x_0}_x$ and $\boldsymbol{\omega}^{x_0}_y$ are absolutely continuous with respect to each other (cf. [23, (5.3) Théorème]). Hence, the condition "almost everywhere" in the last part of the statement of Theorem 13.7.1 is independent of the choice of the points of the definition for the pluriharmonic measures.

Corollary 13.7.1 (Poisson Integral Formula) *Let f be a holomorphic function or a pluriharmonic function on a domain containing the Bers compactification. Then,*

$$f(x) = \int_{\partial\mathcal{T}^B_{x_0}} f(\varphi)\mathbb{P}(x_0, x, \varphi)d\boldsymbol{\omega}^{x_0}_{x_0}(\varphi)$$

for $x \in \mathcal{T}_g$ after identifying $\mathcal{T}_g \cong \mathcal{T}^B_{x_0}$ via the Bers embedding with the base point x_0.

We give a sketch of the proof of Theorem 13.7.1. The details appear in [62]. We first notice the following which is cofirmed in [62].

Proposition 13.7.1 (APT's are Null) *For any $x \in \mathcal{T}_g$, the pluriharmonic measure $\boldsymbol{\omega}^{x_0}_x$ is supported in $\partial^{min}\mathcal{T}^B_{x_0}$. Namely, we have $\boldsymbol{\omega}^{x_0}_x(\partial^{APT}\mathcal{T}^B_{x_0}) = 0$.*

Recall that $\partial^{min}\mathcal{T}^B_{x_0}$ is a subset of $\partial\mathcal{T}^B_{x_0}$ consisting of boundary groups whose ending lamination is minimal (cf. Sect. 13.4.5.2). Proposition 13.7.1 is proved as follows. For any $\gamma \in \pi_1(\Sigma_g)$, the trace function Tr_γ of γ is a holomorphic function on the ambient space of the Bers slice whose values are in $\mathbb{C} - [0, 4)$ on the Bers compactification (cf. (13.8.1)). Furthermore, $\mathrm{Tr}_\gamma(\varphi) = 4$ if and only if γ is an APT of ρ_φ. Composing an affine conjugation of the Joukowsky transform, we get a continuous function λ on the Bers compactification which is holomorphic in $\mathcal{T}_g$ such that the values are in $\mathbb{D} \cup \{1\}$ and $\lambda(\varphi) = 1$ if and only if γ is an APT of ρ_φ. As $n \to \infty$, the n-th power λ^n tends to the characteristic function on the set consisting of $\varphi \in \partial\mathcal{T}^B_{x_0}$ such that γ is an APT of ρ_φ. Since $\pi_1(\Sigma_g)$ is a countable group, the Poisson integral formula (13.7.2) for λ^n and the Lebesgue dominated convergence theorem guarantee that the pluriharmonic measure of $\partial^{APT}\mathcal{T}^B_{x_0}$ is zero.

Fix $x, y \in \mathcal{T}_g$. We define a measurable function Λ on $\partial\mathcal{T}^B_{x_0}$ by the limit

$$\Lambda(\varphi) = \limsup_{z \to \varphi} \frac{\log\tanh(d_T(y, z))}{\log\tanh(d_T(x, z))}$$

at $\varphi \in \partial\mathcal{T}^B_{x_0}$. Then, Demailly observes that

$$d\boldsymbol{\omega}^{x_0}_y \leq \Lambda^{3g-3} d\boldsymbol{\omega}^{x_0}_x \tag{13.7.6}$$

on $\partial\mathcal{T}^B_{x_0}$ (cf. [23, (3.8) Théorème] and (13.7.4)). When $\varphi \in \partial^{ue}\mathcal{T}^B_{x_0}$, we can see that any sequence in $\mathcal{T}_g$ converging to φ in the Bers compactification converges to the projective class $[F_\varphi]$ in the Gardiner–Masur compactification by applying the argument in [60] and the characterization of uniquely ergodic points in the Gardiner–Masur compactification developped in [57, 59]. Since

$$\frac{\log\tanh(d_T(y, z))}{\log\tanh(d_T(x, z))} = \exp(2(d_T(x, z) - d_T(y, z)))(1 + o(1))$$

as $z \to \varphi$, we obtain

$$\Lambda(\varphi) = \frac{\mathrm{Ext}_x(F_\varphi)}{\mathrm{Ext}_y(F_\varphi)}$$

for $\varphi \in \partial^{ue}\mathcal{T}^B_{x_0}$ from the formulae in [59, §5.1] (see also Sect. 13.8.1).

Let us finish the sketch of the proof of Theorem 13.7.1. Let $\mathcal{PMF}^{min} \subset \mathcal{PMF}$ be the set of projective classes of minimal measured foliations. From Klarreich's result [38], there is a continuous projection $\varpi : \mathcal{PMF}^{min} \to \partial\mathcal{C}_g$ (see also [29]). Notice from Proposition 13.7.1 that $\boldsymbol{\omega}^{x_0}_{x_0}$ is supported on $\partial^{min}\mathcal{T}^B_{x_0}$.

A key for proving Theorem 13.7.1 is that $\boldsymbol{\omega}^{x_0}_{x_0}$ is absolutely continuous with respect to the push-forward measure $(\Phi \circ \varpi)_*(d\nu^{x_0}_{Th})$ of the Thurston measure (Sect. 13.5.2). In fact, we see in [62] that $\boldsymbol{\omega}^{x_0}_{x_0}$ coincides with $(\Phi \circ \varpi)_*(\nu^{x_0}_{Th})$ on $\partial\mathcal{T}^B_{x_0}$ (see Sect. 13.8.1).

From the absolute continuity, we conclude that $\partial^{ue}\mathcal{T}^B_{x_0}$ is of full measure with respect to the pluriharmonic measure $\boldsymbol{\omega}^{x_0}_{x_0}$ on $\mathcal{T}_g$, since from Masur's result [48], $\partial^{ue}\mathcal{T}^B_{x_0}$ is a set of full measure on $\partial\mathcal{T}^B_{x_0}$ with respect to $(\Phi\circ\varpi)_*(\nu^{x_0}_{Th})$. Therefore, the function Λ defined above coincides with $\mathbb{P}(x, y, \cdot)$ almost everywhere on $\partial\mathcal{T}^B_{x_0}$ with respect to the pluriharmonic measures. Since the function $\mathbb{P}$ is reciprocal in the sense that $\mathbb{P}(y, x, \varphi) = \mathbb{P}(x, y, \varphi)^{-1}$ for $(x, y, \varphi) \in \mathcal{T}_g \times \mathcal{T}_g \times \partial\mathcal{T}^B_{x_0}$, from (13.7.6), the identity (13.7.5) holds almost everywhere on $\partial\mathcal{T}^B_{x_0}$ with respect to the pluriharmonic measures.

13.7.6 Toy Model: The Case of Tori

The (pluricomplex) Green function on $\mathcal{T}_1 \cong \mathbb{D}$ with pole at $\xi_0 \in \mathcal{T}_1$ (satisfying the properties given in Sect. 13.7.1) is

$$u_{\mathcal{T}_1}(X_\xi, X_{\xi_0}) = \log\left|\frac{\xi - \xi_0}{1 - \overline{\xi_0}\xi}\right| = \log\tanh d_T(X_\xi, X_{\xi_0}).$$

After identifying the Thurston (Gardiner–Masur) compactification and the Bers compactification via Minsky's ending lamination theorem for once-punctured torus groups, the (pluri)harmonic measure at $\xi \in \mathcal{T}_1 \cong \mathbb{D}$ is equal to the usual harmonic measure

$$d\boldsymbol{\omega}^0_\xi = \frac{1}{2\pi}\frac{1 - |\xi|^2}{|\xi - e^{\sqrt{-1}\Theta}|}d\Theta, \tag{13.7.7}$$

which coincides with the (normalized) Thurston measure (13.5.4). In this case, the function defined by

$$\mathbb{P}(\xi, \eta, [a, b]) = \frac{\mathrm{Ext}_\xi(F_{[a,b]})}{\mathrm{Ext}_\eta(F_{[a,b]})} = \frac{1 - |\eta|^2}{1 - |\xi|^2}\frac{|\xi - e^{\sqrt{-1}\Theta([a,b])}|}{|\eta - e^{\sqrt{-1}\Theta([a,b])}|}$$

for $(\xi, \eta, [a, b]) \in \mathcal{T}_1 \times \mathcal{T}_1 \times \mathcal{PMF}$ satisfies

$$d\boldsymbol{\omega}^0_\eta = \mathbb{P}(\xi, \eta, \cdot)d\boldsymbol{\omega}^0_\xi$$

on $\mathcal{PMF}$ for $\xi, \eta \in \mathcal{T}_1$ as (13.7.5).

13.8 Toward Complex Analysis with Thurston Theory

13.8.1 Pluriharmonic Measures

After [6, §2.3.1], we consider the *cocycle function* defined by

$$\beta(x, y; \varphi) = \begin{cases} \frac{1}{2}(\log \mathrm{Ext}_x(F_\varphi) - \log \mathrm{Ext}_y(F_\varphi)) & (\varphi \in \partial^{ue}\mathcal{T}_{x_0}^B) \\ 0 & (\text{otherwise}). \end{cases}$$

for $(x, y, \varphi) \in \mathcal{T}_g \times \mathcal{T}_g \times \partial\mathcal{T}_{x_0}^B$. The cocycle function β is also understood as the *horofunction* for the Teichmüller distance when $\varphi \in \partial^{ue}\mathcal{T}_{x_0}^B$ (cf. [46]. See also [58, 59]). Theorem 13.7.1 implies that the pluriharmonic measures $\{\boldsymbol{\omega}_x^{x_0}\}_{x\in\mathcal{T}_g}$ on $\partial\mathcal{T}_{x_0}^B$ are regarded as *conformal densities* of dimension $\delta = 6g - 6$ for the cocycle β in the sense of Athreya, Bufetov, Eskin and Mirzakhani: that is, for $x, y \in \mathcal{T}_g$ and a measurable subset $U \subset \partial\mathcal{T}_{x_0}^B$, we have

$$d\boldsymbol{\omega}_y^{x_0} = \exp(\delta\beta(x, y; \cdot))\, d\boldsymbol{\omega}_x^{x_0}$$
$$\boldsymbol{\omega}_{\omega(x)}^{x_0}(\omega(U)) = \boldsymbol{\omega}_x^{x_0}(U)$$

(see the discussion after [23, (5.2) Définition]). Notice from the ending lamination theorem that the mapping class group $\mathrm{MCG}(\Sigma_g)$ acts on $\partial^{min}\mathcal{T}_{x_0}^B$ homeomorphically. From [45], the conformal density of dimension $6g - 6$ on $\mathcal{PMF}$ are unique up to scale, and equal to the system of Thurston measures on $\mathcal{PMF}$.

We actually have an analogy with the case of Teichmüller space of tori concerning the pluriharmonic measures (cf. [62]). Namely, as observations (13.5.4) and (13.7.7) in our toy model, the pluriharmonic measure coincides with the pushforward measure of the Thurston measure via the mapping (13.4.1). The coincidence is a *topological characterization* of the pluriharmonic measures in Teichmüller theory.

Kaimanovich and Masur [34] discussed the Poisson boundary of Teichmüller space and obtained the Poisson integral formula in their setting. Though the author does not know any connection between their setting and our setting at this time, any connection will be interesting.

13.8.2 Trace Functions

The Bers slice is contained in the space of projective structures as mentioned in Sect. 13.4.4. For $\gamma \in \pi_1(\Sigma_g)$, the *trace function*

$$\mathrm{Tr}_\gamma(\varphi) = \mathrm{tr}^2(\rho_\varphi(\gamma)) \quad (\varphi \in A_2(\mathbb{H}^*, \Gamma)) \tag{13.8.1}$$

is a holomorphic function on the ambient space $A_2(\mathbb{H}^*, \Gamma)$. From the Poisson integral formula (Corollary 13.7.1),

$$\mathrm{Tr}_\gamma(x) = \int_{\partial \mathcal{T}_{x_0}^B} \mathrm{Tr}_\gamma(\varphi)\mathbb{P}(x_0, x, \varphi)d\boldsymbol{\omega}_{x_0}^{x_0}(\varphi) \tag{13.8.2}$$

for $\gamma \in \pi_1(\Sigma_g)$ under the identification $\mathcal{T}_g \cong \mathcal{T}_{x_0}^B$. The trace function is the quadruple of the square of cosh of the half of the *complex length* of γ of the representation. As the pivot theorem in [52], the short curve theorem in [53], and the length bound theorem in [19], the behavior of the complex length is studied in terms of end invariants from the combinatorial view point. We expect, with Extremal length geometry, the behavior of the trace function in $\mathcal{T}_g$ is understood with the averaging (13.8.2) of the combinatorial descriptions of the trace functions at the Bers boundary.

13.8.3 Holomorphic Functions

Since Teichmüller space is a Stein manifold (cf. [13]), the complex analytical structure is essentially understood from the $\mathbb{C}$-algebra of the space of holomorphic functions (cf. [33]). The discussion in the previous section will provide a new approach to the study of holomorphic functions on Teichmüller space.

Since the Bers slice $\mathcal{T}_{x_0}^B$ is polynomially convex in $A_2(\mathbb{H}^*, \Gamma)$ (cf. [72]), any holomorphic function on Teichmüller space can be approximated by holomorphic functions on the ambient space $A_2(\mathbb{H}^*, \Gamma)$. To almost all points in the Bers boundary, the ending laminations which are topological invariants are associated. Hence, the boundary values of given holomorphic functions stand for the "combinatorial" descriptions of the holomorphic functions in Teichmüller theory. Thus, any holomorphic function on Teichmüller space will be understood by the behavior of the boundary values (combinatorial descriptions) of the approximations together with the Poisson integral formula (Corollary 13.7.1).

In [64], the author defines real analytic charts of Teichmüller space associated with the extremal length functions and describes the complex structure with respect to these charts. We also describe the CR-equations on the horosphere of the extremal length functions. The boundary values of the holomorphic functions on the horosphere satisfy the CR-equations (cf. [15]). To characterize the combinatorial descriptions of holomorphic functions at the boundary, we hope to develop a kind of "CR-equations" on the boundary by (for example) extending the CR-equations on the horospheres.

Acknowledgments The author would like to thank the organizers of the conference "99e rencontre entre mathématiciens et physiciens théoriciens : Géométrie et physique" held at IRMA (Strasbourg) in June 2017 for their warm hospitality. The author also thanks Professor Athanase Papadopoulos for his fruitful comments.

References

1. A. A'Campo-Neuen, N. A'Campo, L. Ji, A. Papadopoulos. A commentary on Teichmüller's paper *Veränderliche Riemannsche Flächen* (Variable Riemann surfaces). Deutsche Mathematical, vol. 7 (1944), 344–359, in *Handbook of Teichmüller Theory*, vol. IV. IRMA Lectures in Mathematics and Theoretical Physics, vol. 19 (European Mathematical Society, Zürich, 2014), pp. 805–814
2. N. A'Campo, L. Ji, A. Papadopoulos. On Grothendieck's construction of Teichmüller space, in *Handbook of Teichmüller theory*, vol. VI. IRMA Lectures in Mathematics and Theoretical Physics, vol. 27 (European Mathematical Society, Zürich, 2016), pp. 35–69
3. A. A'Campo-Neuen, N. A'Campo, V. Alberge, A. Papadopoulos. A commentary on Teichmüller's paper "Bestimmung der extremalen quasikonformen Abbildungen bei geschlossenen orientierten Riemannschen Flächen", in *Handbook of Teichmüller theory*, vol. V. IRMA Lectures in Mathematics and Theoretical Physics, vol. 26 (European Mathematical Society, Zürich, 2016), pp. 569–580
4. L.V. Ahlfors, The complex analytic structure of the space of closed Riemann surfaces, in *Analytic Functions* (Princeton University, Princton, 1960), pp. 45–66
5. L. Ahlfors, L. Bers, Riemann's mapping theorem for variable metrics. Ann. of Math. **72**(2), 385–404 (1960)
6. J. Athreya, A. Bufetov, A Eskin, M. Mirzakhani, Lattice point asymptotics and volume growth on Teichmüller space. Duke Math. J. **161**(6), 1055–1111 (2012)
7. A. Belkhirat, A. Papadopoulos, M. Troyanov, Thurston's weak metric on the Teichmüller space of the torus. Trans. Am. Math. Soc. **357**(8), 3311–3324 (2005)
8. L. Bers, Spaces of Riemann surfaces, in *Proceedings of the Internationalia Congress Mathematical 1958* (Cambridge University Press, New York, 1960), pp. 349–361
9. L. Bers, Correction to "Spaces of Riemann surfaces as bounded domains". Bull. Am. Math. Soc. **67**, 465–466 (1961)
10. L. Bers, On boundaries of Teichmüller spaces and on Kleinian groups. I. Ann. of Math. **91**(2), 570–600 (1970)
11. L. Bers, Fiber spaces over Teichmüller spaces. Acta. Math. **130**, 89–126 (1973)
12. L. Bers, The action of the modular group on the complex boundary, in *Riemann Surfaces and Related Topics: Proceedings of the 1978 Stony Brook Conference (State University New York, Stony Brook, New York, 1978)*. Annuals of Mathematical Studies, vol. 97 (Princeton University Press, Princeton, 1981), pp. 33–52
13. L. Bers, L. Ehrenpreis, Holomorphic convexity of Teichmüller spaces. Bull. Amer. Math. Soc. **70**, 761–764 (1964)
14. J.S. Birman C. Series, Algebraic linearity for an automorphism of a surface group. J. Pure Appl. Algebra **52**(3), 227–275 (1988)
15. A. Boggess, in *CR Manifolds and the Tangential Cauchy-Riemann Complex*. Studies in Advanced Mathematics (CRC Press, Boca Raton, 1991)
16. F. Bonahon, Bouts des variétés hyperboliques de dimension 3. Ann. of Math. (2) **124**(1), 71–158 (1986)
17. F. Bonahon, Geodesic laminations on surfaces, in *Laminations and Foliations in Dynamics, Geometry and Topology (Stony Brook, NY, 1998)*. Contemporary Mathematical, vol. 269 (American Mathematical Society, Providence, 2001), pp. 1–37
18. J.F. Brock, K.W. Bromberg, On the density of geometrically finite Kleinian groups. Acta Math. **192**(1), 33–93 (2004)
19. J.F. Brock, R.D. Canary, Y.N. Minsky, The classification of Kleinian surface groups, II: The ending lamination conjecture. Ann. Math. (2) **176**(1), 1–149 (2012)
20. K. Bromberg, Projective structures with degenerate holonomy and the Bers density conjecture. Ann. Math. (2) **166**(1), 77–93 (2007)

21. R.D. Canary, D.B.A. Epstein, P.L. Green, Notes on notes of Thurston, in *Fundamentals of Hyperbolic Geometry: Selected Expositions*. London Mathematical Society Lecture Note Series, vol. 328 (Cambridge University, Cambridge, 2006), pp. 1–115. With a new foreword by Canary
22. A.J. Casson, S.A. Bleiler, *Automorphisms of Surfaces After Nielsen and Thurston*. London Mathematical Society Student Texts, vol. 9 (Cambridge University, Cambridge, 1988)
23. J.-P. Demailly, Mesures de Monge-Ampère et mesures pluriharmoniques. Math. Z. **194**(4), 519–564 (1987)
24. B. Deroin, R. Dujardin, Complex projective structures: Lyapunov exponent, degree, and harmonic measure. Duke Math. J. **166**(14), 2643–2695 (2017)
25. A. Douady, A. Fathi, D. Fried, F. Laudenbach, V. Poénaru, M. Shub, *Travaux de Thurston sur les surfaces*. Astérisque, Société Mathématique de France, Paris (1979), vol. 66. Séminaire Orsay, With an English summary
26. D. Dumas, Skinning maps are finite-to-one. Acta Math. **215**(1), 55–126 (2015)
27. C.J. Earle, Teichmüller theory, in *Discrete Groups and Automorphic Functions (Proceedings of the Conferece, Cambridge, 1975)* (Academic Press, London, 1977), pp. 143–162
28. F.P. Gardiner, H. Masur, Extremal length geometry of Teichmüller space. Complex Variables Theory Appl. **16**(2–3), 209–237 (1991)
29. U. Hamenstädt, Train tracks and the Gromov boundary of the complex of curves, in *Spaces of Kleinian groups*. London Mathematical Society Lecture Note Series (Cambridge University, Cambridge, 2006), pp. 187–207
30. J.L. Harer, The virtual cohomological dimension of the mapping class group of an orientable surface. Invent. Math. **84**(1), 157–176 (1986)
31. J. Hubbard, H. Masur, Quadratic differentials and foliations. Acta Math. **142**(3–4), 221–274 (1979)
32. Y. Imayoshi, M. Taniguchi, *An Introduction to Teichmüller Spaces* (Springer, Tokyo, 1992)
33. H. Iss'sa, On the meromorphic function field of a Stein variety. Ann. of Math. (2) **83**, 34–46 (1966)
34. V.A. Kaimanovich, H. Masur, The Poisson boundary of Teichmüller space. J. Funct. Anal. **156**(2), 301–332 (1998)
35. L. Keen, C. Series, Pleating invariants for punctured torus groups. Topology **43**(2), 447–491 (2004)
36. S.P. Kerckhoff, The asymptotic geometry of Teichmüller space. Topology **19**(1), 23–41 (1980)
37. S.P. Kerckhoff, W.P. Thurston, Noncontinuity of the action of the modular group at Bers' boundary of Teichmüller space. Invent. Math. **100**(1), 25–47 (1990)
38. E. Klarerich, The boundary at infinity of the curve complex and the relative Teichmüller space. preprint (1999)
39. M. Klimek, Extremal plurisubharmonic functions and invariant pseudodistances. Bull. Soc. Math. France **113**(2), 231–240 (1985)
40. M. Klimek, Pluripotential Theory, London Mathematical Society Monographs. New Series, vol. 6 (The Clarendon Press/Oxford University Press, New York, 1991). Oxford Science Publications
41. K. Kodaira, *Complex Manifolds and Deformation of Complex Structures*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 283 (Springer, New York, 1986). Translated from the Japanese by Kazuo Akao, With an appendix by Daisuke Fujiwara
42. S.L. Krushkal, Strengthening pseudoconvexity of finite-dimensional Teichmüller spaces. Math. Ann. **290**(4), 681–687 (1991)
43. S.L. Krushkal, The Green function of Teichmüller spaces with applications. Bull. Am. Math. Soc. (N.S.) **27**(1), 143–147 (1992)
44. C.J. Leininger, S. Schleimer, Connectivity of the space of ending laminations. Duke Math. J. **150**(3), 533–575 (2009)
45. E. Lindenstrauss, M. Mirzakhani, Ergodic theory of the space of measured laminations. Int. Math. Res. Not. IMRN **2008**(4), Art. ID rnm126, 49 (2008)

46. L. Liu, W. Su, The horofunction compactification of the Teichmüller metric, in *Handbook of Teichmüller theory*, vol. IV. IRMA Lectures in Mathematics and Theoretical Physics, vol. 19 (European Mathematical Society, Zürich, 2014), pp. 355–374
47. L. Liu, W. Su, Variation of extremal length functions on Teichmüller space. Int. Math. Res. Not. IMRN **2017**(21), 6411–6443 (2017)
48. H. Masur, Interval exchange transformations and measured foliations. Ann. of Math. (2) **115**(1), 169–200 (1982)
49. H.A. Masur, Y.N. Minsky, Geometry of the complex of curves. I. Hyperbolicity. Invent. Math. **138**(1), 103–149 (1999)
50. H.A. Masur, Y.N. Minsky, Geometry of the complex of curves. II. Hierarchical structure. Geom. Funct. Anal. **10**(4), 902–974 (2000)
51. R. Mineyama, H. Miyachi, A characterization of biholomorphic automorphisms of Teichmüller space. Math. Proc. Cambridge Philos. Soc. **154**(1), 71–83 (2013)
52. Y.N. Minsky, The classification of punctured-torus groups. Ann. of Math. (2) **149**(2), 559–626 (1999)
53. Y.N. Minsky, The classification of Kleinian surface groups. I. Models and bounds. Ann. of Math. (2) **171**(1), 1–107 (2010)
54. M. Mirzakhani, Ergodic theory of the earthquake flow. Int. Math. Res. Not. IMRN **2008** (3), Art. ID rnm116, 39
55. H. Miyachi, On Gardiner-Masur boundary of Teichmüller space, in *Complex Analysis and Its Applications*. OCAMI Studies, vol. 2 (Osaka Municipal University Osaka, 2007), pp. 295–300
56. H. Miyachi, Teichmüller rays and the Gardiner-Masur boundary of Teichmüller space. Geom. Dedicata **137**, 113–141 (2008)
57. H. Miyachi, Teichmüller rays and the Gardiner-Masur boundary of Teichmüller space II. Geom. Dedicata **162**, 283–304 (2013)
58. H. Miyachi, Extremal length geometry, in Handbook of Teichmüller theory, vol. IV. IRMA Lectures in Mathematics and Theoretical Physics, vol. 19 (European Mathematical Society, Zürich, 2014), pp. 197–234
59. H. Miyachi, Unification of extremal length geometry on Teichmüller space via intersection number. Math. Z. **278**(3–4), 1065–1095 (2014)
60. H. Miyachi, A rigidity theorem for holomorphic disks in Teichmüller space. Proc. Amer. Math. Soc. **143**(7), 2949–2957 (2015)
61. H. Miyachi, Extremal length functions are log-plurisubharmonic, in *In the Tradition of Ahlfors–Bers, VII*. Contemp. Mathematical, vol. 696 (American Mathematical Socity, Providence, 2017), pp. 225–250
62. H. Miyachi, Pluripotential theory on Teichmüller space II—Poisson integral formula (2018). Submitted, arXiv: https://arxiv.org/abs/1810.04343
63. H. Miyachi, Pluripotential theory on Teichmüller space I: Pluricomplex Green function. Conform. Geom. Dyn. **23**, 221–250 (2019)
64. H. Miyachi, Toward the complex geometry of Teichmüller space with Extremal length geometry—A complex chart associated with extremal length, in *Proceedings of Workshop on Grothendieck-Teichmülller Theories*, Chern Institute of Mathematics, June 24–30, 2016, Tianjin, China, ed. by L. Ji, A. Papadopoulos, W. Su (International Press, Boston, and Higher Education Press, Beijing, 2021)
65. H. Namazi, J. Souto, Non-realizability and ending laminations: proof of the density conjecture. Acta Math. **209**(2), 323–395 (2012)
66. K. Ohshika, Realising end invariants by limits of minimally parabolic, geometrically finite groups. Geom. Topol. **15**(2), 827–890 (2011)
67. K. Ohshika, Reduced Bers boundaries of Teichmüller spaces. Ann. Inst. Fourier (Grenoble) **64**(1), 145–176 (2014)
68. R.C. Penner, J.L. Harer, *Combinatorics of Train Tracks*. Annals of Mathematics Studies, vol. 125 (Princeton University, Princeton, 1992)
69. M. Rees, An alternative approach to the ergodic theory of measured foliations on surfaces. Ergod. Theory Dyn. Syst. **1**(4), 461–488 (1982/1981)

70. H.L. Royden, Automorphisms and isometries of Teichmüller space, in *Advances in the Theory of Riemann Surfaces (Proceedings of the Conference, Stony Brook, New York, 1969)*. Annuals of Mathematical Studies, No. 66 (Princeton University, Princeton, 1971), pp. 369–383
71. *Séminaire Henri Cartan, 13ième année: 1960/61. Familles d'espaces complexes et fondements de la géométrie analytique. Fasc. 1 et 2: Exp. 1–21*. 2ième édition, corrigée. École Normale Supérieure. Secrétariat mathématique, Paris (1962)
72. H. Shiga, On analytic and geometric properties of Teichmüller spaces. J. Math. Kyoto Univ. **24**(3), 441–452 (1984)
73. H. Shiga, Projective structures on Riemann surfaces and Kleinian groups. J. Math. Kyoto Univ. **27**(3), 433–438 (1987)
74. J.-L. Stehlé, Fonctions plurisousharmoniques et convexité holomorphe de certains fibrés analytiques. C. R. Acad. Sci. Paris Sér. A **279**, 235–238 (1974)
75. O. Teichmüller, Extremale quasikonforme Abbildungen und quadratische Differentiale. Abh. Preuss. Akad. Wiss. Math.-Nat. Kl. **1939**(22), 197 (1940)
76. O. Teichmüller, Bestimmung der extremalen quasikonformen Abbildungen bei geschlossenen orientierten Riemannschen Flächen. Abh. Preuss. Akad. Wiss. Math.-Nat. Kl. **1943**(4), 42 (1943)
77. O. Teichmüller, Veränderliche Riemannsche Flächen. Deutsche Math. **7**, 344–359 (1944)
78. O. Teichmüller, Variable Riemann surfaces, in *Handbook of Teichmüller Theory*, vol. IV. IRMA Lectures in Mathematics and Theoretical Physics, vol. 19 (European Mathematical Society, Zürich, 2014), pp. 787–803. Translated from the German by Annette A'Campo-Neuen
79. O. Teichmüller, Determination of extremal quasiconformal mappings of closed oriented Riemann surfaces, in *Handbook of Teichmüller Theory*, vol. V. IRMA Lectures in Mathematics and Theoretical Physics, vol. 26 (European Mathematical Society, Zürich, 2016), pp. 533–567. Translated from the German by Annette A'Campo-Neuen
80. O. Teichmüller, Extremal quasiconformal mappings and quadratic differentials, in *Handbook of Teichmüller Theory*, vol. V. IRMA Lectures in Mathematics and Theoretical Physics, vol. 26 (European Mathematical Society, Zürich, 2016), pp. 321–483. Translated from the German by Guillaume Théret
81. W. Thurston, The Geometry and Topology of Three-Manifolds (1980). Lecture Note at Princeton University. http://library.msri.org/nonmsri/gt3m/
82. W.P. Thurston, Three-dimensional manifolds, Kleinian groups and hyperbolic geometry. Bull. Am. Math. Soc. (N.S.) **6**(3), 357–381 (1982)
83. W.P. Thurston, Hyperbolic Structures on 3-manifolds, II: Surface groups and 3-manifolds which fiber over the circle. ArXiv Mathematics e-prints (1998)
84. S.-K. Yeung, Bounded smooth strictly plurisubharmonic exhaustion functions on Teichmüller spaces. Math. Res. Lett. **10**(2–3), 391–400 (2003)

Chapter 14
Signatures of Monic Polynomials

Norbert A'Campo

Abstract Let $P : \mathbb{C} \to \mathbb{C}$ be a monic polynomial map of degree $d \geq 1$. We call the inverse image of the union of the real and imaginary axes the geometric picture of the polynomial P. The geometric picture of a monic polynomial is a piecewise smooth planar graph. Smooth isotopy classes relative to the $4d$ asymptotic ends at infinity of geometric pictures are called signatures. The set of signatures Σ_d of monic degree-d polynomials is finite. We give a combinatorial characterization of the set of signatures Σ_d and prove that the space of monic polynomials of given signature is contractible. This construction leads to a real semi-algebraic cell-decomposition

$$\mathrm{Pol}_d = \bigcup_{\sigma \in \Sigma_d} \{P \mid \sigma(P) = \sigma\}$$

of the space Pol_d of monic polynomials of degree d. In this cell-decomposition the classical discriminant locus Δ_d appears as a union of cells. The complement of the classical discriminant $B_d := \mathrm{Pol}_d \setminus \Delta_d$ is a union of cells. The face operators of this cell-decomposition of the space B_d are explicitly given. Since B_d is a classifying space for the braid group, we obtain a finite complex that computes the group cohomology of the braid group with integral coefficients.

The picture of the polynomial P is in fact a union of leaves of the pair of orthogonal foliations of the quadratic differential dP^2. Clearly, our inspiration on this work came from William Thurston's work.

Keywords Cell-decomposition · Spaces of polynomials · Semi-algebraic cell decomposition · Discriminant

AMS Classification: 26C10, 54B15, 14P10

N. A'Campo (✉)
Departement Mathematik und Informatik, Fachbereich Mathematik, Universität Basel, Basel, Switzerland
e-mail: Norbert.ACampo@unibas.ch

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_14

14.1 Introduction

Let $P : \mathbb{C} \to \mathbb{C}$ be a polynomial mapping. We assume that P is monic, i.e. with leading coefficient 1. We call a polynomial P *balanced* if its sub-leading coefficient vanishes which says that the sum of its roots $P^{-1}(0)$ weighted by multiplicity equals 0. A unique Tschirnhausen substitution $z = z - t$ will transform a monic polynomial into a monic and balanced one. We call the inverse image by the map P of the union of the real and the imaginary axis the geometric picture π_P of the monic polynomial P.

Geometric pictures[1] of monic polynomials are special graphs in the Gaussian plane $\mathbb{C}$. Their combinatorial restrictions are listed in the following statement.

Theorem 14.1.1 *Let $P(z)$ be a monic polynomial of degree $d > 0$. Its geometric picture π_P is a smooth graph in $\mathbb{C}$ with the following properties:*

1. *The graph has no cycles. The graph is a forest. The non-compact edges are properly embedded in $\mathbb{C}$.*
2. *The complementary regions have a 4-colouring by symbols A, B, C, D. The colouring is proper, meaning that regions having a boundary edge in common have different colors.*
3. *The edges are oriented smooth curves and have a 2-colouring by symbols R, I. They carry the symbol R if the edge separates D and A or B and C coloured regions. They carry the symbol I if the edge separates A and B or C and D coloured regions. The colouring may be not proper. The orientation is right-handed if one crosses the edge from D to A or A to B, and left-handed if one crosses B to C or C to D.*
4. *The picture has $4d$ edges that, near infinity, are asymptotic to the rays $re^{k\pi i/2d}$, $r > 0$, $k = 0, 1, \cdots, 4d - 1$. The colours R, I alternate and the orientations of the R coloured and also the I coloured alternate between outgoing and in-going.*
5. *Near infinity the sectors are coloured in the counterclockwise orientation by the 4-periodic sequence of symbols $A, B, C, D, A, B, \cdots$.*
6. *The graph can have 5 types of vertices: for the first 4 types only A, B or B, C or C, D or D, A regions are incident and only edges of one color are incident, moreover for the fifth type, regions of all 4 colours are incident and the colours appear in the counterclockwise orientation as $A, B, C, D, A, B, \cdots$. So, in particular the graph has no terminal vertices.*
7. *At all points $p \in \pi_P$ the germ of the graph π_P is smoothly diffeomorphic to the germ at $0 \in \mathbb{C}$ of $\{z \in \mathbb{C} \mid \mathrm{Re} z^k = 0\}$ for some $k = 1, 2, \cdots$.*

Proof The real and imaginary axis decompose the complex plane $\mathbb{C}$ in four regions coloured in cyclic counterclockwise order by A, B, C, D according to the signs of the real and imaginary part respectively $++, -+, --, +-$. The real and imaginary

[1] François Bergeron informed us that he also introduced geometric pictures and studied them for the same purpose, see arXiv:0901.4030.

axes are coloured by R, I and are oriented by the gradients of the real and imaginary parts. In fact, this is the colouring and the orientation of the picture π_z of the degree-1 polynomial z. The picture π_P inherites the colouring for its regions and the colouring together with orientation of its edges by pulling back via the map P the colouring and orientation of π_z. Properties 2, 3, $\cdots$ 7 are clear.

For property 1 we first observe that the function $\mathbf{Re}(P) * \mathbf{Im}(P) : \mathbb{C} \to \mathbb{R}$ is harmonic. Indeed, $\mathbf{Re}(P) * \mathbf{Im}(P) = \frac{1}{2}\mathbf{Im}(P^2)$ and the imaginary part of the holomorphic map P^2 is harmonic. A minimal cycle Z in π_P is a simply closed curve and would bound an open bounded region U. Since the function $\mathbf{Re}(P) * \mathbf{Im}(P)$ vanishes along Z, we would have $\mathbf{Re}(P) * \mathbf{Im}(P) = 0$ on U. It follows that the image $P(U)$ of $P = \mathbf{Re}(P) + i\mathbf{Im}(P)$ is contained in the union of the real and imaginary axes in $\mathbb{C}$, contradicting the openness of the non constant holomorphic mapping P [11]. ■

Theorem 14.1.2 *For a given degree d, there exist only finitely many isotopy classes of graphs satisfying the seven properties of Theorem 14.1.1.*

Proof We compactify the graph by adding $4d$ ideal vertices at infinity, one for each ray, see Property 4. Let v be the number of vertices and v' be the number of inner (non-ideal) vertices of the compactified graph. Since the degree of an inner vertex is at least 4, the number of incidence pairs of a vertex, finite or ideal, and edge is at least $4v' + 4d$. So the number e of edges is at least $2v' + 2d$. Since the graph is an non-empty forest, its Euler number is at least 1. Hence

$$1 \leq v - e \leq (v' + 4d) - (2v' + 2d)$$

showing that $v' \leq 2d - 1$ and $v \leq 6d - 1$. The statement follows, since the number of isotopy classes, relative to infinity, of planar forests with $4d$ ideal fixed terminal vertices at infinity and at most $2d - 1$ finite vertices is finite. ■

Definition A *signature* of degree d is a smooth isotopy class of graphs that satisfy the seven properties of Theorem 14.1.1.

Equipped with the Hausdorff topology of proper closed subsets in $\mathbb{C}$, a signature becomes a topological space of planar graphs. Classical theorems of Rheinhold Baer [1], David Epstein [6] and Jean Cerf [3] in planar topology tell us that a signature is a contractible space. See the thesis of Yves Ladegaillerie [7] for the study of spaces of graphs in surfaces.

The following theorems are the main results. Every signature is realized by a monic polynomial.

Theorem 14.1.3 *Let σ be a signature of degree $d > 0$. Then there exists some $P \in \mathrm{Pol}_d$ whose geometric picture belongs to σ.*

The space of monic polynomials with given signature is contractible.

Theorem 14.1.4 *Let σ be a signature of degree $d > 0$. The space $\{P \in Pol_d \mid \sigma(P) = \sigma\}$ is contractible.*

14.2 Bi-regular Polynomials

As intermezzo we first study the most generic monic polynomials. The corresponding cells are the open cells. We call the map P *bi-regular* if $0 \in \mathbb{R}$ is a regular value for both mappings, the real as well as the imaginary part $\mathbf{Re}(P) : \mathbb{C} \to \mathbb{R}$ and $\mathbf{Im}(P) : \mathbb{C} \to \mathbb{R}$. The geometric picture of a bi-regular polynomial P of degree $d > 0$ is the union of the oriented inverse images of $0 \in \mathbb{R}$ for the maps $\mathbf{Re}(P)$ and $\mathbf{Im}(P)$. It has d vertices of valency 4 at the roots of P and $4d$ non compact terminal edges. Here we show as an example the geometric picture of the bi-regular polynomial

$$P = z^{13} - 6z^7 + z^4 - z^3 + 5z^2 + z + 3 + 2i$$

See the picture in Fig. 14.1 that we made with SAGE [14]. The green lines are the inverse image by P of the oriented (from $-\infty$ to $+\infty$) real axis, and are also the inverse image by $\mathbf{Im}(P)$ of $0 \in \mathbb{R}$. The blue ones are the inverse image by P of the oriented (from $-i\infty$ to $+i\infty$) imaginary axis and also the inverse image by $\mathbf{Re}(P)$ of $0 \in \mathbb{R}$. There are 13 transversal intersection points of a green and blue line, which of course are the roots of the polynomial P. At each root a blue and a green line intersect orthogonally since the polynomial map P is conformal and hence its differential at regular points preserves angles. Each blue or green line is a properly embedded copy of the oriented real line in the plane. Near infinity those

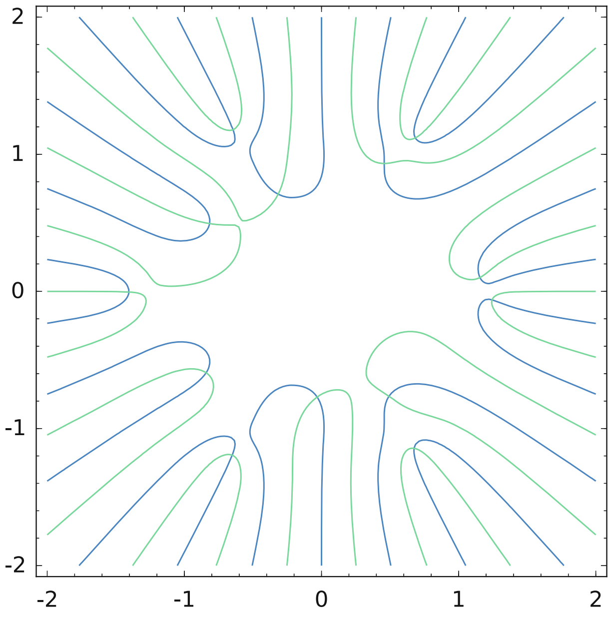


Fig. 14.1 $P = z^{13} - 6z^7 + z^4 - z^3 + 5z^2 + z + 3 + 2i$

lines are asymptotic to rays emanating from the origin. For a bi-regular polynomial of degree d, the inverse image of the real axis is a disjoint union of d copies of an oriented real line, having $2d$ ends that are asymptotic to rays directed by the $2d$-roots of unity θ with $\theta^d = \pm 1$. The inverse image of the imaginary axis is a similar disjoint union of d copies of a real line, except that the ends are asymptotic to the directions of the $4d$-roots of unity θ with $\theta^d = \pm i$. We orient the asymptotic rays from 0 to ∞. The orientation of a curve of the picture and its asymptotic ray match if $\theta^d = +1, +i$ and are opposite if $\theta^d = -1, -i$. We call the geometric picture of a bi-regular polynomial a bi-regular picture.

We say that two bi-regular pictures π, π' are combinatorially equivalent if there exists a regular proper ambient isotopy that keeps the direction of the asymptotics fixed and that moves π to π'. A bi-regular signature is a combinatorial equivalence class of bi-regular pictures.

In the next section we will count the number of bi-regular signatures of degree d. From the combinatorial viewpoint, a bi-regular signature is a signature such that every vertex has valence 4 at which the incident 4 sectors have moreover 4 different colours.

This is especially interesting for following a root r_t continuously given by a family P_t of polynomials. What is still missing, is an understanding of the wall crossings phenomena between different connected components of bi-regular polynomials. In particular we do not know the dual graph of those components for which a component becomes a vertex and a pair of vertices is connected by an edge if one gets from one component to the other by a transversal wall crossing. We plan applications to computer graphics and robotics in the future.

14.3 Counting Bi-regular and Sub Bi-regular Signatures

Let $P : \mathbb{C} \to \mathbb{C}$ be a monic polynomial mapping of degree d. We assume that $0 \in \mathbb{R}$ is a regular value of the imaginary part mapping of $\mathbf{Im}(P) : \mathbb{C} \to \mathbb{R}$. The inverse image $P^{-1}(\mathbb{R}) \subset \mathbb{C}$ is a system of d disjoint smoothly embedded copies of the real line. The orientations of this system can be reconstructed, since the positive end of the real axis is an asymptotic ray with matching orientations and since the matching and non-matching ends alternate if one goes from one $2d$-root of unity to the next. So we can forget the orientations of the components of $P^{-1}(\mathbb{R}) \subset \mathbb{C}$ without loosing information. Combinatorially we can think of $P^{-1}(\mathbb{R}) \subset \mathbb{C}$ as a system of d disjoint diagonals and edges in a $2d$-gon. The number of possible systems $D(d)$ of d non-intersecting diagonals or edges in a $2d$-gon is given by a Catalan number. We put $D(0) = 1$, and have $D(1) = 1$, $D(2) = 2$ and for $d \geq 3$ the number of d non-intersecting diagonals or edges in a $2d$-gon is given by a Catalan number $D(d)$. Here in particular *disjoint* means having no common vertices! Moreover, for $d \geq 3$

the recurrence relation

$$D(d+1) = \sum_{0 \le i \le d} D(i)D(d-i)$$

holds. This recurrence relation is obtained by splitting a $2(d+1)$-gon along the curve that has the first vertex as end. This is the recurrence relation for the Catalan numbers, hence $D(d) = \frac{1}{d+1}\binom{2d}{d}$ [4]. The first Catalan numbers for $d = 1, 2, \cdots$ are

$$1, 2, 5, 14, 42, 132, 429, 1430.$$

For the inverse image of the imaginary axis we also have $D(d)$ possibilities. The two possibilities are very dependent, since each component of the inverse image of the real axis intersects transversally the inverse image of the imaginary axis once. So we need a combined counting. Let $\mathrm{Pict}(d)$ be the number of possible combinatorial types of pictures. We put $\mathrm{Pict}(0) = 1$ and have $\mathrm{Pict}(1) = 1$. For $d \ge 2$ we have the recurrence relation

$$\mathrm{Pict}(d+1) = \sum_{0 \le i, 0 \le j, 0 \le k, 0 \le l, i+j+k+l=d} \mathrm{Pict}(i)\mathrm{Pict}(j)\mathrm{Pict}(k)\mathrm{Pict}(l),$$

which is obtained from the following splitting: let α be the curve in $P^{-1}(\mathbb{R})$ that is asymptotic to the positive real axis and β be the curve in $P^{-1}(i\mathbb{R})$ that intersects A. The pair of curves (α, β) splits the complex plane into four regions. The summing indices i, j, k, l are the number of roots of a bi-regular polynomial in these regions. The Catalan recurrence expresses $D(d+1)$ as a sum of products $D(a)D(b)$ with $a + b = d$. The recurrence for $\mathrm{Pict}(d+1)$ is similar, except that $\mathrm{Pict}(d+1)$ is a sum of 4-factor products. In order to integrate this recursion we first computed with a PARI program [10] the first 15 terms. The result was

$$1, 4, 22, 140, 969, 7084, 53820, 420732, 3362260, 27343888,$$

$$225568798, 1882933364, 15875338990, 134993766600, 1156393243320$$

A search in the on-line Encyclopedia of Integral Sequences founded in 1964 by N.J.A. Sloane, [9] see https:/oeis.org/ identifies this sequence with the sequence $A002293$ and shows to us many interesting interpretations. Also we learn, that the closed formula is of Fuss–Catalan type:

$$\mathrm{Pict}(d) = \frac{1}{3d+1}\binom{4d}{d}$$

By induction upon d we check that the proposed expression satisfies the recursion relation of $\mathrm{Pict}(d)$.

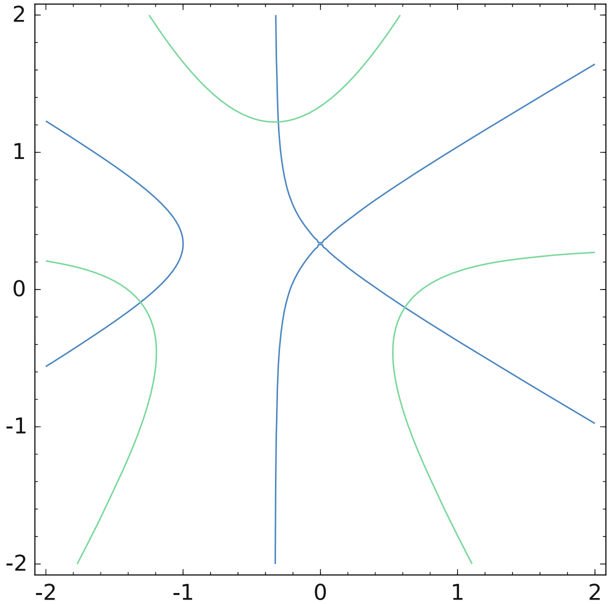


Fig. 14.2 $P(z) = (z - i/3)^3 + (z - i/3)^2 + i$

The space of bi-regular polynomials of degree d is an open subset in the space of all degree-d polynomials.

The connected components correspond bijectively to pictures of bi-regular polynomials. We say that two components are neighbours if they are separated by a wall of real co-dimension 1. In this case we also say that two bi-regular pictures or bi-regular signatures are neighbours. The signature of the picture of Fig. 14.2 defines such a wall that separates two bi-regular components. The picture of Fig. 14.2 allows two smoothings that yield bi-regular pictures.

The discriminant $\Delta \subset \mathbb{C}^d$ is the space of monic polynomial mappings $P : \mathbb{C} \to \mathbb{C}$ having $0 \in \mathbb{C}$ as critical value. Clearly, a polynomial P belongs to Δ if and only if the mappings $\mathbf{Re}(P) : \mathbb{C} \to \mathbb{R}$ and $\mathbf{Im}(P) : \mathbb{C} \to \mathbb{R}$ have a critical point with critical value 0 in common. It follows that each cell of bi-regular polynomials is contained in the complement of Δ. It also follows that the co-dimension 1 walls are contained in the complement of Δ.

The polynomial $P(z) = z^3 - z/3 + 28/27$ is regular above 0 as mapping from $\mathbb{C}$ to $\mathbb{C}$, but the map $\mathbf{Re}(P) : \mathbb{C} \to \mathbb{R}$ has two critical points with 0 as value. The polynomial P belongs to a stratum of real co-dimension 2. See Fig. 14.3. The two critical points of $\mathbf{Re}(P) : \mathbb{C} \to \mathbb{R}$ can fuse together in a stratum of real co-dimension 3. See the picture of $Q(z) = z^3 + 1$ in Fig. 14.4.

The polynomial $P(z) = z^3 - \dfrac{1}{10}z + 1$ is regular above 0 as mapping from $\mathbb{C}$ to $\mathbb{C}$, but the map $\mathbf{Re}(P) : \mathbb{C} \to \mathbb{R}$ has two critical points with 0 as value. The polynomial P belongs to a stratum of real co-dimension 2. See Fig. 14.3. The two critical points

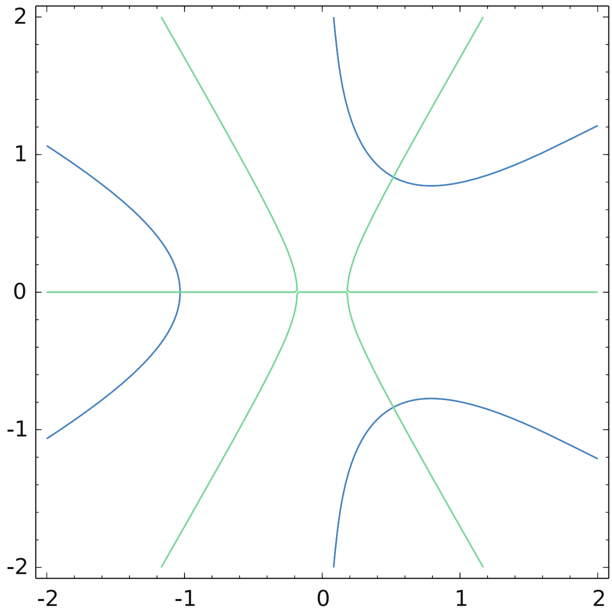


Fig. 14.3 $P(z) = z^3 - \frac{1}{10}z + 1$

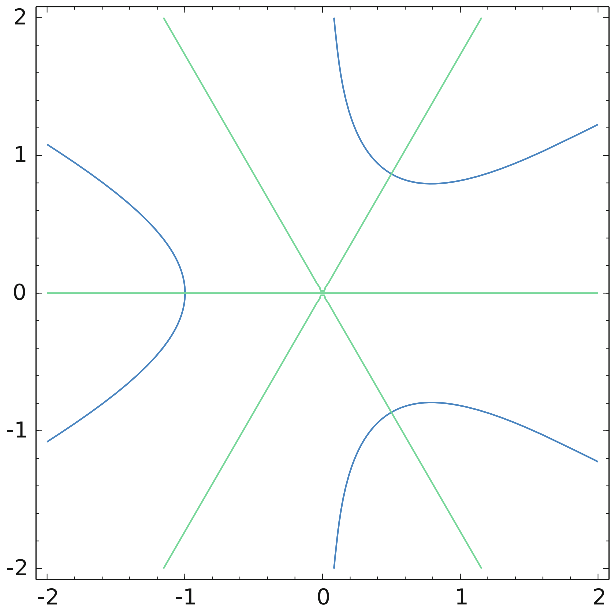


Fig. 14.4 $Q(z) = z^3 + 1$

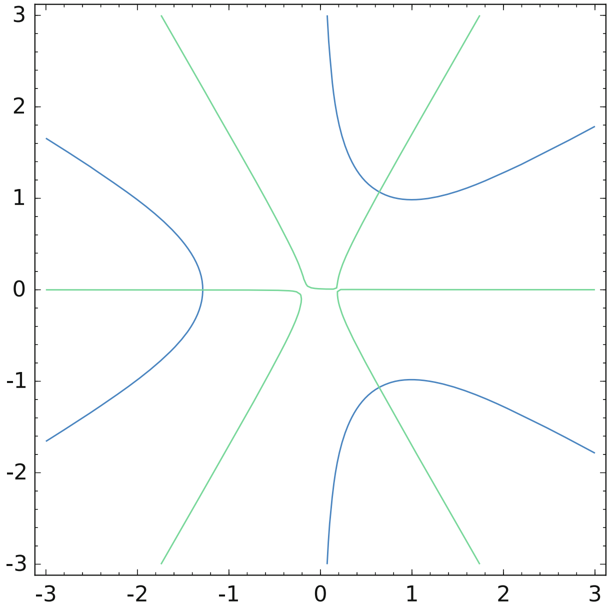


Fig. 14.5 $P(z) = z^3 - (\frac{1}{10} + \frac{i}{200})z + 1 + \frac{i}{1000}$

of $\mathbf{Re}(P) : \mathbb{C} \to \mathbb{R}$ can fuse together in a stratum of real co-dimension 3. See the picture of $Q(z) = z^3 + 1$ in Fig. 14.4.

The two critical points of $\mathbf{Re}(P) : \mathbb{C} \to \mathbb{R}$ can be smoothed, one a lot, the other less, see Fig. 14.5.

Again with a PARI program we could compute the numbers of co-dimension 1 walls in degree $d = 1, 2, 3, \dots$. We get:

$$0, 4, 48, 480, 4560, 42504, 393120, 3624768, 33390720, 307618740$$

This sequence is not identified by the Sloane data base.

After having put a previous version of the present work on the arxiv, Alin Bostan has communicated to me by email the following very interesting observations [2]. The sequence

$$0, 4, 48, 480, 4560, 42504, 393120, 3624768, 33390720, 307618740, \cdots$$

is equal to

$$d \mapsto 4\binom{4d}{d-2}$$

Its generating function is an algebraic hypergeometric function, namely:

$$4x^2 \, {}_3F_2([9/4, 5/2, 11/4], [10/3, 11/3], 2^8/3^3 x)$$

Problems Let $B(d, c)$ be the number of cells in B_d of codimension c. Study the generating series

$$C(x, y) = \sum_{d,c} B(d, c) x^d y^c \in \mathbb{Z}[[x, y]]$$

and the coefficients

$$C_c(x) = \sum_{d} B(d, c) x^d \in \mathbb{Z}[[x]]$$

Study the differential operators that annilate $C(x, y)$, $C_c(x)$. Find closed expressions for $B(d, c)$.

Special polynomials have typical pictures. As example see the fifth Chebyshev polynomial of the first kind in Fig. 14.6. One observes that its picture can be smoothed at 4 places. So, the fifth Chebyshev polynomial belongs to a cell of codimension 4. This cell is in the closure of 2^4 cells of bi-regular polynomials. This holds for all degrees: the Chebyshev polynomial T_n of degree n belongs to a cell of

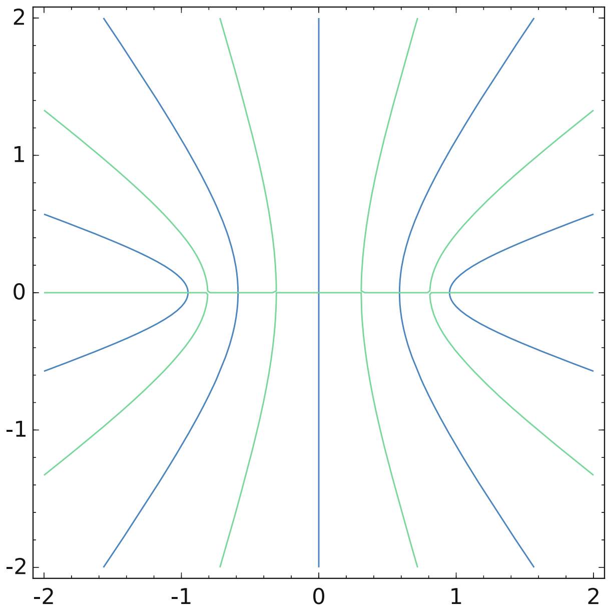


Fig. 14.6 $T(5, z) = 16z^5 - 20z^3 + 5z$

codimension $n - 1$ along which 2^{n-1} bi-regular cells meet. Incidently, observe that 2^{n-1} is the leading coefficient of the polynomial T_n. The cell of the signature $\sigma(T_n)$ is the space of all real monic Morse deformations of the polynomial z^n with $n - 1$ real critical points and 2 critical values.

14.4 Proofs

The proofs are based on the Riemann Mapping Theorem in combination with theorems of Baer [1], Epstein [6] and Cerf [3] on homotopy versus isotopy and theorems of C.J. Earle and J. Eells [5] on the contractability of the connected components of groups of diffeomorphism in dimension two.

Proof of Theorem 14.1.3 Let σ be a signature and let γ be a smooth oriented, coloured embedded graph in the class σ. Let $4d$ be the number of ideal vertices. The seven properties allow to construct a smooth function $f : \mathbb{C} \to \mathbb{C}$ such that the following holds.

1. The graph γ is the inverse image by f of the union of the real and imaginary axes.
2. The map f is open with at most $d - 1$ critical points. The determinant of the tangent map Df is positive at all regular points of f. At each critical point of f the germ of f is smoothly equivalent to the germ of $z \in \mathbb{C} \mapsto z^k + t \in \mathbb{C}$ for some $k = 1, 2, \cdots$ and some $t \in \{+1, -1, +i, -i, 0\}$.
3. The restriction of f to an edge of γ is regular and injective.
4. The colourings of regions and edges of γ are the pull-backs by f of the colourings of P_z.
5. $\lim_{z \in \mathbb{C}, |z| \to +\infty} \dfrac{f(z)}{z^d} = 1$.

The construction of the map f goes in steps. First choose $r_1 > 0$ big enough such that all circles with center $0 \in \mathbb{C}$ and radius $r \geq r_1$ intersect γ transversally. This is possible since the non-bounded edges of γ are asymptotic to rays of constant argument. Define f to be z^d in the complement of the disk of radius r_1. Clearly, f will satisfy 5.

Next label the regions of the complement of γ. Remember that every region goes to infinity. Label the sectors counterclockwise at infinity from 1 to $4d$. The first 4 regions with labels $1, 2, 3, 4$ are those going to infinity along the sectors with label $1, 2, 3, 4$. Let S be the sector with lowest number that does not belong to one of the preceding regions with labels $1, 2, 3, 4$. The region with label 5 will be that region that goes to infinity along the sector S. Now look again for the sector S with lowest label not belonging to an already labelled region. The region going to infinity along S gets label 6, etc.

The degree k_l of a region R_l is defined as its number of sectors at infinity. The number k_l is also the number of connected boundary components of color R or I of R_l.

Extend f, satisfying 1., 2., 3., 4., over the closure of the first region by putting a critical point of type $(z-a) \mapsto z^{k_1} + (i+1)$ at a point in the interior of R_1. Now f is already defined on coloured components of the boundary of the closure of the region R_2. Extend f over the closure of R_2 by putting a critical point of type $(z-a) \mapsto z^{k_2} + (i-1)$, etc.

Let J be the pull-back by f of the standard conformal structure J_0 on $\mathbb{C}$ to $\mathbb{C}$. The map $f : (\mathbb{C}, J) \to (\mathbb{C}, J_0)$ is holomorphic. By the Riemann mapping theorem a biholomorphic map $\rho : (\mathbb{C}, J) \to (\mathbb{C}, J_0)$ exists. Indeed, by property 5 for f, the map extends to a self-map of $\mathbb{C} \cup \{\infty\}$. Replacing ρ finally by a positive real multiple $\lambda\rho$ the composition $f \circ \rho : \mathbb{C} \to \mathbb{C}$ by Rouché's Theorem will be a monic polynomial having a picture in the class of γ. ■

Proof Theorem 14.1.4 Consider the signature σ as a space Γ of smooth oriented planar graphs. The space Γ, if equipped with the topology induced by the topology of the oriented arc-length parametrizations of the edges, is contractible by the theorems of Baer [1] and Epstein [6]. Given $\gamma \in \Gamma$, the space E_γ with the smooth topology of functions $f : \mathbb{C} \to \mathbb{C}$ satisfying the five properties stated in the previous theorem is contractible. The space E_Γ of pairs (f, γ) with $\gamma \in \Gamma$ and $f \in E_\gamma$ by a theorem of Cerf [3] is the total space of a fiber bundle $\pi : E_\Gamma \to \Gamma,\ (f, \gamma) \mapsto \gamma$. It follows that the space E_Γ is contractible.

The group $G_{\mathbb{C},\infty}$ of orientation preserving diffeomorphisms of $\mathbb{C}$ extending to $\mathbb{C} \cup \{\infty\}$ as a diffeomorphism with the identity as differential at ∞, is contractible. The group $G_{\mathbb{C},\infty}$ acts with closed orbits and without fixed points on E_Γ. So the space $E_\Gamma / G_{\mathbb{C},\infty}$ is contractible. By the Riemann mapping Theorem [12], there exists in every $G_{\mathbb{C},\infty}$-orbit a unique pair (f, γ) such that the pull back by f of the standard conformal structure on $\mathbb{C}$ is again the standard structure. In order to achieve uniqueness of the pair (f, γ), we require moreover that f, now by Rouché's Theorem a monic polynomial, is balanced. It follows that the space of monic and balanced polynomials with picture in Γ is a space of representatives for the quotient $E_\Gamma / G_{\mathbb{C},\infty}$.

We conclude that the space of monic balanced polynomial mappings P with picture in the isotopy class Γ is contractible. The group of Tschirnhausen substitutions is contractible and acts fixed point free on the space of monic polynomial mappings P with picture in the isotopy class Γ. Hence, the space of monic polynomial mappings P with signature σ is contractible too. ■

Labelling Roots Let r be a root of a monic polynomial P. The root r belongs to a connected component T of the picture of P. The component T is an coloured oriented tree. We define as label the pair (α, β) consisting of the $4d$-root of unity. The root of unity $\alpha = e^{2\pi i k}/2d$, with $k \in \{0, 1, \cdots, 2d-1\}$ minimal, is in fact the $2d$-root of unity, that we get by starting at r and by following in T the oriented edges in $P^{-1}(\mathbb{R})$. The root of unity $\beta = e^{2\pi i (2k+1)/4d}$ with minimal $k \in \{0, 1, \cdots, 2d-1\}$

is the root of unity that we get by starting at r and by following in T the oriented edges in $P^{-1}(i\mathbb{R})$.

Essentially, from its label we can find back the corresponding root by solving differential equations. The map from root to label is constant in each cell of the cell-decomposition by signatures. This property has clearly applications, each time one wishes to follow roots of polynomials continuously in families of polynomials. Robotics typically encounters this wish.

14.5 Pictures of Meromorphic Functions

The real axis $\mathbb{R} \cup \{\infty\}$ and the imaginary axis $i\mathbb{R} \cup \{\infty\}$, both extended by the point ∞, divide the Riemann sphere $P^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$ in four regions, that again we label by the colours A, B, C, D. Define the picture of a rational map $f : P^1(\mathbb{C}) \to P^1(\mathbb{C})$ to be the inverse image of the union of the extended axis. Similarly, define the picture of a holomorphic map $f : S \to P^1(\mathbb{C})$ on a Riemann surface. Call a rational map or more generally a meromorphic function f on a Riemann surface *very bi-regular* if the critical values do not belong to the extended real or the extended imaginary axis. Call a function f bi-regular if its critical values do not belong to the (non extended) real or imaginary axis. The bi-regular polynomials remain according to this definition bi-regular. We plan to study in future from the combinatorial viewpoint these more general settings.

14.6 Face Operations, Remarks, Questions

The most non-bi-regular polynomial are $P = (z - r)^d, d \geq 2, r \in \mathbb{C}$. Each connected component of bi-regular polynomials of degree $d > 1$ has such a polynomial in its closure. Indeed, let P be bi-regular of degree $d > 1$. The family of bi-regular polynomials $t^d P(z/t),\ t \in \mathbb{R},\ t > 0$, has as limit at $t = 0$ the polynomial z^d. The family is an orbit of the weighted homogeneous action of the group of positive real numbers

$$(t, P) \mapsto t \bullet P = t^{\text{degree}(P)} P(z/t)$$

In fact the larger group of affine substitutions

$$z \mapsto z/t + a,\ t > 0,\ a \in \mathbb{C}$$

acts on monic polynomials. We can use this action for simplification and normalization. We define as norm $\|P\|$ the l^2-norm of the vector of coefficients of P. First, for a polynomial $P \neq (z - r)^d$ of degree $d > 1$ we simplify by a Tschirnhausen substitution $z \mapsto z - a$ killing the coefficient of the term z^{d-1} in P, next we choose $t > 0$ in order to get a polynomial of norm 1. This is possible since for $P \neq z^d$ there exists precisely one $t > 0$ with $\|t \bullet P\| = 1$. So instead of studying the chambers of bi-regular polynomials in the vector space of complex dimension d of all monic polynomials of degree d, one can restrict this study to the unit sphere of real dimension $2d - 3$ in the space of degree d Tschirnhausen simplified polynomials. For instance, it would be interesting to study what happens for $d = 3$. One gets a decomposition of the sphere S^3 in 22 contractible components. What is the dual graph?

Face operators correspond to the following two operations on signatures. Let σ be a signature. Let π be a picture in the class σ. The picture π decomposes the plane $\mathbb{C}$ in polygonal regions. The regions have piecewise smooth curves as boundaries.

The first operation consists in contracting diagonals of regions. Let D be a smooth generic diagonal in such a region connecting two boundary points of the same color, such that the graph $\pi \cup D$ is still a forest. The endpoints of D are smooth points of edges. The new graph $\pi \cup D$ obtained by adding D to the picture π does not satisfy the seven properties. In particular, two vertices are of degree 3. Let π_D be the planar graph obtained by contracting the diagonal D to a new vertex of degree 4. The graph π_D, together with its colouring of edges, labelling of regions, satisfies the seven properties and the class of π_D is again a signature σ_D.

The second operation consists of contracting an edge of π connecting two vertices which are no roots, i.e. vertices not incident with all four colors A, B, C, D. Contracting the edge E in π transforms π to a new picture with the seven properties, so constructs a new signature σ_E.

The operation of adding and contracting a diagonal D to σ or the operation of contracting an edge E such that σ_D or σ_E is again a signature corresponds to a co-dimension one face operation for the cell-decomposition of the space B_d.

It is a challenging problem to understand the combinatorics of these face operations and to describe the corresponding cell and co-chain complex for the spaces B_d.

By a theorem of Stanisław Łojasiewicz [8], the semi-algebraic cell-decomposition of B_d is compatible with a triangulation. Working in the second barycentric subdivisions allows to construct regular open neighborhoods U_σ in B_d of closures in B_d of cells $\{P \mid \sigma(P) = \sigma\} \subset B_d$. The integral Čech-cohomology of the acyclic covering $\{U_\sigma\}$ of B_d computes the group cohomology $H^*((Br(d), \mathbb{Z})$ of the braid group $Br(d)$. It is a challenging problem to do this computation effectively.

A third face operation is needed in spaces of polynomials that have roots of multiplicities exceeding one. Instead of contracting edges, now also contracting minimal subtrees T with two or more edges in π such that the class σ_T of π/T is again a signature gives a face operation.

The unit sphere S^{2d-2} in the space of complex Tschirnhausen simplified polynomials of degree d has a natural probability measure. A natural question is about the probability that a random polynomial P realizes a given picture. What picture has highest probability?

The notion bi-regularity suggests two notions of discriminants for polynomials. We define as real discriminant the set $\Delta_{\mathbb{R}}$ of all monic polynomials P of a given degree such that $0 \in \mathbb{R}$ is a critical value of the real part $\mathbf{Re}(P)$, and accordingly $\Delta_{i\mathbb{R}}$ all polynomials with $0 \in \mathbb{R}$ as critical value for $\mathbf{Im}(P)$. Recall that the classical discriminant Δ is the set of polynomials P such that the mapping $P : \mathbb{C} \to \mathbb{C}$ has $0 \in \mathbb{C}$ as critical value. The complement of the union $\Delta_{\mathbb{R}} \cup \Delta_{i\mathbb{R}}$ is the set of bi-regular polynomials and the classical discriminant Δ is included in the intersection $\Delta_{\mathbb{R}} \cap \Delta_{i\mathbb{R}}$.

We get a braid invariant as follows. A braid b defines an isotopy class of a closed path of monic polynomials in the complement of Δ. What is the minimal number of bi-regular chambers that such a path for a given braid has to visit?

14.7 Sage and Pari Scripts

```
''Hello Pari'' Computes the vector of the number of possible
pictures of the monic degree deg <= g bi-regular polynomials.

vector_numb_pict(g)=
{
X=vector(g,i,0);
X[1]=1;
for(deg=2,g,
        for(a=0,deg-1,
        for(b=0,deg-1-a,
        for(c=0,deg-1-a-b,
        for(d=deg-1-a-b-c,deg-1-a-b-c,
    X[deg]=X[deg]+
                  if(a==0,1,X[a])*
                         if(b==0,1,X[b])*
                              if(c==0,1,X[c])*
                                     id(d==0,1,X[d]);
               ); ); ); );
    );
  X
}
numb_pict(deg)=
{
X=vector_numb_pict(deg);
X[deg]
}

Computes the number of codimension 1 walls in the
space of monic degree deg polynomials.

numb_wall(deg)=
{
```

```
if(deg<3,Res=(deg-1)*4,
       X=vector_numb_pict(deg-1);
       Res=2*deg*sum(a=1,deg-1,X[a]*X[deg-a]);
      );
Res
}

``Hello Sage'' Draw the picture of a polynomial.
Here as example P=z^13-6z^7 .....

import matplotlib
p=Graphics()
var('x','y',domain=RR)
z=x+i*y
f=expand(z^13-6*z^7+z^4-z^3+5*z^2+z+3+2*i)
u=(f+conjugate(f))/2
v=-i*(f-u)
p1=implicit_plot(u==0,(x,-4,4),(y,-4,4),color=rainbow(5)[2])
p2=implicit_plot(v==0,(x,-4,4),(y,-4,4),color=rainbow(5)[3])
p=sum([p1,p2])
p.show()
```

Acknowledgments This work started in 2014 at the Graduate School of Sciences, Hiroshima University during the Conference "Branched Coverings, Degenerations and Related Topics". The author thanks Professors Makoto Sakuma and Ichiro Shimada for the warm hospitality and for providing stimulating mathematical environment. The author thanks the anonymous referee for his suggestions that have improved the exposition.

References

1. R. Baer, Isotopien von Kurven auf orientierbaren, geschlossenen Flächen. J. Reine Angew. Math. **159**, 101–116 (1928)
2. A. Bostan, Communication Email, 21 Feb 2017
3. J. Cerf, La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie. Inst. Hautes Études Sci. Publ. Math. **39**, 5–173 (1970)
4. L. Comtet, *Analyse Combinatoire I, II, Collection SUP* (Presses Universitaires de France, Paris, 1970)
5. C.J. Earle, J. Eells, The diffeomorphism group of a compact Riemann surface. Bull. Am. Math. Soc. **73**(4), 557–559 (1967)
6. D.B.A. Epstein, Curves on 2-manifolds and isotopies. Acta Math. **115**, 83–107 (1966)
7. Y. Ladegaillerie, Découpes et isotopies de surfaces topologiques, Thèse de Doctorat d'État, Faculté des Sciences, Montpellier (1976)
8. S. Łojasiewicz, Triangulation of semi-analytic sets. Annali della Scuola Normale Superiore di Pisa - Classe di Scienze **18**(4), 449–474 (1964)
9. OEIS Foundation Inc., The On-Line Encyclopedia of Integer Sequences (2011), http://oeis.org
10. PARI/GP, version `2.5.0`, Bordeaux (2013), http://pari.math.u-bordeaux.fr/
11. R. Remmert, *Classical Topics in Complex Function Theory* (Springer, New York, 1998)

12. B. Riemann, Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse, Göttingen (1851)
13. N.J.A. Sloane, *A Handbook of Integer Sequences* (Academic, New York, 1973)
14. W.A. Stein et al., SAGE Mathematics Software (Version 6.1.1). The Sage Development Team, http://www.sagemath.org

Chapter 15
Anti-de Sitter Geometry and Teichmüller Theory

Francesco Bonsante and Andrea Seppi

Abstract The aim of this chapter is to provide an introduction to Anti-de Sitter geometry, with special emphasis on dimension three and on the relations with Teichmüller theory, whose study has been initiated by the seminal paper of Geoffrey Mess in 1990. In the first part we give a broad introduction to Anti-de Sitter geometry in any dimension. The main results of Mess, including the classification of maximal globally hyperbolic Cauchy compact manifolds and the construction of the Gauss map, are treated in the second part. Finally, the third part contains related results which have been developed after the work of Mess, with the aim of giving an overview on the state-of-the-art.

Keywords Anti-de Sitter geometry · Globally hyperbolic Lorentzian manifolds · Hyperbolic geometry · Teichmüller theory

2010 Mathematics Subject Classification 53C50, 57M50, 30F60

Introduction

At the end of last century the interest around Lorentzian geometry in low dimension, and in particular Lorentzian manifolds of constant sectional curvature, grew significatively. Among them, the most interesting ones are those of constant *negative*

The authors Francesco Bonsante and Andrea Seppi are members of the national research group GNSAGA.

F. Bonsante (✉)
Dipartimento di Matematica "Felice Casorati", Università degli Studi di Pavia, Pavia, Italy
e-mail: bonfra07@unipv.it

A. Seppi
CNRS and Université Grenoble Alpes, Gières, France
e-mail: andrea.seppi@univ-grenoble-alpes.fr

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_15

sectional curvature, which are called *Anti-de Sitter* manifolds and have been largely studied until nowadays.

There were at least two different motivations behind this increased interest for Lorentzian geometry of constant sectional curvature. The first motivation was the study of proper affine actions on $\mathbb{R}^n$. Affine actions which preserve the Euclidean structure of $\mathbb{R}^n$ are well-understood since the work of Bieberbach of 1912. On the other hand the general case seems considerably more difficult and there are still important open questions in the area. It was natural to consider proper actions which preserve the Minkowski structure as an intermediate problem, which already contains some deep cases, like proper actions of free groups. In particular in dimension three, the classification of free group actions was shown to be crucial towards a complete understanding of three-dimensional affine manifolds, see [65, Theorem 2.1]. This problem has been studied by several authors, see for instance [37, 38, 57, 58, 73], and a complete classification has been given only recently [41, 49, 51]. Similar problems have been studied in the more general setting of proper isometric actions on Lorentzian manifolds of constant sectional curvature [49, 80, 95, 96]. See [54] for a recent and complete survey on these topics.

In a different direction, another motivation arose from the study of gravity in dimension three. In mathematical physics, this consists in the study of Lorentzian metrics on manifolds which obey to the so-called Einstein equation. In dimension three the problem is considerably simpler, since solutions of Einstein equations are precisely Lorentzian metrics of constant sectional curvature (whose sign depends on the choice of the cosmological constant which appears in the Einstein equation). The study of the space of constant sectional curvature metrics was therefore considered as the first step towards a quantization of the three-dimensional gravity, and as a toy model which could help in the understanding of the four-dimensional situation. See for instance the inspiring work of Witten [113]. Unlike its Riemannian counterpart, this classification is expected to include Lorentzian metrics which are not geodesically complete, in light of the relevant notion of initial and final singularity. A standard assumption is to consider *globally hyperbolic metrics*. Roughly speaking, these are metrics which admit foliations by Riemannian hypersurfaces, recovering the idea of a space evolving in time. By a result of Choquet-Bruhat, any globally hyperbolic metric solving Einstein's equation can be isometrically embedded in a maximal one, see [42, 43], which reduces the problem to the classification of maximal globally hyperbolic Einstein spacetimes. In dimension three this problem was addressed by Andersson, Moncrief and others by means of analytic methods (see for instance [3, 4, 91]).

The seminal work of Geoffrey Mess [89], which originally appeared in 1990, represented a very striking, and successful, attempt to link these two different areas. On the one hand Mess proved one of the main achievements in the classification of proper isometric actions of discrete groups on Minkowski space, showing that the action is necessarily by a free group. On the other hand he gave a noteworthy classification of the moduli space of maximal globally hyperbolic spacetimes of constant sectional curvature. Mess' approach, unlike that of Andersson and Moncrief, was based on geometric constructions inspired by the work of Thurston in the 1980s. Indeed a remarkable aspect of his work is the link between three-dimensional gravity and hyperbolic geometry in dimension two, with particular regard to connections with Teichmüller theory. While those connections were expected, and partially contained in the previous work of other authors, it is really the paper of Mess that deeply clarified the picture.

The work of Mess, now considered "classical", provided a new perspective for the study of Lorentzian geometric structures and Teichmüller space. It inspired many lines of investigation which have been developed until the very recent years and seem to be still very promising.

Scope and Organization

The purpose of this chapter is threefold. The first goal is providing an introduction to Anti-de Sitter geometry, first in any dimension and then specifically in dimension three, and this is the content of Part 1. More concretely, in Chap. 15.1.1 we provide some general preliminaries on Lorentzian geometry, with focus on Lorentzian manifolds of constant sectional curvature and maximal isometry group. This serves also as a motivation for the models of Anti-de Sitter space to be introduced later, by explaining in what sense they represent the model spaces for constant negative curvature in the Lorentzian setting. In Chap. 15.1.1.4 we introduce various models of Anti-de Sitter space in arbitrary dimension, and study their geometry and their properties. Chapter 15.1.2 focuses on three-dimensional Anti-de Sitter geometry, by introducing the $\mathrm{PSL}(2, \mathbb{R})$-model which is peculiar to this dimension.

The second goal, achieved in Part 2, is to provide a self-contained exposition of the results of Mess, published in [89], which concern Anti-de Sitter three-dimensional geometry. These can be divided into two main directions: the classification of maximal globally hyperbolic Anti-de Sitter three-manifolds containing a closed Cauchy surface and the construction of the Gauss map. Chapter 15.2.1 contains a number of preliminary results necessary to develop the theory, in particular about causal properties of Anti-de Sitter geometry and isometric actions, which constitute the fundamental setup for the proofs of Mess' classification results. In Chap. 15.2.2 we then prove the classification result of maximal globally hyperbolic manifolds containing a Cauchy surface of genus g. For genus $g = 1$, we describe the deformation space of these structures, which is essentially identified with the deformation space of semi-translation structures on the 2-torus. The situation is extremely more interesting in genus $g \geq 2$. Here the main classification result of Mess, whose proof is concluded in Theorem 15.2.2.5.4, is that the deformation space of maximal globally hyperbolic manifolds is homeomorphic to the product of two copies of the Teichmüller space of the closed surface of genus g. In Chap. 15.2.3 we discuss the construction of the Gauss map associated with spacelike surfaces in Anti-de Sitter space, an idea whose main application in the work of Mess is a proof of Thurston's Earthquake Theorem, using pleated surfaces. We will sketch Mess' proof of the Earthquake Theorem, again in an essentially self-contained fashion, and at the same time we develop further tools, for instance a differential geometric approach to the Gauss map for smooth spacelike surfaces, which have been proved useful in many applications.

Indeed, in Part 3 we survey more recent results on Anti-de Sitter three-dimensional geometry, with special interest in the relations with Teichmüller theory, which somehow rely on the legacy of Mess' paper. In Chap. 15.3.1 we still focus on maximal globally hyperbolic manifolds with closed Cauchy surfaces. We give further results on their structure, for instance on foliations by surfaces with special properties of curvatures, and on the understanding of invariants such as the volume, in relation with their deformation space. As an outcome, we obtain applications in Teichmüller theory, and new parameterizations of the deformation space in terms of holomorphic objects. Finally, in Chap. 15.3.2 we discuss the case of spacelike surfaces with a different topology. We explain a number of results which can be seen as the "universal" version of the analogue problems in the closed case, and derive applications for the theory of universal Teichmüller theory. As a conclusion we mention very briefly Anti-de Sitter structures with timelike cone singularities ("particles") and with multi-black holes, and how they are related to the Teichmüller theory of hyperbolic surfaces with cone points and with geodesic boundary respectively.

Other Research Directions

As mentioned already at the beginning, the aim of this paper is not to provide a comprehensive treatment of Anti-de Sitter geometry, and we decided to focus on three-dimensional geometry, in the spirit of Mess, and to the relations with Teichmüller theory of hyperbolic surfaces. A variety of related topics are not included here, as a result of our necessity to make certain choices in the exposition, but would certainly deserve their own place. Among others, we would like to mention:

- The study of properly discontinuous actions on Anti-de Sitter three-space, a natural problem to consider in light of the results we mentioned in this introduction about proper actions on affine space, for which much work towards a complete classification has been developed in recent times. See [48, 50, 54, 55, 72, 108, 109].
- Higher-dimensional Anti-de Sitter geometry, in particular the study of globally hyperbolic manifolds: [8–11, 85, 90]
- The relations of Anti-de Sitter geometry with other geometric structures, both in dimension three and in higher dimensions, for instance given by the Wick rotation [20, 21, 98] and by geometric transition [45–47, 50, 62, 93, 99].
- The study of dynamical properties of isometric actions on Anti-de Sitter space, for instance in terms of Anosov representations, and the generalizations of these properties to other types of geometric structures. See for instance [8, 10, 52, 69,

70, 74, 76, 84, 112]. It is also worth remarking here the recents works which highlighted the Higgs bundles perspective, see [1, 2, 44].

15.1 Part 1: Anti-de Sitter Space

15.1.1 Preliminaries on Lorentzian Geometry

The aim of this preliminary section is to briefly recall some basic facts about Lorentzian manifolds. We will introduce Lorentzian manifolds of constant sectional curvature and we will see that, as in the Riemannian case, two Lorentzian manifolds of constant sectional curvature K are locally isometric. In particular, we focus on those with maximal isometry group, as they provide models of manifolds of constant sectional curvature: if M is a Lorentzian manifold with constant sectional curvature K and maximal isometry group, then any Lorentzian manifold with constant sectional curvature K carries a natural $(\mathrm{Isom}(M), M)$-atlas made of local isometries. Simply connected space forms have maximal isometry group, but in general there are manifolds with maximal isometry group which are not simply connected. In particular we will focus on the case $K = -1$ and in that case it will be convenient to use models which are not simply connected.

15.1.1.1 Basic Definitions

A *Lorentzian metric* on a manifold of dimension $n+1$ is a non-degenerate symmetric 2-tensor g of signature $(n, 1)$. A *Lorentzian manifold* is a connected manifold M equipped with a Lorentzian metric g.

In a Lorentzian manifold M we say that a non-zero vector $v \in TM$ is *spacelike*, *lightlike*, *timelike* if $g(v, v)$ is respectively positive, zero or negative. More generally, we say that a linear subspace $V \subset T_xM$ is *spacelike, lightlike, timelike* if the restriction of g_x to V is positive definite, degenerate or indefinite.

The set of lightlike vectors, together with the null vector, disconnects T_xM into 3 regions: two convex open cones formed by timelike vectors, one opposite to the other, and the region of spacelike vectors. As a consequence the set of timelike vectors in the total space TM is either connected or is made by two connected components. In the latter case M is said *time-orientable*, and a time orientation is the choice of one of these components. Vectors in the chosen component are said *future-directed*, vectors in the other component are said *past-directed*.

A differentiable curve is *spacelike, lightlike, timelike* if its tangent vector is spacelike (resp. lightlike, timelike) at every point. It is *causal* if the tangent vector is either timelike or lightlike. Given a point x in a time-oriented Lorentzian manifold M, the *future* of x is the set $\mathrm{I}^+(x)$ of points which are connected to x by a future-directed causal curve. The *past* of x, denoted $\mathrm{I}^-(x)$, is defined similarly, for past-directed causal curves.

An *orthonormal basis* of $T_x M$ is a basis $v_1, \dots v_{n+1}$ such that $|g(v_i, v_j)| = \delta_{ij}$, with $v_1, \dots v_n$ spacelike, and v_{n+1} timelike. If $v_1, \dots v_{n+1}$ is an orthonormal basis then for $v = \sum_{i=1}^{n+1} x_i v_i$ we have $g(v, v) = \sum_{i=1}^{n} x_i^2 - x_{n+1}^2$.

As in the Riemannian setting, on a Lorentzian manifold M there is a unique linear connection ∇ which is symmetric and compatible with the Lorentzian metric g. We refer to it as the *Levi-Civita connection* of M. The Levi-Civita connection determines the Riemann curvature tensor defined by

$$R(u, v)w = \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w \ .$$

We then say that a Lorentzian manifold M has constant sectional curvature K if

$$g(R(u, v)v, u) = K \left(g(u, u)g(v, v) - g(u, v)^2 \right) \tag{15.1.1}$$

for every pair of vectors $u, v \in T_x M$ and every $x \in M$. This definition is strictly analogous to the definition given in the Riemannian realm. However in this setting the sectional curvature can be defined only for planes in $T_x M$ where g is non-degenerate.

Finally, we say that M is *geodesically complete* if every geodesic is defined for all times, or in other words, the exponential map is defined everywhere.

15.1.1.2 Maximal Isometry Groups and Geodesic Completeness

Two Lorentzian manifolds M and N of constant curvature K are locally isometric, a fact which is well-known in the Riemannian setting. More precisely, the following holds:

Lemma 15.1.1.2.1 *Let M and N be Lorentzian manifolds of constant curvature K. Then every linear isometry $L : T_x M \to T_y N$ extends to an isometry $f : U \to V$, where U and V are neighbourhoods of x and y respectively. Any two extensions $f : U \to V$ and $f : U' \to V'$ of L coincide on $U \cap U'$. Moreover L extends to a local isometry $f : M \to N$ provided that M is simply connected and N is geodesically complete.*

Exactly as in the Riemannian case the proof is a simple consequence of the classical Cartan–Ambrose–Hicks Theorem (see for instance [92] for a reference). A direct consequence of Lemma 15.1.1.2.1 is the following:

Corollary 15.1.1.2.2 *Let M and N be simply connected, geodesically complete Lorentzian manifolds of constant curvature K. Then any linear isometry $L : T_x M \to T_y N$ extends to a global isometry $f : M \to N$.*

In particular, there is a unique simply connected geodesically complete Lorentzian manifold of constant curvature K up to isometries. For instance for $K = 0$ a model is the Minkowski space $\mathbb{R}^{n,1}$, that is $\mathbb{R}^{n+1}$ provided with the standard metric

$$g = dx_1^2 + \ldots + dx_n^2 - dx_{n+1}^2 \,.$$

In Sect. 15.1.1.7 we will construct an explicit model for $K = -1$.

Another consequence of Lemma 15.1.1.2.1 is that, fixing a point $x_0 \in M$, the set of isometries of M, which we will denote by $\mathrm{Isom}(M)$, can be realized as a subset of $\mathrm{ISO}(T_{x_0}M, TM)$, namely the fiber bundle over M whose fiber over $x \in M$ is the space of linear isometries of $T_{x_0}M$ into T_xM. It can be proved that $\mathrm{Isom}(M)$ has the structure of a Lie group with respect to composition so that the inclusion $\mathrm{Isom}(M) \hookrightarrow \mathrm{ISO}(T_{x_0}M, TM)$ is a differentiable proper embedding, see [78, Theorem 4.1]. It follows that the maximal dimension of $\mathrm{Isom}(M)$ is $\dim \mathrm{O}(n, 1) + n + 1 = (n + 1)(n + 2)/2$.

Definition 15.1.1.2.3 A Lorentzian manifold M has *maximal isometry group* if the action of $\mathrm{Isom}(M)$ is transitive and, for every point $x \in M$, every linear isometry $L : T_xM \to T_xM$ extends to an isometry of M.

Equivalently M has maximal isometry group if the above inclusion of $\mathrm{Isom}(M)$ into $\mathrm{ISO}(T_{x_0}M, TM)$ is a bijection. Hence if M has maximal isometry group, then the dimension of the isometry group is maximal.

From Corollary 15.1.1.2.2, every simply connected Lorentzian manifold M has maximal isometry group if it has constant sectional curvature and is geodesically complete. The converse holds even without the simply connectedness assumption. Namely:

Lemma 15.1.1.2.4 *If M is a Lorentzian manifold with maximal isometry group then M has constant sectional curvature and is geodesically complete.*

Proof Let us show that the sectional curvature is constant. First fix a point $x \in M$. As the identity component of $\mathrm{O}(T_xM) \cong \mathrm{O}(n, 1)$ acts transitively on spacelike planes, there exists a constant K such that Eq. (15.1.1) holds for for every pair (u, v) of vectors tangent at x which generate a spacelike plane. Now, for every point $x \in M$ both sides of Eq. (15.1.1) are polynomial in $u, v \in T_xM$. Since the set of pairs (u, v) which generate spacelike planes is open in $T_xM \times T_xM$, we conclude that Eq. (15.1.1) must hold for every pair $(u, v) \in T_xM \times T_xM$. Since $\mathrm{Isom}(M)$ acts transitively on M, Eq. (15.1.1) holds for every $(u, v) \in T_xM \times T_xM$ independently of x, with the same constant K.

To prove geodesic completeness, we have to show that every geodesic is defined for all times. Suppose γ is a parameterized geodesic with $\gamma(0) = x$ and $\gamma'(0) = v \in T_xM$, which is defined for a finite maximal time $T > 0$. Let $T_0 = T - \epsilon > 0$. Then by assumption one can find an isometry $f : M \to M$ which maps x to $\gamma(T_0)$ and v to $\gamma'(T_0)$. Then $t \mapsto f \circ \gamma(t - T_0)$ is a parameterized geodesic which provides a continuation of γ, thus contradicting the assumption that $T < +\infty$ is the maximal time of definition. □

15.1.1.3 A Classification Result

Simply connected Lorentzian manifolds with maximal isometry group play an important role, in light of the following result of classification.

Proposition 15.1.1.3.1 *Let M_K be a simply connected Lorentzian manifold of constant sectional curvature K with maximal isometry group, and let M be a Lorentzian manifold of constant sectional curvature K. Then:*

- *M is geodesically complete if and only if there is a local isometry $p : M_K \to M$ which is a universal covering.*
- *M has maximal isometry group if and only if $\mathrm{Aut}(p : M_K \to M)$ is normal in $\mathrm{Isom}(M_K)$.*

Proof If M is geodesically complete, then lifting the metric to the universal cover $\widetilde{M}$ one gets a simply connected geodesically complete Lorentzian manifold of constant sectional curvature K, which by Corollary 15.1.1.2.2 is isometric to M_K. The covering map $p : M_K \to M$ is then a local isometry by construction. The converse is straightforward.

Now, let $\Gamma = \mathrm{Aut}(p : M_K \to M)$, which is a discrete subgroup of $\mathrm{Isom}(M_K)$. Thus M is obtained as the quotient $M = M_K/\Gamma$, where Γ acts freely and properly discontinuously on M_K. The isometry group of M is isomorphic to $N(\Gamma)/\Gamma$, where $N(\Gamma)$ is the normalizer of Γ in $\mathrm{Isom}(M_K)$. The isomorphism is based on the observation that any isometry of $\widetilde{M}$ which normalises Γ descends to an isometry of M, and conversely the lifting of any isometry of M must be in $N(\Gamma)$.

Hence the condition that M has maximal isometry group is equivalent to the condition that every element f of $\mathrm{Isom}(M_K)$ descends in the quotient to an isometry of M. This is in turn equivalent to the condition that $f\Gamma f^{-1} = \Gamma$ for every $f \in \mathrm{Isom}(M_K)$, namely, that Γ is normal in $\mathrm{Isom}(M_K)$. □

Remark 15.1.1.3.2 Since $\Gamma = \mathrm{Aut}(p : M_K \to M)$ is discrete, being normal in $\mathrm{Isom}(M_K)$ implies that elements of Γ commute with the elements of the identity component of $\mathrm{Isom}(M_K)$. This remark suggests that there are usually not many Lorentzian manifolds of constant sectional curvature with maximal isometry group.

Finally, any isometry between connected open subsets of a Lorentzian manifold M with maximal isometry group extends to a global isometry. In particular if M_K is a Lorentzian manifold of constant sectional curvature K with maximal isometry group, than any Lorentzian manifold M of constant sectional curvature K admits a natural $(\mathrm{Isom}(M_K), M_K)$-structure whose charts are isometries between open subsets of M and open subsets of M_K.

We will sometimes refer to Lorentzian manifolds of constant sectional curvature K with maximal isometry group as *models* of constant sectional curvature K. After these preliminary motivations, in the following we will study several models of constant sectional curvature -1, or in other words, models of *Anti-de Sitter geometry*.

15.1.1.4 Models of Anti de Sitter $(n + 1)$-Space

We construct here models of Lorentzian manifolds with constant sectional curvature -1 and maximal isometry group in any dimension, by stressing the analogies with the models of hyperbolic space.

15.1.1.5 The Quadric Model

Let us start by the so-called quadric model, which is the analogue of the hyperboloid model of hyperbolic space. Denote by $\mathbb{R}^{n,2}$ the real vector space $\mathbb{R}^{n+2}$ equipped with the quadratic form

$$q_{n,2}(x) = x_1^2 + \ldots + x_n^2 - x_{n+1}^2 - x_{n+2}^2 \,,$$

and by $\langle v, w\rangle_{n,2}$ the associated symmetric form. Finally let $\mathrm{O}(n, 2)$ be the group of linear transformations of $\mathbb{R}^{n+2}$ which preserve $q_{n,2}$.

Then we define

$$\mathbb{H}^{n,1} := \{x \in \mathbb{R}^{n,2} \mid q_{n,2}(x) = -1\} \,.$$

It is immediate to check that $\mathbb{H}^{n,1}$ is a smooth connected submanifold of $\mathbb{R}^{n,2}$ of dimension $n+1$. The tangent space $T_x\mathbb{H}^{n,1}$ regarded as a subspace of $\mathbb{R}^{n+2}$ coincides with the orthogonal space $x^\perp = \{y \in \mathbb{R}^{n+2} \mid \langle x, y\rangle_{n,2} = 0\}$. A simple signature argument shows that the restriction of the symmetric form $\langle \cdot, \cdot\rangle_{n,2}$ to $T\mathbb{H}^{n,1}$ has Lorentzian signature, hence it makes $\mathbb{H}^{n,1}$ a Lorentzian manifold. We remark that this model is somehow the analogue of the hyperboloid model of hyperbolic space, that is

$$\mathbb{H}^n = \{y \in \mathbb{R}^{n,1} \mid q_{n,1}(y) = -1\,,\ y_{n+1} > 0\} \,, \tag{15.1.2}$$

and in fact $\mathbb{H}^n$ is isometrically embedded in $\mathbb{H}^{n,1}$ as the submanifold defined by $x_{n+2} = 0$, $x_{n+1} > 0$.

The natural action of $\mathrm{O}(n, 2)$ on $\mathbb{R}^{n,2}$ preserves $\mathbb{H}^{n,1}$ and in fact $\mathrm{O}(n, 2)$ acts by isometries on $\mathbb{H}^{n,1}$. We remark that if $x \in \mathbb{H}^{n,1}$ and $v_1, \ldots, v_{n+1}$ is an orthonormal basis of $T_x\mathbb{H}^{n,1}$ then the linear transformation of $\mathbb{R}^{n+2}$ sending the standard basis to the basis $v_1, \ldots, v_{n+1}, x$ is in $\mathrm{O}(n, 2)$. This in particular shows that $\mathrm{O}(n, 2)$ acts transitively on $\mathbb{H}^{n,1}$ and that the stabilizer of a point x acts transitively on the space of orthonormal bases of $T_x\mathbb{H}^{n,1}$. These facts imply that $\mathbb{H}^{n,1}$ has maximal isometry group and that the isometry group is indeed identified to $\mathrm{O}(n, 2)$. Notice that $\mathbb{H}^{n,1}$ can be regarded as the non-Riemannian symmetric space $\mathrm{O}(n, 2)/\mathrm{O}(n, 1)$, where $\mathrm{O}(n, 1)$ is identified with the stabilizer of $(0, 0, \ldots, 0, 1)$.

The Sectional Curvature By Lemma 15.1.1.2.4, $\mathbb{H}^{n,1}$ has constant sectional curvature. Let us now check that the sectional curvature is negative (actually,

$K = -1$). For this purpose, observe that the normal line in $\mathbb{R}^{n,2}$ of $\mathbb{H}^{n,1}$ at a point $x \in \mathbb{H}^{n,1}$ is identified with the line generated by x itself. It follows that, if v, w are tangent vector fields along $\mathbb{H}^{n,1}$, we have the orthogonal decomposition (we will omit here the subscript in the bilinear form $\langle v, w\rangle_{n,2}$, and simply write $\langle v, w\rangle$):

$$(D_v w)(x) = (\nabla_v w)(x) + \langle v, w\rangle x \ ,$$

where D is the flat connection of $\mathbb{R}^{n+2}$ and ∇ is the Levi-Civita connection of $\mathbb{H}^{n,1}$. Using the flatness of D, one easily computes that

$$R(u, v)w = \langle u, w\rangle v - \langle v, w\rangle u \ ,$$

so that

$$\langle R(u, v)v, u\rangle = -\left(\langle u, u\rangle\langle v, v\rangle - \langle v, u\rangle^2\right) \ ,$$

and this shows that $\mathbb{H}^{n,1}$ has constant sectional curvature -1. Finally, let us remark that $\mathbb{H}^{n,1}$ is not simply connected, being homeomorphic to $\mathbb{R}^n \times S^1$. (See Fig. 15.1.)

15.1.1.6 The "Klein Model" and Its Boundary

Let us now introduce a projective model, or "Klein model", for Anti-de Sitter geometry. Let us define

$$\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} := \mathbb{H}^{n,1}/\{\pm\mathbb{1}\} \ .$$

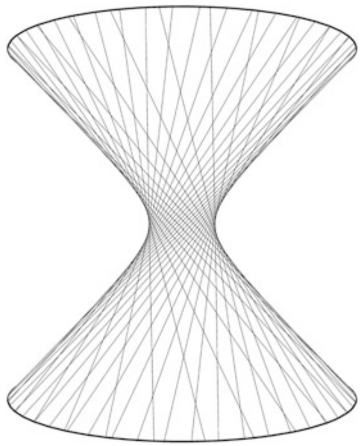


Fig. 15.1 For $n = 1$, $\mathbb{H}^{1,1}$ is the one-sheeted hyperboloid in $\mathbb{R}^{1,2}$, which is homeomorphic to the annulus $\mathbb{R} \times S^1$. The lines in the left and right rulings are lightlike geodesics

Since $\{\pm\mathbb{1}\}$ is the center of $O(n,2)$ (hence normal), $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ (endowed with the Lorentzian metric induced by the quotient) has maximal isometry group by Proposition 15.1.1.3.1 and is therefore a model of constant sectional curvature -1. It can also be shown that the center of the isometry group of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is trivial, or equivalently that $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is the quotient of its universal covering by the center of its isometry group. It follows (see also Remark 15.1.1.3.2) that it is the *minimal* model of AdS geometry, in the sense that any other model is a covering of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$.

By definition $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is naturally identified with a subset of real projective space $\mathbb{R}\mathrm{P}^{n+1}$, more precisely the subset of timelike directions of $\mathbb{R}^{n,2}$:

$$\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} = \{[x] \in \mathbb{R}\mathrm{P}^{n+1} \mid q_{n,2}(x) < 0\} .$$

Like in hyperbolic geometry, the boundary of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ in projective space $\mathbb{R}\mathrm{P}^{n+1}$ is a quadric of signature $(n,2)$, that is the projectivization of the set of lightlike vectors in $\mathbb{R}^{n,2}$. Namely

$$\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} = \{[x] \in \mathbb{R}\mathrm{P}^{n+1} \mid q_{n,2}(x) = 0\} .$$

Isometries of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ induce projective transformations which preserve $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$.

The Conformal Lorentzian Structure of the Boundary In the rest of this subsection, in analogy with hyperbolic geometry, we shall equip $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ with a conformal Lorentzian structure that extends the conformal Lorentzian structure defined inside. This will be obtained by means of the following construction.

Given a point $\ell = \mathrm{Span}(x)$ of $\mathbb{R}\mathrm{P}^{n+1}$, the tangent space of real projective space has the canonical identification

$$T_\ell\mathbb{R}\mathrm{P}^{n+1} \cong \mathrm{Hom}(\ell, \mathbb{R}^{n+2}/\ell) .$$

As a preliminary remark, when ℓ is timelike (namely $q_{n,2}(x) < 0$), the quotient $\mathbb{R}^{n+2}/\ell$ is canonically identified with $\ell^\perp$. For any local section $\sigma : \mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \to \mathbb{R}^{n,2}$ of the canonical projection, one can then define a metric over $T\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ by

$$\langle\langle f, g\rangle\rangle_\sigma = \langle f(\sigma[x]), g(\sigma[x])\rangle_{n,2} , \tag{15.1.3}$$

for $f, g \in T_{[x]}\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \cong \mathrm{Hom}(\ell, \ell^\perp)$. It is an exercise to check that if σ_0 is a section with values in $\mathbb{H}^{n,1}$, then one recovers the previously constructed metric of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, which coincides with the pull-back of the metric over $\mathbb{R}^{n,2}$ since the differential of σ_0 identifies $T_{[x]}\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ with $T_x\mathbb{H}^{n,1} = x^\perp$. This does not hold for a general section σ, but one still recovers a conformal metric as a consequence of the easy formula

$$\langle\langle f, g\rangle\rangle_{\lambda\sigma} = \lambda^2\langle\langle f, g\rangle\rangle_\sigma \tag{15.1.4}$$

for any function λ.

Let us now turn our attention to the case that $\ell = \mathrm{Span}(x)$ is lightlike, namely $q_{n,2}(x) = 0$. In this case there is no way to define a natural metric on $\mathbb{R}^{n+2}/\ell$. However, if we let

$$\mathbb{L} = \{x \in \mathbb{R}^{n,2} \mid q_{n,2}(x) = 0\}$$

be the space of lightlike vectors, then $T_x\mathbb{L}$ is precisely $\ell^\perp$ and contains ℓ itself. In fact $T_\ell \partial \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is canonically identified to $\mathrm{Hom}(\ell, \ell^\perp/\ell)$. Moreover the bilinear form of $\mathbb{R}^{n,2}$, restricted to $\ell^\perp$, induces a well-defined, non degenerate bilinear form (of signature $(n-1, 1)$) on $\ell^\perp/\ell$, which we denote by $\langle v, w\rangle_{\ell^\perp/\ell}$.

Hence one can define a metric on $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ for any section $\sigma : \partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \to \mathbb{L}$ of the canonical projection by the formula

$$(\!(f, g)\!)_\sigma = \langle f(\sigma[x]), g(\sigma[x])\rangle_{\ell^\perp/\ell} \ , \tag{15.1.5}$$

where now $f, g \in \mathrm{Hom}(\ell, \ell^\perp/\ell)$. Here this metric can be indeed be expressed as the pull-back

$$(\!(f, g)\!)_\sigma = \langle \sigma_*(f), \sigma_*(g)\rangle_{n,2} \ , \tag{15.1.6}$$

since the degenerate metric on $T_x\mathbb{L} = \ell^\perp$ is by construction the pull-back of the metric of $\ell^\perp/\ell$ by the projection along the degenerate direction ℓ.

One again has the formula

$$(\!(f, g)\!)_{\lambda\sigma} = \lambda^2 \, (\!(f, g)\!)_\sigma \tag{15.1.7}$$

similarly to (15.1.4), and therefore the induced conformal class over $T\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is independent of the choice of σ and equips $T\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ with a conformal Lorentzian metric. The naturality of the construction implies that the isometry group of $\mathbb{H}^{n,1}$ acts by conformal transformations over $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$. Finally, let us show that this conformal Lorentzian metric is naturally the conformal compactification of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$. In fact, if σ is a section of the canonical projection $\pi : \mathbb{R}^{n,2} \to \mathbb{R}\mathrm{P}^{n+1}$, defined in a neighborhood U of a point of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, by construction the metric $(\!(\cdot,\cdot)\!)_\sigma$ over $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \cap U$ is the limit of the conformal metric associated to σ defined in $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \cap U$: this means that if (p_n, v_n) is a sequence in $T\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ that converges to $(p_\infty, v_\infty) \in T\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, then $\langle\!\langle v_n, v_n\rangle\!\rangle_\sigma(p_n) \to (\!(v_\infty, v_\infty)\!)_\sigma(p_\infty)$.

In the physics literature, the conformal Lorentzian manifold $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is known as *Einstein universe*. See for instance [13, 63, 64] for more details.

Remark 15.1.1.6.1 A conformal Lorentzian structure is equivalent to the smooth field of lightlike directions, which is also called the *light cone*. More precisely, a diffeomorphism $f : (M, g) \to (N, g')$ between Lorentzian manifold is conformal, meaning that $f^*g' = e^{2\lambda}g$ for some smooth function $\lambda : M \to \mathbb{R}$, if and only if the differentials of f and f^{-1} map causal vectors to causal vectors. If M and N have dimension ≥ 3, this is indeed equivalent to the condition that the differential of f maps lightlike vectors to lightlike vectors.

Remark 15.1.1.6.2 In order to better understand the light cone in the case of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, let us notice that if $[y] \in \partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ formula (15.1.6) implies that the lightlike vectors in $T_{[y]}\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ are the projection of vectors $x \in \mathbb{R}^{n+2}$ such that $\langle x, y\rangle_{n,2} = 0$ and $\langle x, x\rangle_{n,2} = 0$. These vectors are such that $\mathrm{Span}(x, y)$ are totally degenerate planes in $\mathbb{R}^{n,2}$, or equivalently give projective lines contained in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$. Thus the light cone in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ through $[y]$ is the union of all the projective lines contained in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ which pass through $[y]$.

15.1.1.7 The "Poincaré Model" for the Universal Cover

We have already observed that $\mathbb{H}^{n,1}$, and its double quotient $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, are not simply connected. Let us now construct a simply connected model of Anti-de Sitter geometry, namely the universal cover of $\mathbb{H}^{n,1}$ and $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$. For this purpose, let $\mathbb{H}^n$ be the hyperboloid model of hyperbolic space (defined in (15.1.2)). Then

$$\pi(y, t) = (y_1, \ldots, y_n, y_{n+1}\cos t, y_{n+1}\sin t) \tag{15.1.8}$$

defines a map $\pi : \mathbb{H}^n \times \mathbb{R} \to \mathbb{H}^{n,1}$ and it is immediate yo check that this map is a covering with deck transformations of the form $(y, t) \mapsto (y, t + 2k\pi)$ for $k \in \mathbb{Z}$. See Fig. 15.6 for a picture in dimension 3. Clearly $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$ is the universal cover also for the projective model $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, the covering map being the composition of (15.1.8) with the double quotient $\mathbb{H}^{n,1} \to \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$.

Pulling back the Lorentzian metric over $\mathbb{H}^n \times \mathbb{R}$ we get a simply connected Lorentzian manifold of constant curvature -1, which we denote by $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$. The metric of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$ is a warped product of the form

$$\pi^* g_{\mathbb{H}^{n,1}} = g_{\mathbb{H}^n} - y_{n+1}^2 dt^2 \ . \tag{15.1.9}$$

Moreover $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$ has maximal isometry group, hence we have obtained a simply connected model for AdS geometry. More precisely, we have a central extension, that is a (non split) short exact sequence

$$0 \to \mathbb{Z} \to \mathrm{Isom}(\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}) \to \mathrm{O}(n, 2) \to 1 \ .$$

It is convenient to express the metric (15.1.9) using the Poincaré model of the hyperbolic space. Recall that the disk model of the hyperbolic space is the unit disk $\mathbb{D}^n = \{\mathsf{x} \in \mathbb{R}^n \mid ||\mathsf{x}|| < 1\}$ equipped with the conformal metric $\frac{4}{(1-\mathsf{r}^2)^2}\sum d\mathsf{x}_i^2$, where $\mathsf{r}^2 := ||\mathsf{x}||^2 = \sum \mathsf{x}_i^2$. The isometry with the hyperboloid model of $\mathbb{H}^n$ is given by the transformation

$$(\mathsf{x}_1, \ldots, \mathsf{x}_n) \mapsto \left(y_1 = \frac{2\mathsf{x}_1}{1-\mathsf{r}^2}, \ldots, y_n = \frac{2\mathsf{x}_n}{1-\mathsf{r}^2},\ y_{n+1} = \frac{1+\mathsf{r}^2}{1-\mathsf{r}^2}\right) \ .$$

In conclusion $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$ has the model $\mathbb{D}^n \times \mathbb{R}$ equipped with the metric

$$\frac{4}{(1-\mathsf{r}^2)^2}(d\mathsf{x}_1^2 + \ldots + d\mathsf{x}_n^2) - \left(\frac{1+\mathsf{r}^2}{1-\mathsf{r}^2}\right)^2 d\mathsf{t}^2 \,. \tag{15.1.10}$$

The "Poincaré model" of the AdS geometry, which has been introduced in [27], is then the cylinder $\mathbb{D}^n \times \mathbb{R}$ equipped with the metric (15.1.10). From Eq. (15.1.10), each slice $\{\mathsf{t} = \mathsf{c}\}$ is a totally geodesic copy of $\mathbb{H}^n$, a fact which will be evident also from other reasons in Sect. 15.1.1.8. The expression (15.1.10) also shows that the vector field $\partial/\partial\mathsf{t}$ is a timelike non-vanishing vector field on $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$, which shows that $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$ is time-orientable. Since any choice of time orientation is preserved by the action of deck transformations of the covering $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1} \to \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, this shows that also $\mathbb{H}^{n,1}$ and $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ are time-orientable. Notice however that vertical lines in the metric are not geodesic (15.1.10), except for the central line, passing through $\mathsf{x}_1 = \ldots = \mathsf{x}_n = \mathsf{t} = 0$.

The Conformal Metric of the Boundary Using the covering map from $\mathbb{H}^n \times \mathbb{R}$ to $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, we can endow $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$ (and similarly any other covering of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$) with a natural boundary. Concretely, the covering map (now in the projective model of $\mathbb{H}^n$)

$$\pi'([y_1 : \ldots : y_n, : y_{n+1}], t) = [y_1 : \ldots : y_n : y_{n+1}\cos t : y_{n+1}\sin t]$$

extends to $\pi' : (\mathbb{H}^n \cup \partial\mathbb{H}^n) \times \mathbb{R} \to \mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \cup \partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$. To compute the conformal Lorentzian structure of the boundary, we consider the map $\tau : \mathbb{H}^n \times \mathbb{R} \to \mathbb{R}^{n+2}$ defined by

$$\tau([y_1 : \ldots : y_n, : y_{n+1}], t) = (y_1/y_{n+1}, \ldots, y_n/y_{n+1}, \cos t, \sin t)$$

which clearly extends to the boundary, and induces a (local) section of the projection $\mathbb{R}^{n+2} \to \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$. In fact, if η is the generator of the group of deck transformations of the covering $\pi' : \widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1} \to \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, then τ has the equivariance $\tau \circ \eta^i = (-1)^i \tau$. Using also (15.1.7), the conformal Lorentzian metric on $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$ induced by σ by means of Eq. (15.1.5) is the pull-back of a Lorentzian metric compatible with the natural conformal structure of the boundary $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$. A direct computation (which becomes very simple by using the metric (15.1.10), the formula (15.1.4) and the observation that τ differs by the hyperboloid section by the factor $y_{n+1} = \frac{1+\mathsf{r}^2}{1-\mathsf{r}^2}$) gives the expression

$$\frac{4}{(1+\mathsf{r}^2)^2}(d\mathsf{x}_1^2 + \ldots + d\mathsf{x}_n^2) - d\mathsf{t}^2 \,. \tag{15.1.11}$$

This metric extends to $\overline{\mathbb{D}}^n \times \mathbb{R}$ and thus the metric $g_{\mathbb{S}^{n-1}} - dt^2$ on $\mathbb{S}^{n-1} \times \mathbb{R}$, where $g_{\mathbb{S}^{n-1}}$ is the round metric over the sphere, is compatible with the conformal Lorentzian structure of $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$. This also shows that the conformal structure of $\partial\mathbb{H}^{n,1} \cong \mathbb{S}^{n-1} \times \mathbb{S}^1$ admits the representative $g_{\mathbb{S}^{n-1}} - g_{\mathbb{S}^1}$, and the conformal structure of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is compatible with the double quotient of the latter, by the involution $(p, q) \mapsto (-p, -q)$ on $\mathbb{S}^{n-1} \times \mathbb{S}^1$.

15.1.1.8 Geodesics

Let us now study more precise properties of AdS geometry, concerning its geodesics.

In the Quadric Model Let us start with the exponential map in the hyperboloid model. Given a point $x \in \mathbb{H}^{n,1}$ and $v \in T\mathbb{H}^{n,1}$ we shall determine the geodesic through x with speed v. Let us distinguish several cases according to the sign of $q_{n,2}(v)$. If v is lightlike, then

$$\gamma(t) = x + tv$$

is a geodesic of $\mathbb{R}^{n,2}$ and is contained in $\mathbb{H}^{n,1}$, hence γ is a geodesic for the intrinsic metric. See Fig. 15.1.

If v is either timelike or spacelike, we claim that the geodesic $\gamma(t) = \exp_x(tv)$ is contained in the linear plane $W = \mathrm{Span}(x, v)$. In fact, the linear transformation T that fixes pointwise W and whose restriction to $W^\perp$ is $-\mathbb{1}_{W^\perp}$ is in $\mathrm{O}(n, 2)$. By the uniqueness of the solutions of the geodesic equation, $T \circ \gamma = \gamma$ hence γ is necessarily contained in $\mathbb{H}^{n,1} \cap W$. One can then easily derive the expressions

$$\gamma(t) = \cosh(t)x + \sinh(t)v \tag{15.1.12}$$

if $q_{n,2}(v) = 1$ and

$$\gamma(t) = \cos(t)x + \sin(t)v \tag{15.1.13}$$

if $q_{n,2}(v) = -1$.

In the Klein Model In analogy with the hyperbolic case, in the Klein model $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ geodesics are intersection of projective lines with the domain $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \subset \mathbb{R}\mathrm{P}^{n+1}$. From the above discussion,

- Timelike geodesics correspond to projective lines that are entirely contained in $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, are closed non-trivial loops and have length π.
- Spacelike geodesics correspond to lines that meet $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ transversally in two points. They have infinite length.
- Lightlike geodesics correspond to lines tangent to $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$.

In particular the light cone through a point $[x] \in \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ coincides with the cone of lines through $[x]$ tangent to $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$. See Fig. 15.2 for a picture (in dimension 3) in an affine chart, where geodesics look like straight lines. For instance in the affine chart $\mathbb{A}_{n+2} = \{x_{n+2} \neq 0\}$, where in coordinates $(y_1, \ldots, y_{n+1}) =$

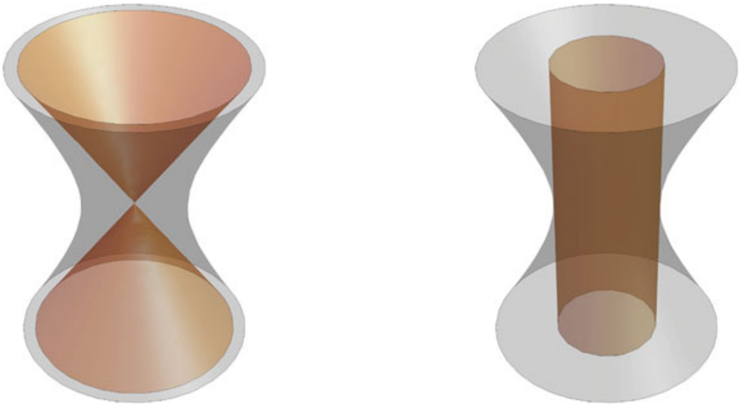


Fig. 15.2 The projective model of three-dimensional AdS space in an affine chart. The interior quadric is the lightcone from the point $[0 : 0 : 0 : 1]$, which is tangent to the boundary as explained in Sect. 15.1.1.8, pictured in the affine charts $x_4 \neq 0$ (left) and $x_3 \neq 0$ (right)

$(x_1/x_{n+2}, \ldots, x_{n+1}/x_{n+2})$, the intersection $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \cap \mathbb{A}_{n+2}$ is the interior of a one-sheeted hyperboloid, that is,

$$\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \cap \mathbb{A}_{n+2} = \{y_1^2 + \ldots + y_n^2 - y_{n+1}^2 < 1\} ,$$

while its boundary is the one-sheeted hyperboloid itself:

$$\partial \mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \cap \mathbb{A}_{n+2} = \{y_1^2 + \ldots + y_n^2 - y_{n+1}^2 = 1\} .$$

In an affine chart, timelike geodesics corresponds to affine lines which are entirely contained in the Anti de Sitter space, and which are not asymptotic to its boundary; lightlike geodesics are tangent to the one-sheeted hyperboloid, or are asymptotic to it (tangent at infinity).

Remark 15.1.1.8.1 An important observation concerns the space of timelike geodesics. Any timelike line is the projectivisation of a negative definite plane. As $\mathrm{Isom}(\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}) \cong \mathrm{PO}(n, 2)$ acts transitively on the space of timelike lines, and since the stabiliser of a timelike line is the group $\mathrm{P}(\mathrm{O}(n) \times \mathrm{O}(2))$ which is the maximal compact subgroup of $\mathrm{PO}(n, 2)$, the space of timelike geodesics of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is naturally identified with the Riemannian symmetric space of $\mathrm{PO}(n, 2)$.

Totally Geodesic Subspaces Before discussing the geodesics in the Poincaré model, let us briefly discuss more in general totally geodesics subspaces. By an argument analogous to the case of geodesics, totally geodesic subspaces of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ of dimension k are obtained as the intersection of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ with the projectivisation $\mathrm{P}(W)$ of a linear subspace W of $\mathbb{R}^{n,2}$ of dimension $k + 1$. The negative index of W can be either 2 or 1, for otherwise the intersection $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \cap \mathrm{P}(W)$ would be empty.

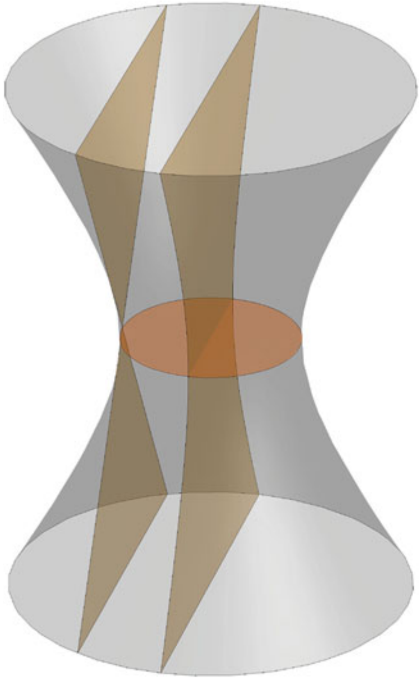


Fig. 15.3 In an affine chart for $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, a spacelike plane (horizontal), which intersects a timelike plane (vertical) in a spacelike geodesic. A lightlike plane (on the left) is tangent to $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ at a point

We have several cases—see Fig. 15.3:

- If W has signature $(k-1, 2)$, then $\mathrm{P}(W) \cap \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is isometric to $\mathbb{A}\mathrm{d}\mathbb{S}^{k-1,1}$.
- If W has signature $(k-2, 1)$, then it is a copy of Minkowski space $\mathbb{R}^{k-2,1}$, hence $\mathrm{P}(W) \cap \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is a copy of the Klein model of hyperbolic space.
- If W is degenerate, then $\mathrm{P}(W) \cap \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is a lightlike subspace foliated by lightlike geodesics tangent to the same point of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$.

A particular case of the last point is when W is degenerate and $\dim W = n + 1$. Then $\mathrm{P}(W) \cap \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is a projective hyperplane tangent to $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ at a point $[x]$ and $\mathrm{P}(W) \cap \partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is the lightlike cone of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ through $[x]$ (Remark 15.1.1.6.2).

In the Universal Cover In the universal cover $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$, geodesics are the lifts of the geodesics of the models $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ or $\mathbb{H}^{n,1}$ which we have just described. Hence every lightlike or spacelike geodesic in $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ and $\mathbb{H}^{n,1}$, which is topologically a line, has a countable number of lifts to $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$. On the other hand timelike geodesics in $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ and $\mathbb{H}^{n,1}$ are topologically circles and are in bijections with timelike geodesics of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$, as the covering map from $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$, restricted to a timelike geodesic, induces a covering map onto the circle.

Using the Poincaré model for the universal cover, introduced in Sect. 15.1.1.7, it is easy to give an explicit description of (unparameterized) lightlike geodesics. In fact, in Lorentzian geometry not only the nature of a vector (i.e. timelike, lightlike or spacelike) is conformally invariant, but also unparameterized lightlike are a conformal invariant. More concretely, the following holds, see for instance [66, Proposition 2.131].

Theorem 15.1.1.8.2 *If two Lorentzian metrics g and g' on a manifold M are conformal, then they have the same unparameterized lightlike geodesics.*

As a consequence of Theorem 15.1.1.8.2, we can replace the Poincaré metric (15.1.10) by the conformal metric given by (15.1.11):

$$\frac{4}{(1+\mathsf{r}^2)^2}(d\mathsf{x}_1^2+\ldots+d\mathsf{x}_n)^2-d\mathsf{t}^2\,. \tag{15.1.14}$$

Now observe that the first term in the expression (15.1.14) is exactly the form of the spherical metric on a hemisphere, pulled-back to the unit disc by means of the stereographic projection. We will call such a metric the *hemispherical* metric and we will denote it, with a small abuse of notation, by $g_{\mathbb{S}^n}$. In other words, the conformal metric (15.1.14) is isometric to $g_{\mathbb{S}^n}-dt^2$ on the product of a hemisphere and the line. The boundary of $\partial\mathbb{D}$ is an equator for the hemispherical metric, and in fact it is the only equator completely contained in $(\mathbb{D}\cup\partial\mathbb{D}, g_{\mathbb{S}^n})$, which justifies the fact that it will be called *the* equator for simplicity.

As a consequence, unparameterized lightlike geodesics of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$ going through a point $(\mathsf{p}_0,\mathsf{t}_0)$ are characterized by the conditions that they are mapped to spherical geodesics under the vertical projection $(\mathsf{p},\mathsf{t})\mapsto\mathsf{p}$ and moreover

$$\mathsf{t}-\mathsf{t}_0=d_{\mathbb{S}^n}(\mathsf{p},\mathsf{p}_0) \tag{15.1.15}$$

on the geodesic. In particular, these lightlike geodesics meet the boundary of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1}$ at the points which satisfy (15.1.15) such that p is on the equator of the hemisphere: as an example, if p_0 is the center of the hemisphere, then the points at infinity of the lightcone over $(\mathsf{p}_0,\mathsf{t}_0)$ are the horizontal slice $\mathsf{t}=\mathsf{t}_0+\pi/2$. This sphere is also the boundary of a hyperplane dual to $(\mathsf{p}_0,\mathsf{t}_0)$, see next section.

The same argument also permits to describe explicitely a lightlike hyperplane in the Poincaré model for the universal cover: the lightlike hyperplane having $(\mathsf{p}_0,\mathsf{t}_0)$ as a past endpoint, (where now p_0 is on the equator) is precisely $\{(\mathsf{p},\mathsf{t})\,|\,\mathsf{t}-\mathsf{t}_0=d_{\mathbb{S}^n}(\mathsf{p},\mathsf{p}_0)\}$, and its future endpoint is $(-\mathsf{p}_0,\mathsf{t}+\pi)$. See Fig. 15.4 for pictures in dimension $2+1$.

15.1.1.9 Polarity in Anti-de Sitter Space

The quadratic form $q_{n,2}$ induces a polarity on the projective space $\mathbb{R}\mathrm{P}^{n+1}$, namely the correspondence which associates to the projective subspace $\mathrm{P}(W)$ the subspace $\mathrm{P}(W^\perp)$. In particular this correspondence induces a duality between spacelike totally geodesic subspaces of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$: the dual of a spacelike k-dimensional subspace is a $n-k+1$ subspace. For instance the dual of a point $[x]$ is a n-dimensional spacelike hyperplane $P_{[x]}=\mathrm{P}(x^\perp)$. Projectively $P_{[x]}$ is characterised as the hyperplane spanned by the intersection of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{n-1,1}$ with the lightcone from

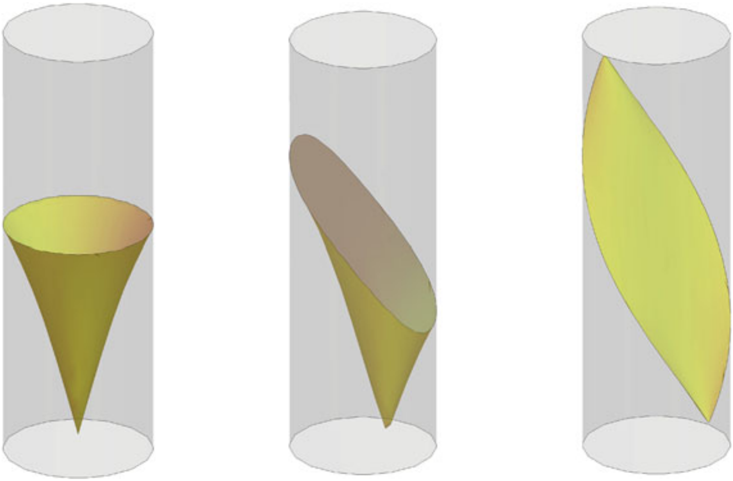


Fig. 15.4 In the left and middle pictures, future lightcones over a point in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$. In the left picture the basepoint of the lightcone projects to the center of the disc, and therefore the closure of the lightcone in the cylinder $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ is a horizontal slice. In the right picture, a lightlike plane, which is actually the degenerate limit of future lightcones as the basepoints tend to the boundary

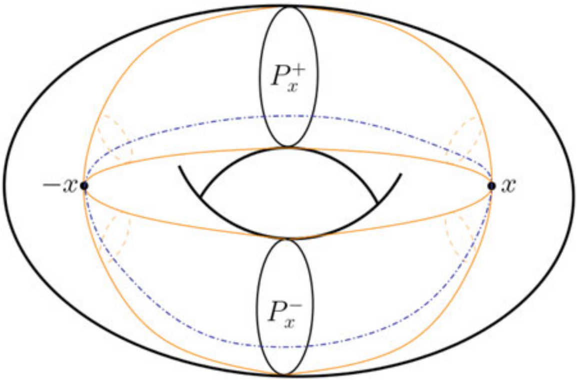


Fig. 15.5 The duality in $\mathbb{H}^{2,1}$, which is the interior of a solid torus. The lightcone from a point x is tangent to $\partial\mathbb{H}^{2,1}$ in two meridians, which span the dual planes $P_x^{\pm}$. Timelike geodesics through x intersect $P_x^{\pm}$ orthogonally and all meet again at the antipodal point $-x$. The region U_x is the solid cylinder bounded by $P_x^{\pm}$ and containing x

$[x]$. More geometrically, it can be checked that $P_{[x]}$ is the set of antipodal points to $[x]$ along timelike geodesics through $[x]$. Also, every timelike geodesic through $[x]$ meets $P_{[x]}$ orthogonally at time $\pi/2$. Conversely, given a totally geodesic spacelike hyperplane H, all the timelike geodesics that meet H orthogonally intersect in a single point, which is the dual point of H.

In the Quadric Model To some extent, the duality between points and planes lifts to the coverings of $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$. In $\mathbb{H}^{n,1}$ there are two dual planes associated to any point x: the sets

$$P_x^{\pm} = \{\exp_x(\pm(\pi/2)v) \mid q_{n,2}(v) = -1,\ v \text{ future-directed}\}\ .$$

Clearly P_x^+ and P_x^- are antipodal and $P_{-x}^{\pm} = P_x^{\mp}$. The planes $P_x^{\pm}$ disconnect $\mathbb{H}^{n,1}$ in two regions U_x and U_{-x}, where U_x is the region containing x. See Fig. 15.5. They can be characterised by

$$U_x = \{y \in \mathbb{H}^{n,1} \mid \langle x, y\rangle_{n,1} < 0\}\ .$$

Spacelike and lightlike geodesics through x do not exit U_x, while all the timelike geodesics through x meet orthogonally $P_x^{\pm}$ and all pass through the point $-x$. More precisely, a point $y \neq x$ is connected to x:

- by a spacelike geodesic if and only if $\langle x, y\rangle_{n,1} < -1$,
- by a lightike geodesic if and only if $\langle x, y\rangle_{(n,1)} = -1$,
- by a timelike geodesic if and only if $|\langle x, y\rangle_{(n,1)}| < 1$.

(To check this, see also the expressions of geodesics in Sect. 15.1.1.8.) An immediate consequence is that if y is connected to x by a spacelike geodesic, there is no geodesic joining y to $-x$. Hence the exponential map of $\mathbb{H}^{n,1}$ is not surjective. But as any point $y \in \mathbb{H}^{n,1}$ can be connected through a geodesic either to x or to $-x$, the exponential over $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is surjective.

In the Universal Cover Finally, let us consider the situation in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1} \cong \mathbb{H}^n \times \mathbb{R}$. Recall that the group of deck transformations for the covering $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1} \to \mathbb{H}^{n,1}$ is $\mathbb{Z}$, where a generator acts by translations of 2π in the $\mathbb{R}$ factor. Hence the preimage of a spacelike plane $P \subset \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$ is the disjoint union of spacelike planes $(P^k)_{k\in\mathbb{Z}}$, enumerated so that the generator η of $\mathbb{Z}$ acts by sending P^k to P^{k+1}. Moreover each connected component of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1} \setminus \bigcup_{k\in\mathbb{Z}} P^k$ is a fundamental domain for the action of deck transformations of the covering $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{n,1} \to \mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$.

Now given a point x, let us apply the previous construction to the plane $P_x = P_{\pi'(x)}$ which is the dual of the image $\pi'(x)$ in $\mathbb{A}\mathrm{d}\mathbb{S}^{n,1}$, and let V_x be the connected component which contains x. We will refer to V_x as the *Dirichlet domain* in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ centered at x, since the construction of V_x is the analogue of a Dirichlet domain in this context. Then the restricted covering map $\pi'|_{V_x} : V_x \to \mathbb{A}\mathrm{d}\mathbb{S}^{n,1} \setminus P_x$ is an isometry. Therefore lightlike and spacelike geodesics through x are entirely contained in V_x. See Fig. 15.6.

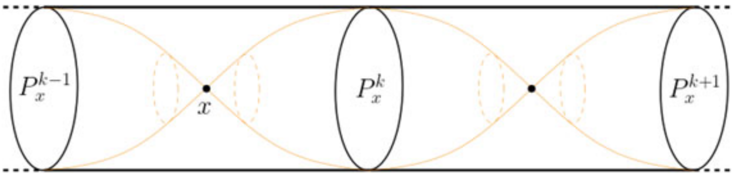


Fig. 15.6 A topological picture of the universal cover $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$. The planes P_x^k are spacelike and differ by deck transformations. The Dirichlet domain V_x is a solid cylinder containing x, bounded by P_x^{k-1} and P_x^k

15.1.2 Anti de Sitter Space in Dimension $(2+1)$

The purpose of this section is to focus on some peculiarites of Anti-de Sitter geometry in dimension three.

15.1.2.1 The PSL(2, $\mathbb{R}$)-Model

The fundamental observation is the existence of a special model in dimension three which naturally endows Anti-de Sitter space with a Lie group structure. To construct this, consider the vector space $\mathcal{M}(2,\mathbb{R})$ of 2×2 matrices with real entries. Then $q = -\det$ is a quadratic form with signature $(2, 2)$, hence there is an isomorphic identification between $(\mathcal{M}(2,\mathbb{R}), -\det)$ and $(\mathbb{R}^{2,2}, q_{2,2})$, unique up to composition by elements in $\mathrm{O}(2, 2)$. Under this isomorphism $\mathbb{H}^{2,1}$ is identified with the Lie group $\mathrm{SL}(2,\mathbb{R})$.

Let us notice that $\mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{R})$ acts linearly on $\mathcal{M}(2,\mathbb{R})$ by left and right multiplication:

$$(A, B) \cdot X := AXB^{-1} . \tag{15.1.16}$$

As a simple consequence of the Binet Formula, this action preserves the quadratic form $q = -\det$ and thus induces a representation

$$\mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{R}) \to \mathrm{O}(\mathcal{M}(2,\mathbb{R}), q) .$$

Since the center of $\mathrm{SL}(2,\mathbb{R})$ is $\pm\mathbb{1}$, the kernel of such a representation is given by $K = \{(\mathbb{1}, \mathbb{1}), (-\mathbb{1}, -\mathbb{1})\}$, and by a dimensional argument it turns out that the image of the representation is the connected component of the identity:

$$\mathrm{Isom}_0(\mathbb{H}^{2,1}) \cong \mathrm{SO}_0(\mathcal{M}(2,\mathbb{R}), q) \cong (\mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{R}))/K .$$

Using this model, one then has a natural identification of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ with the Lie group $\mathrm{PSL}(2,\mathbb{R})$, in such a way that

$$\mathrm{Isom}_0(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}) \cong \mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R}) \tag{15.1.17}$$

acting by left and right multiplication on $\mathrm{PSL}(2, \mathbb{R})$.

The stabilizer of the identity in $\mathrm{Isom}_0(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1})$ is the diagonal subgroup $\Delta < \mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$. Under the obvious identification of $\mathrm{PSL}(2, \mathbb{R})$ and Δ, the action of the identity stabilizer on the Lie algebra $\mathfrak{sl}(2, \mathbb{R}) = T_{\mathbb{1}}\mathrm{PSL}(2, \mathbb{R})$ is the adjoint action of $\mathrm{PSL}(2, \mathbb{R})$. A direct consequence of this construction is the bi-invariance of the quadratic form q. Indeed, denoting by $q_{\mathbb{1}}$ the restriction of q to $T_{\mathbb{1}}\mathrm{SL}(2, \mathbb{R})$, a direct computation shows that $q_{\mathbb{1}}$ equals $(1/8)\kappa$, where $\kappa(X, Y) = 4\mathrm{tr}\,(XY)$ is the Killing form of $\mathfrak{sl}(2, \mathbb{R})$.

Remark 15.1.2.1.1 The Lie algebra $\mathfrak{sl}(2, \mathbb{R})$ equipped with the quadratic form $q_{\mathbb{1}}$ is then a copy of the 3-dimensional Minkowski space, hence the adjoint action yields a representation

$$\mathrm{PSL}(2, \mathbb{R}) \to \mathrm{O}(\mathfrak{sl}(2, \mathbb{R}), q_{\mathbb{1}})$$

which in turn induces the well-known isomorphism

$$\mathrm{SO}_0(2, 1) \cong \mathrm{SO}_0(\mathfrak{sl}(2, \mathbb{R}), q_{\mathbb{1}}) \cong \mathrm{PSL}(2, \mathbb{R}) \; ,$$

which is nothing but the restriction of the isomorphism (15.1.17) to the stabilizer of the identity in the left-hand side $\mathrm{Isom}_0(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1})$, and to the diagonal subgroup Δ in the right-hand side $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$.

Remark 15.1.2.1.2 The identification between $\mathbb{H}^{2,1}$ and $\mathrm{SL}(2, \mathbb{R})$ parallels the more classical identification between the three sphere S^3 and the Lie group $\mathrm{SU}(2)$. The analogy can be deepened by considering the isomorphism of $\mathfrak{gl}(2, \mathbb{R})$ with the algebra of pseudo-quaternions, namely the four-dimensional real algebra generated by $1, i, j, k$ with the relations $-i^2 = j^2 = k^2 = 1$ and $k = ij = -ji$. Under this isomorphism the quadratic form det corresponds to

$$q(a + bi + cj + dk) = a^2 + b^2 - c^2 - d^2 \; ,$$

hence $\mathbb{H}^{2,1}$ is identified to the set of unitary pseudo-quaternions.

15.1.2.2 The Boundary of PSL(2, $\mathbb{R}$)

From the identification between $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ and $\mathrm{PSL}(2, \mathbb{R})$, we obtain an identification of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ with the boundary of $\mathrm{PSL}(2, \mathbb{R})$ into $\mathrm{P}(\mathcal{M}(2, \mathbb{R}))$, which is the projectivization of the cone of rank 1 matrices. Therefore from now on we shall always consider

$$\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} = \{[X] \in \mathrm{P}(\mathcal{M}(2, \mathbb{R})) \,|\, \mathrm{rank}(X) = 1\} \; .$$

We have a homeomorphism

$$\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \to \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$$

which is defined by

$$[X] \mapsto (\mathrm{Im}X, \mathrm{Ker}X) ,$$

and is equivariant under the actions of $\mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$: the obvious action on $\mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$, and the action on $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ induced by (15.1.16).

Lemma 15.1.2.2.1 *The inversion map $\iota[X] = [X]^{-1}$ is a time-reversing isometry of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ which induces the homeomorphism $(x, y) \mapsto (y, x)$ on $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cong \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$.*

Proof Clearly ι is equivariant with respect to the isomorphism of $\mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$ which switches the two factors. To show that it is an isometry it thus suffices to check that its differential at the identity is a linear isometry, which is obvious since $d_{\mathbb{1}}\iota$ is minus the identity, which also shows time-reversal. The second claim is easily checked by observing that for an invertible 2×2 matrix we have $(\det X)X^{-1} = (\mathrm{tr}\, X)\mathbb{1} - X$ by the Cayley–Hamilton theorem, so that projectively $[X^{-1}] = [\mathrm{tr}\, X\mathbb{1} - X]$. This shows that the inversion map of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ extends to the transformation $[X] \to [\mathrm{tr}\, X\mathbb{1} - X]$ along the boundary. If X is a rank 1 matrix, then it is traceless if and only if $X^2 = 0$, that is, if and only if $\mathrm{Ker}X = \mathrm{Im}X$. So in this case the statement is easily proved. If $\mathrm{tr}\, X \neq 0$, then X is diagonalizable with eigenvalues 0, and $\mathrm{tr}\, X$. Moreover $\mathrm{Ker}X$ and $\mathrm{Im}X$ are the corresponding eigenspaces. It is easily seen that $\mathrm{Ker}(\mathrm{tr}\, X\mathbb{1} - X) = \mathrm{Im}X$ and $\mathrm{Im}(\mathrm{tr}\, X\mathbb{1} - X) = \mathrm{Ker}X$. □

Using the upper half-plane model for the hyperbolic space $\mathbb{H}^2$, $\mathbb{R}\mathrm{P}^1$ corresponds to the boundary at infinity $\partial\mathbb{H}^2$ and $\mathrm{PSL}(2,\mathbb{R})$ is identified to $\mathrm{Isom}_0(\mathbb{H}^2)$, which acts on $\mathbb{R}\mathrm{P}^1$ in the canonical way. One can therefore consider $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ as $\partial\mathbb{H}^2 \times \partial\mathbb{H}^2$. We can interpret the convergence to $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ in this setting.

Lemma 15.1.2.2.2 *A sequence $[X_n] \in \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ converges to $(x, y) \in \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cong \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$ if and only if for every $p \in \mathbb{H}^2$, $X_n(p) \to x$ and $X_n^{-1}(p) \to y$.*

Proof Since the action of $\mathrm{PSL}(2,\mathbb{R})$ on $\mathbb{H}^2$ is isometric, if the condition holds for some p, then it holds for all $p \in \mathbb{H}^2$. Hence one can take for instance $p = i$ in the upper half-plane. Assuming X_n converges projectively to a rank 1 matrix X, one checks immediately that $X(p)$ is in the projective class of $x = \mathrm{Im}(X)$. The convergence $X_n^{-1}(p) \to y$ then follows by Lemma 15.1.2.2.1. □

In this dimension, $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is a double ruled quadric, which in an affine chart looks like in Fig. 15.1. We shall now describe geometrically these rulings. Given any $(x_0, y_0) \in \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$,

$$\lambda_{y_0} := \{(x, y_0) \mid x \in \mathbb{R}\mathrm{P}^1\}$$

describes a projective line in $\mathbb{R}\mathrm{P}^3$ which is contained in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, hence lightlike for the conformal Lorentzian structure of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ by Remark 15.1.1.6.2. In fact, λ_{y_0} is the orbit of (x_0, y_0) by the action of $\mathrm{PSL}(2,\mathbb{R}) \times \{\mathbb{1}\}$, or by the (now free) action of $PSO(2) \times \{\mathbb{1}\}$, where PSO(2) corresponds to a 1-parameter elliptic subgroup in $PSL(2, \mathbb{R})$. In short,

$$\lambda_{y_0} = PSL(2, \mathbb{R}) \cdot (x_0, y_0) = PSO(2) \cdot (x_0, y_0) .$$

We refer to λ_{y_0} as the *left ruling* through (x_0, y_0), and similarly the *right ruling* is

$$\mu_{x_0} := \{(x_0, y) \mid y \in \mathbb{R}P^1\} ,$$

for which analogous considerations hold.

We conclude this section by remarking that the conformal Lorentzian structure on $\partial\mathbb{A}d\mathbb{S}^{2,1}$ is easily expressed in terms of the left and right rulings. Let us start by carefully choosing a time-orientation on $\mathbb{A}d\mathbb{S}^{2,1}$. Orienting $\mathbb{R}P^1$ in the usual way, consider the induced orientation on PSO(2). We remark that PSO(2) is a timelike geodesic of $\mathbb{A}d\mathbb{S}^{2,1}$ and we choose the time orientation on $\mathbb{A}d\mathbb{S}^{2,1}$ in such a way that PSO(2) oriented as above is future directed. Observe that the action of $PSO(2) \times \{\mathbb{1}\}$ on $\mathbb{A}d\mathbb{S}^{2,1}$ yields a flow on $\mathbb{A}d\mathbb{S}^{2,1}$ generated by a right-invariant vector field, which at $\mathbb{1}$ is the *positive* tangent vector of PSO(2). So orbits are all timelike and *future directed*. Similarly $\{\mathbb{1}\} \times PSO(2)$ yields a flow generated by a left-invariant vector field, which at $\mathbb{1}$ is the *negative* tangent vector of PSO(2), and its orbits are all timelike and *past directed*.

Proposition 15.1.2.2.3 *Let* $\pi_l, \pi_r : \mathbb{R}P^1 \times \mathbb{R}P^1 \to \mathbb{R}P^1$ *be the canonical projections and* $d\theta$ *the angular form on* $\mathbb{R}P^1 \cong \partial\mathbb{H}^2$. *Then the symmetric product* $\pi_l^*(d\theta)\pi_r^*(d\theta)$ *is in the conformal class of* $\partial\mathbb{A}d\mathbb{S}^{2,1}$.

Proof Since we already know that the left and right rulings are lightlike for the conformal class of $\partial\mathbb{A}d\mathbb{S}^{2,1}$, it only remains to check the sign by Remark 15.1.1.6.1. Notice that λ_{y_0} is the orbit of the action of $PSO(2) \times \{\mathbb{1}\}$, while μ_{x_0} is the orbit of the action of $\{\mathbb{1}\} \times PSO(2)$. Then λ_{y_0} with the obvious parameterization is future directed while μ_{x_0} is past directed. The result follows. □

Therefore a C^1 curve in $\partial\mathbb{A}d\mathbb{S}^{2,1}$ is spacelike when it is locally the graph of an orientation-preserving function, and timelike when it is locally the graph of an orientation-reversing function. Given two intervals I_1 and I_2 in $\partial\mathbb{H}^2$ and assuming θ_1 and θ_2 are angle determinantion over I_1 and I_2, the future $I^+_{I_1 \times I_2}(p_0, q_0)$ of a point (p_0, q_0) in $I_1 \times I_2$ is region where $\theta_1(p) - \theta_1(p_0) > 0$ and $\theta_2(q) - \theta_2(q_0) < 0$, while the past is determined by reversing both inequalities. In conclusion

$$I^+_{I_1 \times I_2}(p_0, q_0) \cup I^-_{I_1 \times I_2}(p_0, q_0) = \{(p, q) \in I_1 \times I_2 \mid (\theta_1(p) - \theta_1(p_0))(\theta_2(q) - \theta_2(q_0)) < 0\} . \quad (15.1.18)$$

15.1.2.3 Levi-Civita Connection

In this section we shall describe the properties of natural metric connections on $\mathbb{A}d\mathbb{S}^{2,1}$, for which the theory of Lie groups permits to give a transparent description. Let us start by some general facts of Lie groups.

Recall that the Lie bracket on the Lie algebra $\mathfrak{g} = T_{\mathbb{1}}G$ of a Lie group G is defined as

$$[V, W]_{\mathfrak{g}} = [\widetilde{V}, \widetilde{W}](\mathbb{1}) = -[\widetilde{V}', \widetilde{W}'](\mathbb{1}) , \tag{15.1.19}$$

where $[\cdot, \cdot]$ now denotes the bracket of vector fields and $\widetilde{V}$, $\widetilde{W}$ (resp. $\widetilde{V}'$, $\widetilde{W}'$) are the left-invariant (resp. right-invariant) vector fields extending V and W respectively.

Now, any Lie group G is equipped with two natural connections, the *left-invariant connection* D^l and the *right-invariant connection* D^r. The former is uniquely determined by the condition that left-invariant vector fields are parallel, and is left-invariant in the sense that, if $L_g : G \to G$ denotes left multiplication by g, then

$$(L_g)_*(D^l_V W) = D^l_{(L_g)_*(V)}(L_g)_*(W) .$$

The left-invariant connection D^l at a point $g \in G$ can be easily expressed as ordinary differentiation in T_gG, after pulling-back a vector field W to g by left multiplication. More precisely,

$$D^l_V W = \left.\frac{d}{dt}\right|_{t=0} (L_{g\gamma(t)^{-1}})_*(W_{\gamma(t)}) , \tag{15.1.20}$$

where $\gamma(t)$ is a path with $\gamma(0) = g$ and $\gamma'(0) = V$.

The analogous definition and property holds for D^r, replacing left-invariant by right-invariant vector fields. Both connections D^l and D^r are flat and are compatible with any metric on G which is left-invariant or right-invariant respectively. Indeed parallel transport of a vector $W \in T_gG$ to $T_{g'}G$ consists just in left (resp. right) multiplication, namely in applying $(L_{g'g^{-1}})_*$ (resp. $(R_{g'g^{-1}})_*$) to W, and is therefore path-independent.

But D^l and D^r are not torsion-free, as can be easily checked by the definition of torsion, which we recall is a tensor of type (2, 1). For instance, computing at the identity and using left-invariant extensions $\widetilde{V}$ and $\widetilde{W}$ of V and W, one obtains

$$\tau^l(V, W) = D^l_V \widetilde{W} - D^l_W \widetilde{V} - [\widetilde{V}, \widetilde{W}](\mathbb{1}) = -[\widetilde{V}, \widetilde{W}](\mathbb{1}) = -[V, W]_{\mathfrak{g}} .$$

Similarly one obtains

$$\tau^r(V, W) = [V, W]_{\mathfrak{g}} .$$

By construction, τ^l is left-invariant and τ^r is right-invariant. But by Ad-invariance of the Lie bracket of $\mathfrak{g}$, the torsions τ^l and τ^r are actually bi-invariant.

Moreover, a direct computation shows that the tensorial quantity $D^r - D^l$ admits the following expression at the identity:

$$D^r_V W - D^l_V W = [V, W]_{\mathfrak{g}} . \tag{15.1.21}$$

To check Eq. (15.1.21), it suffices to consider the right-invariant extension $\widetilde{W}$ of W, so that $D^r_V \widetilde{W} = 0$. Using the expression (15.1.20) for D^l at the identity, we see that

$$\begin{aligned} D^l_V \widetilde{W} &= \frac{d}{dt}\bigg|_{t=0} (L_{\exp(-tV)})_*(\widetilde{W}_{\exp(tV)}) \\ &= \frac{d}{dt}\bigg|_{t=0} (L_{\exp(-tV)})_*(R_{\exp(tV)})_*(W) = -\mathrm{ad}_V(W) = -[V, W]_{\mathfrak{g}} , \end{aligned}$$

which thus shows Eq. (15.1.21).

Now, given a bi-invariant pseudo-Riemannian metric on G, its Levi-Civita connection ∇ can be expressed as the mid-point between D^l and D^r. Namely, using now V and W to denote vector fields,

$$\nabla_V W = \frac{1}{2}\left(D^l_V W + D^r_V W\right) , \tag{15.1.22}$$

which is still a connection on G since the space of connections forms an affine space with underlying vector space the space of $(2, 1)$-tensors. Indeed ∇ is still compatible with the metric and is moreover torsion-free, since its torsion, which equals $(\tau^l + \tau^r)/2$, vanishes.

A direct consequence of Eqs. (15.1.21) and (15.1.22) is the following well-known expression for the Levi-Civita connection in terms of left-invariant vector fields:

Lemma 15.1.2.3.1 *Given left-invariant vector fields V and W on G, the Levi-Civita connection of a bi-invariant metric has the expression:*

$$\nabla_V W = \frac{1}{2}[V, W] .$$

In particular, the Lie group exponential map coincides with the pseudo-Riemannian exponential map.

Proof The first part of the statement follows from Eqs. (15.1.21) and (15.1.22), since for left-invariant vector fields $D^l_V W = 0$. The second part is a direct consequence, since 1-parameter groups $\gamma : I \to G$ integrate left-invariant vector fields, and therefore $\nabla_{\dot\gamma}\dot\gamma = 0$. □

15.1.2.4 Lorentzian Cross-Product

Before a discussion on geodesics in the PSL(2, $\mathbb{R}$)-model, which will rely on the Lie group generalities of the previous section, we discuss here some particular features of the Lie group $G = \mathrm{PSL}(2, \mathbb{R})$. Namely, we have a natural Lorentzian cross product, that is a $T\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$-valued 2-form $(V, W) \mapsto V \boxtimes W$, which is defined by the equality

$$\langle V \boxtimes W, U \rangle = \Omega(V, W, U) , \qquad (15.1.23)$$

where $\langle \cdot, \cdot \rangle$ is the Anti-de Sitter metric and Ω is the associated volume form, namely the unique 3-form taking the value 1 on any positive oriented orthonormal basis. Here we orient PSL(2, $\mathbb{R}$) by declaring that the orthonormal basis

$$V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad W = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad U = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

at the identity is positive. In other words, $V \boxtimes W$ equals $*(X \wedge Y)$, where $* : \Lambda^2 T\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \to T\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is the Hodge star operator defined similarly to the Riemannian case.

At the identity, a very simple equality holds for the Lorentzian cross product and the Lie bracket of $\mathfrak{g}$:

Lemma 15.1.2.4.1 *Given $V, W \in T_{\mathbb{1}}\mathrm{PSL}(2, \mathbb{R})$, $[V, W]_{\mathfrak{g}} = -2V \boxtimes W$.*

Proof We claim that the volume form of the Anti-de Sitter metric equals:

$$\Omega(V, W, U) = -\frac{1}{2}\langle [V, W]_{\mathfrak{g}}, U \rangle . \qquad (15.1.24)$$

The stated equality then follows from Eq. (15.1.23). To see the claim, first let us observe that the expression in (15.1.24) is an alternating three-form, as a consequence of the skew-symmetry of the Lie bracket and of the (infinitesimal version of) Ad-invariance of the Anti-de Sitter metric, namely:

$$\langle [V, W]_{\mathfrak{g}}, U \rangle = -\langle W, [V, U]_{\mathfrak{g}} \rangle . \qquad (15.1.25)$$

Hence Ω is a multiple of the volume form. To check the multiplicative factor, by left-invariance, it suffices to perform the computation at $T_{\mathbb{1}}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} = \mathfrak{sl}(2, \mathbb{R})$ on the positive orthonormal basis

$$V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad W = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad U = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

for which V, W are spacelike and U is timelike. The equality follows since $[V, W]_{\mathfrak{g}} = 2U$. □

Lemma 15.1.2.4.1 permits to rewrite the expression for the Levi-Civita connection of left-invariant vector fields, from Lemma 15.1.2.3.1, simply as $\nabla_V W = -V \boxtimes W$ and, together with Eqs. (15.1.22) and (15.1.21), to obtain the following general expression for the Levi-Civita connection.

$$\nabla_V W = D^l_V W - V \boxtimes W = D^r_V W + V \boxtimes W \, . \tag{15.1.26}$$

Remark 15.1.2.4.2 Using the set-up of this section, one easily gets another computation of the curvature of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, different from that given in Sect. 15.1.1.5. Fix $V, W, U \in \mathfrak{g} = T_{\mathbb{1}}\mathrm{PSL}(2,\mathbb{R})$, and denote by $\widetilde{V}, \widetilde{W}, \widetilde{U}$ the left invariant extensions of V, W, U. From Lemma 15.1.2.3.1 and the Jacobi identity, one gets the following expression for the Riemann tensor:

$$\begin{aligned} R(V,W)U &= \left(\nabla_{\widetilde{V}}\nabla_{\widetilde{W}}\widetilde{U} - \nabla_{\widetilde{W}}\nabla_{\widetilde{V}}\widetilde{U} - \nabla_{[\widetilde{V},\widetilde{W}]}\widetilde{U}\right)(\mathbb{1}) \\ &= \left(\frac{1}{4}[\widetilde{V},[\widetilde{W},\widetilde{U}]] - \frac{1}{4}[\widetilde{W},[\widetilde{V},\widetilde{U}]] - \frac{1}{2}[[\widetilde{V},\widetilde{W}],\widetilde{U}]\right)(\mathbb{1}) \\ &= \frac{1}{4}[V,[W,U]_{\mathfrak{g}}]_{\mathfrak{g}} - \frac{1}{4}[W,[V,U]_{\mathfrak{g}}]_{\mathfrak{g}} - \frac{1}{2}[[V,W]_{\mathfrak{g}},U]_{\mathfrak{g}} = \frac{1}{4}[U,[V,W]_{\mathfrak{g}}]_{\mathfrak{g}} \, . \end{aligned}$$

Hence from Lemma 15.1.2.4.1 and Eq. (15.1.25):

$$\langle R(V,W)W,V\rangle = \frac{1}{4}\langle [W,[V,W]_{\mathfrak{g}}]_{\mathfrak{g}},V\rangle = \frac{1}{4}\langle [V,W]_{\mathfrak{g}},[V,W]_{\mathfrak{g}}\rangle = \langle V\boxtimes W, V\boxtimes W\rangle = -1 \, ,$$

for V, W orthonormal spacelike vectors, hence spanning a spacelike plane. An analogous computation holds for timelike planes, thus showing that the sectional curvature is identically -1.

15.1.2.5 Geodesics in PSL(2, $\mathbb{R}$)

In this section we will describe the geodesics of the PSL(2, $\mathbb{R}$)-model, applying its Lie group structure.

Exponential Map Let us start by understanding the geodesics through the identity. Recalling Remark 15.1.2.1.1, the Lie algebra of PSL(2, $\mathbb{R}$) is isometrically identified to a copy of Minkowski space, where under such an isometry the stabilizer of a point (namely PSL(2, $\mathbb{R}$) acting by means of the adjoint action) corresponds to the group of linear isometries of Minkowski space. In short, this means that we shall distinguish geodesics by their type (timelike, spacelike, lightlike) and those will be equivalent under this action. Moreover, by Lemma 15.1.2.3.1 it suffices to understand the one-parameter groups for the Lie group structure of PSL(2, $\mathbb{R}$). We immediately get the following:

- Timelike geodesics are, up to conjugacy, of the form

$$\begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}$$

  namely, under the identification of $\mathrm{PSL}(2, \mathbb{R})$ with $\mathrm{Isom}(\mathbb{H}^2)$, they are elliptic one-parameter groups fixing a point in $\mathbb{H}^2$. In this example, the tangent vector is the matrix

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} .$$

  These are in fact closed geodesics, parameterized by arclength, of total length π.
- Spacelike geodesics are, again up to conjugacy:

$$\begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix}$$

  with initial velocity

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .$$

  In terms of hyperbolic geometry, these are hyperbolic one-parameter groups, fixing two points in the boundary of $\mathbb{H}^2$ (in this case, ± 1).
- Finally, lightlike geodesics are the parabolic one-parameter groups conjugate to

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} ,$$

  whose initial vector has indeed zero length.

A Totally Geodesic Spacelike Plane Using the above description of timelike geodesics through $\mathbb{1}$, we can also interpret the duality of Sect. 15.1.1.9 in terms of the structure of $\mathrm{PSL}(2, \mathbb{R})$. Recalling that the dual plane of a point A is the set of antipodal points along timelike geodesics through A, one sees that the dual plane of $\mathbb{1}$ consists of elliptic isometries of $\mathbb{H}^2$ which rotate by an angle π. Equivalently, this is the set of (projective classes) of traceless matrices, that is (by the Cayley–Hamilton theorem)

$$P_{\mathbb{1}} = \{[J] \in \mathrm{PSL}(2, \mathbb{R}) \mid J^2 = -\mathbb{1}\} .$$

In other words, $P_{\mathbb{1}}$ is identified with the space of *linear almost-complex structures* on $\mathbb{R}^2$, up to sign reversing. The boundary at infinity of $P_{\mathbb{1}}$ is made of traceless matrices of rank 1, that is, the projectivization of the set of nilpotent 2×2 matrices.

Recall that the stabilizer of $\mathbb{1}$ is the diagonal subgroup of $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$, and it also acts on the dual plane $P_{\mathbb{1}}$ by conjugation. The following statement is then straightforward:

Lemma 15.1.2.5.1 *The map from $\mathbb{H}^2$ to $P_{\mathbb{1}}$, sending $p \in \mathbb{H}^2$ to the elliptic order-two element in* $\mathrm{PSL}(2, \mathbb{R})$ *fixing p, is a* $\mathrm{PSL}(2, \mathbb{R})$*-equivariant isometry.*

Proof Equivariance with respect to the actions of $\mathrm{PSL}(2, \mathbb{R})$ is easy since, for an element $X \in \mathrm{PSL}(2, \mathbb{R})$, the order-two elliptic element fixing $X \cdot p$ is precisely the X-conjugate of the order-two elliptic element fixing p. Using the equivariance, it follows that the pull-back of the metric of $P_{\mathbb{1}}$ is a constant multiple of the hyperbolic metric of $\mathbb{H}^2$. Since both have curvature -1, they must coincide. □

On the double cover $\mathbb{H}^{2,1}$, which is the $\mathrm{SL}(2, \mathbb{R})$-model, $P_{\mathbb{1}}$ lifts to the two planes $P_{\mathbb{1}}^{\pm}$ dual to the identity. One of them consists of the matrices J such that $J^2 = -\mathbb{1}$, namely the linear almost-complex structures on $\mathbb{R}^2$, which are compatible with the standard orientation of $\mathbb{R}^2$; the other contains the linear almost-complex structures on $\mathbb{R}^2$ compatible with the opposite orientation of $\mathbb{R}^2$.

Timelike Geodesics To get a complete description of timelike geodesics (not only those through the identity) it suffices to let (the identity component of) the isometry group of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, namely $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$ act on $\mathrm{PSL}(2, \mathbb{R})$ by left and right multiplication. In particular an important description of the space of timelike geodesics of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ (which is also the space of timelike geodesics of each finite-index cover of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$) can be obtained, see [9].

Proposition 15.1.2.5.2 *There is a homeomorphism between the space of (unparameterized) timelike geodesics of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ and $\mathbb{H}^2 \times \mathbb{H}^2$. The homeomorphism is equivariant for the action of* $\mathrm{Isom}_0(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}) \cong \mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$.

Proof The homeomorphism is defined as follows. Given $(p, q) \in \mathbb{H}^2 \times \mathbb{H}^2$, we associate to it the subset

$$L_{p,q} = \{X \in \mathrm{PSL}(2, \mathbb{R}) \mid X \cdot q = p\} .$$

By the previous discussion, geodesics through the identity are precisely of the form $L_{p,p}$ for some $p \in \mathbb{H}^2$. It is easy to check that the map $(p, q) \mapsto L_{p,q}$ is equivariant for the natural actions of $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$, namely $(A, B) \cdot L_{p,q} = L_{A \cdot p, B \cdot q}$, which also implies that $L_{p,q}$ is an unparameterized geodesic and that all unparameterized geodesics are of this form, namely the map we defined is surjective. It remains to see the injectivity: if $L_{p,q} = L_{p',q'}$ for $(p, q) \neq (p', q')$ then in particular there exists an isometry of $\mathbb{H}^2$ sending p to q and p' to q'. But such an isometry is necessarily unique since the identity is the only isometry of $\mathbb{H}^2$ fixing two different points. This gives a contradiction and thus concludes the proof. □

Spacelike Geodesics Let us conclude this section by an analysis of spacelike geodesics. Let us consider an oriented geodesic ℓ of $\mathbb{H}^2$. From the discussion at the beginning of this section, the one-parameter group of hyperbolic transformations fixing ℓ *as an oriented geodesic* constitutes a spacelike geodesic through the origin. By an argument very similar to Proposition 15.1.2.5.2, relying on the equivariance of the construction by the actions of $\mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$, one then proves that every spacelike geodesic is of the form:

$$L_{\ell,\jmath} = \{X \in \mathrm{PSL}(2,\mathbb{R}) \mid X \cdot \jmath = \ell \text{ as oriented geodesics}\} ,$$

where ℓ and $\jmath$ denote oriented geodesics of $\mathbb{H}^2$. We remark that every (unparameterized, unoriented) spacelike geodesic can be expressed in the above form in two ways, as one can change the orientation of both ℓ and $\jmath$. Every such choice corresponds to a choice of orientation for the spacelike geodesic. In other words, one can show:

Proposition 15.1.2.5.3 *There is a homeomorphism between the space of (unparameterized) oriented spacelike geodesics of* $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ *and the product of two copies of* $\partial\mathbb{H}^2 \times \partial\mathbb{H}^2 \setminus \Delta$*, the space of oriented geodesics of* $\mathbb{H}^2$*. The homeomorphism is equivariant for the action of* $\mathrm{Isom}_0(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}) \cong \mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$.

However, for our purpose, we will mostly deal with *unoriented* geodesics, hence we will have $L_{\ell,\jmath} = L_{\ell',\jmath'}$ where ℓ' equals ℓ but endowed with the opposite orientation. Given a spacelike geodesic, there is a natural notion of *dual* spacelike geodesic, which is defined using the projective duality between points and planes from Sect. 15.1.1.9:

Definition 15.1.2.5.4 Given a spacelike geodesic $L_{\ell,\jmath}$ in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, the *dual spacelike geodesic* is the intersection of all spacelike planes dual to points of $L_{\ell,\jmath}$.

The construction of the dual geodesic is involutive. Let us now see an explicit example. For the case of the geodesic $L_{\ell,\ell}$ through the origin, which consists of the one-parameter hyperbolic group of $\mathrm{PSL}(2,\mathbb{R})$ translating along ℓ, it can be checked that the dual geodesic consists of all elliptic order-two elements (hence contained in $P_{\mathbb{1}}$, as it is expected from the definition) whose fixed point lies in ℓ. In other words, the dual spacelike geodesic of $L_{\ell,\ell}$ is $L_{\ell,\ell'}$.

We can easily describe the points at infinity in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ of these geodesics. Using Lemma 15.1.2.2.2, if x and y are the endpoints at infinity of ℓ in $\partial\mathbb{H}^2$, then clearly any sequence diverging towards an end of $L_{\ell,\ell} \subset \mathrm{PSL}(2,\mathbb{R})$ maps an interior point towards x, and the sequence of inverses towards y (up to switching x and y). In other words, under the identification of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ with $\mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$ (Sect. 15.1.2.2), the endpoints of $L_{\ell,\ell}$ are (x, y) and (y, x). A similar argument applied to $L_{\ell,\ell'}$, which consists of order-two elliptic elements with fixed point in ℓ, shows that its endpoints are (x, x) and (y, y).

Recalling the descriptions of the left and right rulings of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, we conclude that the endpoints of a spacelike geodesic and its dual are mutually connected by lightlike segments in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. See also Fig. 15.8 in Sect. 15.2.2, where this configuration is studied and applied more deeply.

By naturality of the construction with respect to the action of $\mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$, one has:

Proposition 15.1.2.5.5 *Given a spacelike geodesic $L_{\ell,\jmath}$ of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, its endpoints in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ are (x_1, y_2) and (y_1, x_2), where x_1 and y_1 are the final and initial endpoints of ℓ in $\partial\mathbb{H}^2$, and x_2 and y_2 are the final and initial endpoints of $\jmath$ (where final and initial refers to the orientation of ℓ and $\jmath$). The dual geodesic is $L_{\ell,\jmath'}$ and has endpoints (x_1, x_2) and (y_1, y_2).*

15.2 Part 2: The Seminal Work of Mess

The aim of this part is to describe Mess' work, including the classification of maximal globally hyperbolic spacetimes with compact Cauchy surface and the Gauss map of spacelike surfaces. The material is organized in the following way. Chapter 15.2.1 analyses various properties of causality and convexity in Anti-de Sitter space, which are preliminary to the proof of Mess' classification result. The latter is given in Chap. 15.2.2. In Chap. 15.2.3 we then treat the Gauss map and its first properties, and discuss Mess' proof of the Earthquake Theorem.

15.2.1 Causality and Convexity Properties

Here we will first study achronal setsin the conformal compactification of Anti-de Sitter space, a notion that makes sense in the universal cover $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, and then adapt the notion for subsets of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. Then we introduce the fundamental notions of invisible domain and of domain of dependence, and describe their properties.

15.2.1.1 Achronal and Acausal Sets

Let us begin with the first definitions.

Definition 15.2.1.1.1 A subset X of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \cup \partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ is *achronal* (resp. *acausal*) if no pair of points in X is connected by timelike (resp. causal) lines in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$.

Since acausality and achronality are conformally invariant notions, it will be often convenient to consider the metric $g_{\mathbb{S}^2} - dt^2$ on $\mathbb{D} \times \mathbb{R}$ we introduced in (15.1.14) (for $g_{\mathbb{S}^2}$ the hemispherical metric on the disc), which is conformal to the Poincaré model of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$.

Lemma 15.2.1.1.2 *A subset X of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \cup \partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ is achronal (resp. acausal) if and only if it is the graph of a function $\mathsf{f} : D \to \mathbb{R}$ that is 1-Lipschitz (resp. strictly 1-Lipschitz) with respect to the distance induced by the hemispherical metric $g_{\mathbb{S}^2}$.*

Clearly here $\mathsf{D} = \pi_{\mathbb{D}}(X)$ denotes the projection of X to the $\mathbb{D}$ factor.

Proof Assume that X is an achronal subset. Since vertical lines in the Poincaré model are timelike, the restriction of the projection $\pi_{\mathbb{D}} : \mathbb{D} \times \mathbb{R} \to \mathbb{D}$ to X is injective. So X can be regarded as the graph of a function $\mathsf{f} : \mathsf{D} \to \mathbb{R}$. Imposing that $(\mathsf{x}, \mathsf{f}(\mathsf{x}))$ and $(\mathsf{y}, \mathsf{f}(\mathsf{y}))$ are not related by a timelike curve we deduce that

$$|\mathsf{f}(\mathsf{x}) - \mathsf{f}(\mathsf{y})| \le d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{y}) , \tag{15.2.1}$$

where d_{S^2} is the hemispherical distance (see also Sect. 15.1.1.8). The same argument shows that conversely the graph of a 1-Lipschitz function defined on some subset of $\mathbb{D}$ is achronal.

Moreover, two points (x, t) and (y, s) are on the same lightlike geodesic if and only if $|\mathsf{t}-\mathsf{s}| = d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{y})$. Hence X is acausal if and only if the inequality in (15.2.1) is strict. □

Observe that a 1-Lipschitz function on a region $\mathsf{D} \subset \mathbb{D}$ extends uniquely to the boundary of D. As a simple consequence of the previous lemma, we thus have:

Lemma 15.2.1.1.3 *An achronal subset X in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is properly embedded if and only if it is a global graph over $\mathbb{D}$, and in this case it extends uniquely to the global graph of a 1-Lipschitz function over $\mathbb{D} \cup \partial\mathbb{D}$.*

In light of Lemma 15.2.1.1.3, in the following we will refer to an *achronal surface* as an achronal subset X in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ which is the graph of a 1-Lipschitz function defined on a domain in $\mathbb{D}$.

Before studying more detailed properties, we shall remark that achronality and acausality are global conditions. Let us first recall the definition of spacelike surface:

Definition 15.2.1.1.4 Given a surface S and a Lorentzian manifold (M, g), a C^1 immersion $\sigma : S \to M$ is *spacelike* if the pull-back metric $\sigma^* g$ is a Riemannian metric. If σ is an embedding, we refer to its image as a *spacelike surface*.

A spacelike surface S is locally acausal (in the sense that any point admits a neighborhood in S that is acausal), but there are examples of spacelike surfaces which are not achronal (hence a fortiori not acausal), a fact which highlights the global character of Definition 15.2.1.1.1. On the other hand, we have this global result.

Lemma 15.2.1.1.5 *Any properly embedded spacelike surface in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ is acausal.*

Proof By Lemma 15.2.1.1.3, any properly embedded spacelike surface S in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ disconnects the space in two regions U and V, whose common boundary is S, and we can assume that the outward pointing normal from U (resp. V) is past-directed (resp. future directed). It then turns out that any future oriented causal path that meets S passes from V towards U. This implies that any causal path meets S at most once. □

Recall from Theorem 15.1.1.8.2 that unparameterized lightlike geodesics only depend on the conformal class of the Lorentzian metric, hence in the following we will simply refer to lightlike geodesics in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, although we very often use the conformal metric (15.1.14).

Lemma 15.2.1.1.6 *Let S be a properly embedded achronal surface of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \cup \partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ and assume that a lightlike geodesic segment γ joins two points of S. Then γ is entirely contained in S.*

Proof Recalling Lemma 15.2.1.1.3, let $\mathsf{f}^S : \overline{\mathbb{D}} \to \mathbb{R}$ be the function defining S, which is 1-Lipschitz with respect to the hemispherical metric. Now if γ joins $(\mathsf{x}, \mathsf{f}^S(\mathsf{x}))$ to $(\mathsf{y}, \mathsf{f}^S(\mathsf{y}))$, then (up to switching the role of x and y) $\mathsf{f}^S(\mathsf{y}) = \mathsf{f}^S(\mathsf{x}) + d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{y})$. Moreover γ consists of points of the form $(\mathsf{z}, \mathsf{f}^S(\mathsf{x}) + d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{z}))$, for z lying on the $g_{\mathbb{S}^2}$-geodesic segment joining x to y. For such a point z on the geodesic segment joining x to y, by achronality of S we have:

$$\mathsf{f}^S(\mathsf{z}) - \mathsf{f}^S(\mathsf{x}) \leq d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{z}) \qquad \text{and} \qquad \mathsf{f}^S(\mathsf{y}) - \mathsf{f}^S(\mathsf{z}) \leq d_{\mathbb{S}^2}(\mathsf{z}, \mathsf{y}) = d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{y}) - d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{z}) \,.$$

But the second inequality implies that $\mathsf{f}^S(\mathsf{z}) \geq \mathsf{f}^S(\mathsf{x}) + d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{z})$ so we conclude that $\mathsf{f}^S(\mathsf{z}) = \mathsf{f}^S(\mathsf{x}) + d_{\mathbb{S}^2}(\mathsf{z}, \mathsf{x})$, proving that γ is contained in S. □

Given a function $\mathsf{f} : \overline{\mathbb{D}} \to \mathbb{R}$, we define its oscillation as

$$\mathrm{osc}(\mathsf{f}) := \max_{\mathsf{y} \in \overline{\mathbb{D}}} \mathsf{f}(\mathsf{y}) - \min_{\mathsf{y} \in \overline{\mathbb{D}}} \mathsf{f}(\mathsf{y}) \,.$$

It is important to stress that this quantity is not invariant under the isometry group of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$.

Lemma 15.2.1.1.7 *Let S be a properly embedded achronal surface, defined as the graph of $f^S : \overline{\mathbb{D}} \to \mathbb{R}$. Then $\mathrm{osc}(f^S) \leq \pi$. Moreover $\mathrm{osc}(f^S) = \pi$ if and only if S is a lightlike plane.*

Proof As f^S is 1-Lipschitz for the hemispherical metric, and the diameter of $\mathbb{D}$ for $g_{\mathbb{S}^2}$ is π we easily see that $\mathrm{osc}(\mathsf{f}^S)$ is bounded by π. Moreover if the value π is attained it follows that there are two antipodal points $\mathsf{y}, \mathsf{y}' \in \partial\mathbb{D}$ such that $\mathsf{f}^S(\mathsf{y}') = \mathsf{f}^S(\mathsf{y}) + \pi$. Recall from Sect. 15.1.1.8 (see also Fig. 15.4) that the lightlike plane with past and future points $(\mathsf{y}, \mathsf{f}^S(\mathsf{y}))$ and $(\mathsf{y}', \mathsf{f}^S(\mathsf{y}) + \pi)$ is

$$P = \{(\mathsf{x}, \mathsf{t}) \mid \mathsf{t} = \mathsf{f}^S(\mathsf{y}) + d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{y})\}$$

and is foliated by lightlike geodesics joining $(\mathsf{y}, \mathsf{f}^S(\mathsf{y}))$ to $(\mathsf{y}', \mathsf{f}^S(\mathsf{y}) + \pi)$. By Lemma 15.2.1.1.6, P is included in S. Since both P and S are global graphs over $\overline{\mathbb{D}}$, $S = P$. □

15.2.1.2 Invisible Domains

The first part of this subsection will be devoted to the definition and first properties of invisible domains, which was first given in [6], and in the last part we will focus on the case that X is a subset of $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$.

Definition 15.2.1.2.1 Given an achronal domain X in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \cup \partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, the *invisible domain* of X is the subset $\Omega(X) \subset \widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ of points which are connected to X by no causal path.

Recall that by McShane's Theorem ([87]) any 1-Lipschitz function on a subset of a metric space admits a 1-Lipschitz extension everywhere. Hence any achronal set X, which by Lemma 15.2.1.1.2 is the graph of a 1-Lipschitz function $\mathsf{f}^X : \mathsf{D} \to \mathbb{R}$ for $\mathsf{D} = \pi_{\mathbb{D}}(X)$, is a subset of a properly embedded achronal surface.

Here we introduce two particular extensions $\mathsf{f}^X_\pm : \mathbb{D} \cup \partial\mathbb{D}$, to which we sometimes refer as the *extremal* extensions:

$$\mathsf{f}^X_-(\mathsf{y}) = \sup\{\mathsf{f}^X(\mathsf{x}) - d_{\mathbb{S}^2}(\mathsf{x},\mathsf{y}) \mid \mathsf{x} \in \pi_{\mathbb{D}}(X)\} \qquad \mathsf{f}^X_+(\mathsf{y}) = \inf\{\mathsf{f}^X(\mathsf{x}) + d_{\mathbb{S}^2}(\mathsf{x},\mathsf{y}) \mid \mathsf{x} \in \pi_{\mathbb{D}}(X)\}\,.$$

Clearly $\mathsf{f}^X_\pm$ coincide with f^X on X and are 1-Lipschitz.

Lemma 15.2.1.2.2 *Let X be any closed achronal subset X of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \cup \partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ and let $S_\pm(X)$ be the graphs of the extremal extensions $\mathsf{f}^X_\pm$.*

(1) The properly embedded surfaces $S_-(X)$ and $S_+(X)$ are achronal with $S_-(X) \subset \overline{\mathrm{I}^-(S_+(X))}$, and $\Omega(X) = \mathrm{I}^+(S_-(X)) \cap \mathrm{I}^-(S_+(X))$.
(2) Every achronal subset containing X is contained in $S_-(X) \cup \Omega(X) \cup S_+(X)$.
(3) Every point of $S_\pm(X)$ is connected to X by at least one lightlike geodesic segment, which is entirely contained in $S_\pm(X)$. Finally, $S_+(X) \cap S_-(X)$ is the union of X and all lightlike geodesic segments joining points of X.

Proof Let us first show that $S_-(X) \subset \overline{\mathrm{I}^-(S_+(X))}$. Given a point (y,t), $\mathsf{t} \le \mathsf{f}^X_+(\mathsf{y})$ if and only if $\mathsf{t} \le \mathsf{f}^X(\mathsf{x}) + d_{S^2}(\mathsf{x},\mathsf{y})$ for every $\mathsf{x} \in \pi_{\mathbb{D}}(X)$, that is, if and only if (y,t) lies outside $\mathrm{I}^+(X)$. Similarly (y,t) lies outside $\mathrm{I}^-(X)$ if and only if $\mathsf{t} \ge \mathsf{f}^X_-(\mathsf{y})$. By achronality, $S_+(X)$ does not meet the past of X, so we deduce that $\mathsf{f}^X_+(\mathsf{y}) \ge \mathsf{f}^X_-(\mathsf{y})$ for all $\mathsf{y} \in \overline{\mathbb{D}}$, that is, $S_-(X)$ is contained in $\overline{\mathrm{I}^-(S_+(X))}$.

As a similar observation, given a point (y,t), $\{(\mathsf{y},\mathsf{t})\} \cup X$ is achronal if and only if $\mathsf{f}^X_-(\mathsf{y}) \le \mathsf{t} \le \mathsf{f}^X_+(\mathsf{y})$. Moreover (y,t) is connected to X by no causal curve if and only if $\mathsf{f}^X_-(\mathsf{y}) < \mathsf{t} < \mathsf{f}^X_+(\mathsf{y})$. This shows that

$$\Omega(X) = \{(\mathsf{y},\mathsf{t}) \mid \mathsf{f}^X_-(\mathsf{y}) < \mathsf{t} < \mathsf{f}^X_+(\mathsf{y})\}\,,$$

and also the second item, by applying the previous observation to any point of an achronal set containing X which is not in X itself.

To prove the third item, fix a point in $(\mathsf{y},\mathsf{t}) \in S_+(X)$. As we are assuming that X is closed in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \cup \partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, the fact that f^X is 1-Lipschitz implies that

$\pi_{\mathbb{D}}(X)$ is closed in $\mathbb{D} \cup \partial\mathbb{D}$, so it is compact. In particular there exists $\mathsf{x} \in \partial\mathbb{D}$ such that $\mathsf{t} = \mathsf{f}_+^X(\mathsf{y}) = \mathsf{f}^X(\mathsf{x}) + d_{S^2}(\mathsf{x}, \mathsf{y})$. Thus (y, t) is connected to $(\mathsf{x}, \mathsf{f}^X(\mathsf{x}))$ by a lightlike geodesic segment. By Lemma 15.2.1.1.6, this geodesic segment is entirely contained in $S_+(X)$. Clearly the proof for $S_-(X)$ is analogous.

It remains to compute $S_-(X) \cap S_+(X)$. For this purpose, notice that if two points of X are connected by a lightlike geodesic segment γ, applying Lemma 15.2.1.1.6 we deduce that $\gamma \subset S_-(X) \cap S_+(X)$. Conversely let $(\mathsf{y}, \mathsf{t}) \in S_-(X) \cap S_+(X)$ so that $\mathsf{f}_-^X(\mathsf{y}) = \mathsf{f}_+^X(\mathsf{y})$. There exist x and x' in $\pi_{\mathbb{D}}(X)$ such that

$$\mathsf{f}_+^X(\mathsf{y}) = \mathsf{f}^X(\mathsf{x}) + d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{y}) \qquad \text{and} \qquad \mathsf{f}_-^X(\mathsf{y}) = \mathsf{f}^X(\mathsf{x}') - d_{\mathbb{S}^2}(\mathsf{x}', \mathsf{y})$$

Using that $\mathsf{f}_-^X(\mathsf{y}) = \mathsf{f}_+^X(\mathsf{y})$, the triangle inequality and the fact that f^X is 1-Lipschitz we deduce that

$$\mathsf{f}^X(\mathsf{x}) - \mathsf{f}^X(\mathsf{x}') = d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{x}') = d_{\mathbb{S}^2}(\mathsf{x}, \mathsf{y}) + d_{\mathbb{S}^2}(\mathsf{y}, \mathsf{x}') \,. \tag{15.2.2}$$

Hence the points $(\mathsf{x}, \mathsf{f}^X(\mathsf{x}))$ and $(\mathsf{x}', \mathsf{f}^X(\mathsf{x}'))$ are joined by a lightlike segment. If x, x' are not antipodal points on $\partial\mathbb{D}$ there, there is a unique hemispherical geodesic η in $\overline{\mathbb{D}}$ joining x to x', which must pass through y by (15.2.2), and which we may assume parameterized by arclength. In this case the geodesic segment joining $(\mathsf{x}, \mathsf{f}^X(\mathsf{x}))$ to $(\mathsf{x}', \mathsf{f}^X(\mathsf{x}'))$ takes the form $t \mapsto (\eta(t), \mathsf{f}^X(\mathsf{x}') + t)$, so it passes through $(\mathsf{y}, \mathsf{f}_+^X(\mathsf{y})) = (\mathsf{y}, \mathsf{f}_-^X(\mathsf{y}))$.

If x and x' are antipodal, then there are infinitely many geodesics joining x to x', and we can pick one going through y. Then the same argument as above applies. □

Remark 15.2.1.2.3 Given a point (y, t), the set of points (x, s) satisfying $|\mathsf{s} - \mathsf{t}| < d_{S^2}(\mathsf{x}, \mathsf{y})$ coincides with the region of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ which is connected to (y, t) by a spacelike geodesic for the Anti-de Sitter metric. It coincides also with the region of points connected to (y, t) by a spacelike geodesic for the conformal metric $g_{\mathbb{S}^2} - dt^2$, although in general spacelike geodesics for the two metrics do not coincide.

Now, since $\mathsf{f}_-^X(\mathsf{y}) \leq \mathsf{t} \leq \mathsf{f}_+^X(\mathsf{y})$ is equivalent to the condition that $|\mathsf{s} - \mathsf{t}| \leq d_{S^2}(\mathsf{x}, \mathsf{y})$ for all $(\mathsf{x}, \mathsf{s}) \in X$, the region

$$S_+(X) \cup \Omega(X) \cup S_-(X) = \{(\mathsf{y}, \mathsf{t}) \mid \mathsf{f}_-^X(\mathsf{y}) \leq \mathsf{t} \leq \mathsf{f}_+^X(\mathsf{y})\}$$

consists of points that are connected to any point of X by spacelike or lightlike geodesics. Moreover $\Omega(X)$ consists of points connected to any point of X by a spacelike geodesic.

We remark that in general $\Omega(X)$ could be empty. For instance if X is a global graph then $S_+(X) = S_-(X) = X$ and $\Omega(X)$ is empty.

Remark 15.2.1.2.4 Since any point of $S_\pm(X)$ is connected to X by a lightlike geodesic, it follows by Lemma 15.2.1.1.6 that the intersection of any properly embedded achronal surface containing X with $S_\pm(X)$ is a union of lightlike geodesic segments with an endpoint in X. In particular any properly embedded acausal surface containing X is contained in the region $\Omega(X)$.

15.2.1.3 Achronal Meridians in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$

We will be mainly interested in the invisible domains of achronal meridians Λ in the boundary of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, that are graphs of 1-Lipschitz functions $\mathsf{f} : \partial\mathbb{D} \to \mathbb{R}$. Let us study more closely this case.

Lemma 15.2.1.3.1 *Let Λ be an achronal meridian in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$. Then either Λ is the boundary of a lightlike plane, or $S_+(\Lambda) \cap S_-(\Lambda) = \Lambda$. In the latter case there is an achronal properly embedded surface in $\Omega(\Lambda)$ whose boundary in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ is Λ.*

Proof Let $\mathsf{f} : \partial\mathbb{D} \to \mathbb{R}$ be the function whose graph is Λ. Recall from Lemma 15.2.1.1.7 that $\mathrm{osc}(\mathsf{f}) \le \pi$. If there are points $\mathsf{x}_0, \mathsf{x}_0'$ such that $\mathsf{f}(\mathsf{x}_0') = \mathsf{f}(\mathsf{x}_0)+\pi$, then combining Lemmas 15.2.1.1.7 and 15.2.1.2.2 we deduce that Λ is the boundary of a lightlike plane, and this lightlike plane coincides with $S_+(\Lambda)\cap S_-(\Lambda)$.

Assume now that the maximal oscillation of f is smaller than π, and let us show that $S_+(\Lambda) \cap S_-(\Lambda) = \Lambda$. By the assumption, if a lightlike geodesic connects $(\mathsf{x}_0, \mathsf{f}(\mathsf{x}_0))$ to $(\mathsf{x}_0', \mathsf{f}(\mathsf{x}_0'))$, then x_0 and x_0' are not antipodal. But then $\mathsf{x}_0, \mathsf{x}_0'$ are connected by a unique length-minimizing geodesic in $\overline{\mathbb{D}}$ for the hemispherical metric, which lies in $\partial\mathbb{D}$. So the lightlike line connecting $(\mathsf{x}_0, \mathsf{f}(\mathsf{x}_0))$ to $(\mathsf{x}_0', \mathsf{f}(\mathsf{x}_0'))$ is contained in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$. By Lemma 15.2.1.2.2 we conclude that $S_-(\Lambda)$ and $S_+(\Lambda)$ do not meet in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ and therefore $S_+(\Lambda) \cap S_-(\Lambda) = \Lambda$.

Finally, in this latter case the function $F = (\mathsf{f}_-^\Lambda + \mathsf{f}_+^\Lambda)/2$ is 1-Lipschitz and defines an achronal properly embedded surface contained in $\Omega(\Lambda)$, whose boundary is Λ.

□

We remark that in fact for any achronal meridian there is a spacelike surface whose boundary at infinity is Λ, see Remark 15.2.1.4.7 below.

Recall from Sect. 15.1.1.9 that, given a point x in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, the Dirichlet domain of x is the region R_x containing x and bounded by two spacelike planes "dual" to x. Namely the planes, which by a small abuse we denote by P_x^+ and P_x^-, consisting of points at timelike distance $\pi/2$ in the future (resp. past) along timelike geodesics with initial point x.

Proposition 15.2.1.3.2 *Let Λ be an achronal meridian in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ different from the boundary of a lightlike plane. Then*

(1) A point $x \in \widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ lies in $\Omega(\Lambda)$ if and only if Λ is contained in the interior of the Dirichlet region R_x.

(2) For any $z \in \Lambda$, let $L_-(z)$ and $L_+(z)$ be the two lightlike planes such that z is the past vertex of $L_+(z)$ and the future vertex of $L_-(z)$. Then

$$\Omega(\Lambda) = \bigcap_{z\in\Lambda} \mathrm{I}^+(L_-(z)) \cap \mathrm{I}^-(L_+(z)) .$$

(3) The length of the intersection of $\Omega(\Lambda)$ with any timelike geodesic of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ is at most π. Moreover, there exists a timelike geodesic whose intersection with $\Omega(\Lambda)$ has length π if and only if Λ is the boundary at infinity of a spacelike plane.

Proof By Remark 15.2.1.2.3 a point x lies in $\Omega(\Lambda)$ if and only if it is connected to any point of Λ by a spacelike geodesic. The region of points connected to x by a spacelike geodesic has boundary the lightcone from x, whose intersection with $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ coincides with $P_x^{\pm} \cap \partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$. This proves the first statement.

Similarly the region bounded by $L_+(z)$ and $L_-(z)$ contains exactly points connected to z by a spacelike geodesic. Using the characterization of $\Omega(\Lambda)$ as above, we conclude the proof of the second statement.

For the third statement, if a timelike geodesic γ meets $\Omega(\Lambda)$ at a point x, then, $\Omega(\Lambda) \subset R_x$, so that the length of $\gamma \cap \Omega(\Lambda)$ is smaller than the length of $\gamma \cap R_x$. But the latter is π. Assume there exists a geodesic γ such that the length of $\gamma \cap \Omega(\Lambda)$ is π. Up to applying an isometry of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ we may assume that γ is vertical in the Poincaré model of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ and the mid-point of $\gamma \cap \Omega(\Lambda)$ is $(0, 0)$. Thus $(0, -\pi/2)$ and $(0, \pi/2)$ lie on $S_-(\Lambda)$ and $S_+(\Lambda)$ respectively. By Remark 15.2.1.2.3 points of Λ are connected to $(0, -\pi/2)$ by a spacelike or lightlike geodesic, hence $\mathsf{s} \leq 0$ for all $(\xi, \mathsf{s}) \in \Lambda$. Analogously using the point $(0, \pi/2)$ we deduce that $\mathsf{s} \geq 0$ for all $(\xi, \mathsf{s}) \in \Lambda$, so that $\Lambda = \partial\mathbb{D} \times \{0\}$. □

With similar arguments, we obtain that the invisible domain of an achronal meridian which is not the boundary of a lightlike plane is always contained in a Dirichlet region.

Proposition 15.2.1.3.3 *Given an achronal meridian Λ in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ different from the boundary of a lightlike plane, the invisible domain $\Omega(\Lambda)$ is contained in a Dirichlet region. Moreover the closure of $\Omega(\Lambda)$ is contained in a Dirichlet region unless Λ is the boundary of a spacelike plane.*

Proof In fact let us set $\mathsf{a}_+ = \sup \mathsf{f}_+^{\Lambda}$ and $\mathsf{a}_- = \inf \mathsf{f}_-^{\Lambda}$, and consider the planes

$$Q_{\mathsf{a}_+} = \{(\mathsf{x}, \mathsf{t}) \mid \mathsf{t} = \mathsf{a}_+\} \qquad \text{and} \qquad Q_{\mathsf{a}_-} = \{(\mathsf{x}, \mathsf{t}) \mid \mathsf{t} = \mathsf{a}_-\}$$

in the Poincaré model. Since clearly $\Omega(\Lambda)$ lies in the open region bounded by those planes, it is sufficient to show that $\mathsf{a}_+ - \mathsf{a}_- \leq \pi$. Assume by contradiction that $\mathsf{a}_+ - \mathsf{a}_- > \pi$. Notice that P_{a_+} meets $S_+(\Lambda)$ at some point $p_+ = (\mathsf{x}_+, \mathsf{a}_+)$, and P_{a_-} meets $S_-(\Lambda)$ at some point $p_- = (\mathsf{x}_-, \mathsf{a}_-)$, where x_+ and x_- are points on $\overline{\mathbb{D}}$. For $\epsilon = (\mathsf{a}_+ - \mathsf{a}_- - \pi)/2$ we can find x'_+ and x'_- in $\mathbb{D}$ such that $p'_+ = (\mathsf{x}'_+, \mathsf{a}_+ - \epsilon)$ and $p'_- = (\mathsf{x}'_-, \mathsf{a}_- + \epsilon)$ lie in $\Omega(\Lambda)$ (clearly if $\mathsf{x}_\pm$ lies in $\mathbb{D}$ we can take $\mathsf{x}'_\pm = \mathsf{x}_\pm$). As $(\mathsf{a}_+ - \epsilon) - (\mathsf{a}_- - \epsilon) = \pi$, the geodesic segment γ joining p'_+ and p'_- is timelike of length π. Its end-points are in $\mathrm{I}^+(S_-(\Lambda)) \cap \mathrm{I}^-(S_+(\Lambda))$, so γ is entirely contained in $\Omega(\Lambda)$. As end-points of γ are contained in $\Omega(\Lambda)$, γ can be extended within $\Omega(\Lambda)$ but this contradicts the third point of Proposition 15.2.1.3.2.

The third point of Proposition 15.2.1.3.2 then shows that if $\mathsf{a}_+ - \mathsf{a}_- = \pi$ then Λ is the boundary of a spacelike plane. Hence apart from this case, one has $\mathsf{a}_+ - \mathsf{a}_- < \pi$, so the closure of $\Omega(\Lambda)$ is contained in a Dirichlet region. □

Remark 15.2.1.3.4 When Λ is the boundary of a spacelike plane P, then there are two points x_- and x_+, such that $P = P_{x_-}^{+} = P_{x_+}^{-}$. The previous arguments show

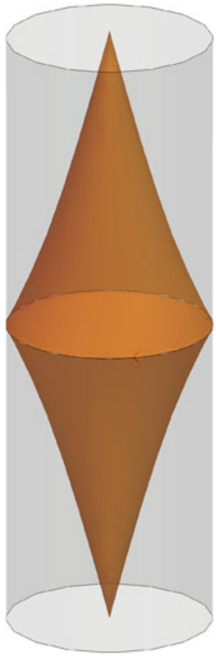


Fig. 15.7 The invisible domain of the boundary of a spacelike plane in the Poincaré model for $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$

that in this case $\Omega(\Lambda)$ is the union of all timelike lines joining x_- to x_+. In this case $S_-(\Lambda)$ is the union of future directed lightlike geodesic rays emanating from x_-, whereas $S_+(\Lambda)$ is the union of future directed lightlike geodesic rays ending at x_+. See Fig. 15.7.

15.2.1.4 Domains of Dependence

We shall now introduce the notion of Cauchy surface and domains of dependence, which is general in Lorentzian geometry, and develop some properties in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$.

Definition 15.2.1.4.1 Given an achronal subset X in a Lorentzian manifold (M, g), the *domain of dependence* of X is the set

$$\mathcal{D}(X) = \{p \in M \mid \text{every inextensible causal curve through } p \text{ meets } X\} .$$

We say that X is a *Cauchy surface* of M if $\mathcal{D}(X) = M$. A spacetime M is said *globally hyperbolic* if it admits a Cauchy surface.

Globally hyperbolic spacetimes have some strong geometric properties, which we summarize in the following theorem. We refer to [19, 23, 24, 68] for an extensive treatment.

Theorem 15.2.1.4.2 *Let M be a globally hyperbolic spacetime. Then*

(1) Any two Cauchy surfaces in M are diffeomorphic.
(2) There exists a submersion $\tau : M \to \mathbb{R}$ whose fibers are Cauchy surfaces.
(3) M is diffeomorphic to $\Sigma \times \mathbb{R}$, where Σ is any Cauchy surface in M.

Remark 15.2.1.4.3 The spacetime $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ is not globally hyperbolic. In fact if X is achronal, it is contained in the graph of a 1-Lipschitz function $\mathsf{f} : (\mathbb{D} \cup \partial\mathbb{D}, g_{\mathbb{S}^2}) \to \mathbb{R}$. If $\mathsf{t}_0 > \sup \mathsf{f}$ and $\xi \in \partial\mathbb{D}$, then any lightlike ray with past end-point (ξ, t_0) does not intersect X.

Remark 15.2.1.4.4 By the usual invariance of causality notions under conformal change of metrics, causal paths in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ are the graphs of 1-Lipschitz functions from (intervals in) $\mathbb{R}$ to $\mathbb{D}$ with respect to the hemispherical metric in the image. Hence an inextensible causal curve in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ is either the graph of a global 1-Lipschitz function from $\mathbb{R}$, or it is defined on a proper interval and has endpoint(s) in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$.

Lemma 15.2.1.4.5 *Given an achronal meridian Λ in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, any Cauchy surface in $\Omega(\Lambda)$ is properly embedded with boundary at infinity Λ.*

Proof Let S be a Cauchy surface in $\Omega(\Lambda)$. For every $x \in \mathbb{D}$, the vertical line through x in the Poincaré model meets $\Omega(\Lambda)$, and its intersection with $\Omega(\Lambda)$ must meet S by definition of Cauchy surface. This shows that S is a graph over $\mathbb{D}$, proving that S is properly embedded, and clearly $\partial S = \Lambda$. □

Proposition 15.2.1.4.6 *Let Λ be an achronal meridian in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ different from the boundary of a lightlike plane. Let S be a properly embedded achronal surface in $\Omega(\Lambda)$. Then $\mathcal{D}(S) = \Omega(\Lambda)$. In particular $\Omega(\Lambda)$ is a globally hyperbolic spacetime.*

Proof Let x be any point in $\Omega(\Lambda)$ and take any inextensible causal path through x. A priori its future endpoint might be either in $S_+(\Lambda)$ or in Λ, but by definition of $\Omega(\Lambda)$, x cannot be connected by a causal path to Λ, hence the latter case is excluded. The same argument applies to show that the past endpoint is in $S_-(\Lambda)$. Since the inextendible causal path meets both $S_+(\Lambda)$ and $S_-(\Lambda)$, it must meet S by Lemma 15.2.1.4.5, hence $x \in \mathcal{D}(S)$.

Conversely, if x is not in $\Omega(\Lambda)$, then one can find a causal path joining x to Λ, which is necessarily inextensible. Hence x is not in $\mathcal{D}(S)$. This concludes the proof. □

Remark 15.2.1.4.7 It follows from Theorem 15.2.1.4.2 and Proposition 15.2.1.4.6 that Λ is the boundary of a spacelike surface in $\Omega(\Lambda)$, namely a Cauchy surface for $\Omega(\Lambda)$. By Lemma 15.2.1.4.5, this surface is properly embedded, hence the graph of a global 1-Lipschitz function. This shows that any proper achronal meridian Λ is the boundary at infinity of a properly embedded spacelike surface, which improves the statement of Lemma 15.2.1.3.1.

The most remarkable consequence of Proposition 15.2.1.4.6 is that the domain of dependence of a properly embedded surface in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ only depends on the boundary at infinity. More precisely we have:

Corollary 15.2.1.4.8 *If S and S' are properly embedded spacelike surfaces in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, then $\mathcal{D}(S) = \mathcal{D}(S')$ if and only if $\partial S = \partial S'$.*

15.2.1.5 Properly Achronal Sets in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$

It will be important for the applications of this theory to consider the model $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. As $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ contains closed timelike lines, it does not contain any achronal subset. However if P is a spacelike plane in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, then $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \setminus P$ does not contain closed causal curves. Indeed it is simply connected, so it admits an isometric embedding into $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, given by a section of the covering map $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, and whose image is a Dirichlet region.

Definition 15.2.1.5.1 A subset X of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cup \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is a *proper achronal subset* if there exists a spacelike plane P such that X is contained in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cup \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \setminus \overline{P}$ and is achronal as a subset of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cup \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \setminus \overline{P}$.

Notice that if X is a proper achronal subset of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cup \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, then it admits a section to $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \cup \partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ and the image is achronal in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \cup \partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$. Conversely if $\widetilde{X}$ is an achronal subset of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ different from a lightlike plane, then it is contained in a Dirichlet region, as a consequence of Lemma 15.2.1.1.7 and the fact that any achronal subset of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ is contained in a properly embedded one. As Dirichlet regions are projected in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ to the complement of a spacelike plane, the image of $\widetilde{X}$ to $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is a proper achronal subset.

Let us provide an important example which will be extensively used later.

Lemma 15.2.1.5.2 *Let $\varphi : \mathbb{R}\mathrm{P}^1 \to \mathbb{R}\mathrm{P}^1$ be an orientation preserving homeomorphism. Then the graph of φ, say $\Lambda_\varphi \subset \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1 \cong \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is a proper achronal subset and any lift $\widetilde{\Lambda}_\varphi$ is an achronal meridian in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$.*

Proof First let us prove that Λ_φ is locally achronal. In fact if U and V are intervals around x and $\varphi(x)$ and θ_1 and θ_2 are positive coordinates on U and V respectively, then timelike curves $\gamma(t) = (\gamma_1(t), \gamma_2(t))$ in $U \times V$ are characterized by the property that $\theta_1'(t)\theta_2'(t) < 0$, where we have put $\theta_i(t) := \theta_i(\gamma_i(t))$. (See Proposition 15.1.2.2.3 and the following paragraph.) In particular points on $\Lambda_\varphi \cap U \times V$ are not related by a timelike curve contained in $U \times V$, by the assumption that φ is orientation-preserving.

Let us prove that there exists a spacelike plane P such that $\overline{P} \cap \Lambda_\varphi = \emptyset$. Let us consider the identification $\mathbb{R}\mathrm{P}^1 = \mathbb{R} \cup \{\infty\}$, and take $\varphi_0 \in \mathrm{PSL}(2, \mathbb{R})$ so that $\varphi_0^{-1}\varphi(0) = 1$, $\varphi_0^{-1}\varphi(1) = \infty$, and $\varphi_0^{-1}\varphi(\infty) = 0$. Then notice that $\varphi_0^{-1}\varphi$ sends the intervals $(\infty, 0)$, $(0, 1)$ and $(1, \infty)$ respectively to $(0, 1)$, $(1, \infty)$, $(\infty, 0)$. Thus $\varphi_0^{-1}\varphi$ has no fixed points, that is, the graph of φ does not meet the graph of φ_0, which is the asymptotic boundary of a spacelike plane P_{φ_0}.

Let us consider now a lift of Λ_φ to the boundary of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, say $\widetilde{\Lambda}_\varphi$. As Λ_φ is contained in a simply connected region of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cup \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, $\widetilde{\Lambda}_\varphi$ is a closed locally achronal curve contained in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$. In particular the projection $\widetilde{\Lambda}_\varphi \to \partial\mathbb{D}$ is locally injective. As $\widetilde{\Lambda}_\varphi$ is compact, the map is a covering. On the other hand, since Λ_φ is homotopic to the boundary of a plane in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, it turns out that $\widetilde{\Lambda}_\varphi$ is homotopic to $\partial\mathbb{D}$ in $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ so that the projection $\widetilde{\Lambda}_\varphi \to \partial\mathbb{D}$ is bijective. It follows that $\widetilde{\Lambda}_\varphi$ is achronal, and the conclusion follows. □

All the results we have proven for achronal sets in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ can be rephrased for proper achronal sets of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. For instance any proper achronal set X can be extended to a properly embedded proper achronal surface and there are two extremal extensions, as in Lemma 15.2.1.2.2.

We will now focus on proper achronal meridians of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, which are proper achronal embedded circles of the boundary of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. They lift to achronal meridians of $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ different from the boundary of a lightlike plane. Indeed the boundary of a lightlike plane is not contained in a Dirichlet region. Conversely any achronal meridian of $\partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ different from the boundary of a lightlike plane projects to an achronal meridian of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$.

Proposition 15.2.1.5.3 *Let Λ be a proper achronal meridian in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ and denote by $\widetilde{\Lambda}$ any lift to the universal covering. Then the universal covering map of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ maps $\Omega(\widetilde{\Lambda})$ injectively to the domain*

$$\Omega(\Lambda) := \{x \in \mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \mid P_x \cap \Lambda = \emptyset\} .$$

Proof If $p : \widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ denotes the covering map, by Proposition 15.2.1.3.3 the invisible domain $\Omega(\widetilde{\Lambda})$ is contained in a Dirichlet region $R_{\tilde{x}}$, hence the restriction of p to $\Omega(\widetilde{\Lambda})$ is injective and its image is contained in $p(R_{\tilde{x}})$, namely the complement in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cup \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ of the spacelike plane P_x dual to $x = p(\tilde{x})$. Moreover by the first point of Proposition 15.2.1.3.2, one can actually pick for $\tilde{x}$ any point in $\Omega(\widetilde{\Lambda})$, which shows that the image $p(\Omega(\widetilde{\Lambda}))$ is contained in $\Omega(\Lambda) := \{x \in \mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \mid P_x \cap \Lambda = \emptyset\}$.

For the converse inclusion, let $x \in \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ be a point whose dual plane P_x does not meet Λ. The preimage $p^{-1}(P_x)$ is a countable disjoint union of planes which disconnect $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \cup \partial\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ in a disjoint union of Dirichlet regions centered at preimages of x. The lift $\widetilde{\Lambda}$ is contained in exactly one such region, say $R_{\tilde{x}}$. By the first point of Proposition 15.2.1.3.2 $\tilde{x} \in \Omega(\widetilde{\Lambda})$ which implies that $x = p(\tilde{x})$ lies in $p(\Omega(\widetilde{\Lambda}))$. □

When Λ is the graph of an orientation-preserving homeomorphism $\varphi : \mathbb{R}P^1 \to \mathbb{R}P^1$, there is a fairly simple characterization of $\Omega(\Lambda)$ using the identification $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} = \mathrm{PSL}(2, \mathbb{R})$.

Corollary 15.2.1.5.4 *Let φ be an orientation-preserving homeomorphism. Then $x \in \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ lies in $\Omega(\Lambda_\varphi)$ if and only if $x \circ \varphi$ has no fixed point as a homeomorphism of $\mathbb{R}P^1$.*

Proof It is easy to check that the dual plane of x, as an element of $\mathrm{PSL}(2, \mathbb{R})$, meets $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ along the graph of x^{-1}, say $\Lambda_{x^{-1}}$. Indeed this is easily checked if $x = \mathrm{id}$ is the identity by the description of $P_{\mathbb{1}}$ we gave in Sect. 15.1.2.5 together with Lemma 15.1.2.2.2. The general case then follows by applying left multiplication by x itself, which maps the graph of the identity to the graph of x^{-1}.

With this remark in hand, we have that $x \in \Omega(\Lambda_\varphi)$ if and only if $\Lambda_{x^{-1}} \cap \Lambda_\varphi = \emptyset$. This condition is equivalent to requiring that $x \circ \varphi$ has no fixed point on $\mathbb{R}P^1$. □

Proposition 15.2.1.5.5 *Let $\sigma : S \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ be a proper spacelike immersion. Then*

- *σ is a proper embedding.*
- *σ lifts to a proper embedding $\widetilde{\sigma} : S \to \widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$.*
- *The boundary at infinity of $\sigma(S)$ is a proper achronal meridian Λ in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$.*
- *$\mathcal{D}(\sigma(S)) = \Omega(\Lambda)$.*

Proof Denote by $\widehat{S}$ the covering of S admitting a lift $\widehat{\sigma} : \hat{S} \to \mathbb{H}^{2,1}$. In general either $\widehat{S} = S$ or it is a $2 : 1$ covering. Since the covering is finite, $\widehat{\sigma}$ is a proper immersion.

Let us consider the identification $\pi : \mathbb{H}^2 \times \mathbb{S}^1 \to \mathbb{H}^{2,1}$ defined in (15.1.8). The induced projection $\mathrm{pr} : \mathbb{H}^{2,1} \to \mathbb{H}^2$ is a proper fibration with timelike fibers. In particular $\widehat{\sigma}$ is trasverse to the fibers of pr. It follows that $\mathrm{pr} \circ \sigma : \widehat{S} \to \mathbb{H}^2$ is a proper local diffeomorphism, hence a covering map. Since $\mathbb{H}^2$ is simply connected, we deduce that the projection $\mathrm{pr} \circ \sigma : \widehat{S} \to \mathbb{H}^2$ is a homeomorphism, $\widehat{\sigma}$ is an embedding, and $\widehat{S}$ is homeomorphic to the plane.

In particular we can lift $\widehat{\sigma}$ to the universal covering, say $\widetilde{\sigma} : \widehat{S} \to \widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, which is still a proper spacelike embedding $\widehat{S} \to \widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$. By Lemmas 15.2.1.1.3 and 15.2.1.1.5 we know that the image is an achronal surface whose boundary is an achronal meridian, and is contained in a Dirichlet domain by Lemma 15.2.1.1.7. It follows that $\widetilde{\sigma}(\widehat{S})$ is contained in a Dirichlet domain of the covering map $\mathbb{H}^{2,1} \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, on which we know that the covering map is injective. In particular σ is also injective, hence $\widehat{S} = S$ and this concludes the proof. □

Remark 15.2.1.5.6 In the proof of Proposition 15.2.1.5.5, once we proved that $\widehat{S}$ is homeomorphic to $\mathbb{R}^2$, then we could have inferred immediately that $\widehat{S} = S$ since it is known, although non-trivial, that $\mathbb{Z}/2\mathbb{Z}$ cannot act freely on $\mathbb{R}^2$ by diffeomorphisms.

We therefore have the following analogue version of Corollary 15.2.1.4.8 in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$.

Corollary 15.2.1.5.7 *If S and S' are properly embedded spacelike surfaces in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, then $\mathcal{D}(S) = \mathcal{D}(S')$ if and only if $\partial S = \partial S'$.*

15.2.1.6 Convexity Notions

Let Λ be a proper achronal meridian in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. In this section we will investigate the convexity properties of $\Omega(\Lambda)$.

Let us recall that $X \subset \mathbb{RP}^3$ is convex if it is contained in an affine chart and it is convex in the affine chart. This notion does not depend on the affine chart containing X. It is a proper convex set if it is moreover compactly contained in an affine chart.

Proposition 15.2.1.6.1 *Given a proper achronal meridian Λ in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, $\Omega(\Lambda)$ is convex. If Λ is different from the boundary of a spacelike plane then $\Omega(\Lambda)$ is a proper convex set.*

Proof By Proposition 15.2.1.3.3 there exists a spacelike plane P such that $\Omega(\Lambda)$ is contained in the affine chart V of $\mathbb{RP}^3$ obtained by removing the projective plane containing P. The domain $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cap V = \mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \setminus P$ is isometric to a Dirichlet region R of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, by an isometry sending Λ to a lifting $\widetilde{\Lambda}$ and $\Omega(\Lambda)$ to $\Omega(\widetilde{\Lambda})$. By the second point of Proposition 15.2.1.3.2 we have

$$\Omega(\widetilde{\Lambda}) = \bigcap_{\tilde{z}\in\widetilde{\Lambda}} \mathrm{I}^+(L_-(\tilde{z})) \cap \mathrm{I}^-(L_+(\tilde{z})) \ .$$

Now if $\tilde{z}$ projects to z, then the images of $L_-(\tilde{z})$ and $L_+(\tilde{z})$ in V are the two components of $L(z)\cap\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, where $L(z)$ is the affine tangent plane of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}\cap V$ at z. It turns out that the image of the region $\mathrm{I}^+(L_-(\tilde{z})) \cap \mathrm{I}^-(L_+(\tilde{z}))$ is the intersection of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ with the open half-space $U(z)$ bounded by $L(z)$ and whose closure contains Λ. This shows:

$$\Omega(\Lambda) = \mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cap \bigcap_{z\in\Lambda} U(z) \ .$$

We now claim that actually

$$\Omega(\Lambda) = \bigcap_{z\in\Lambda} U(z) \subset \mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \ ,$$

which will conclude the proof. As $\bigcap_{z\in\Lambda} U(z)$ is connected and meets $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, to show that it is contained in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ it is sufficient to show that it does not meet the boundary of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. For any $w \in \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ let us consider the leaf of the left ruling through w, which intersects Λ at a point z. It turns out that $L(z)$ contains the leaf of the left ruling through z, hence $w \notin U(z)$.

Now, assume that Λ is not the boundary of a spacelike plane. Then by Proposition 15.2.1.3.3 on the universal covering the compact set $\Omega(\widetilde{\Lambda}) \cup S_+(\widetilde{\Lambda}) \cup S_-(\widetilde{\Lambda})$ is contained in a Dirichlet domain, so its image is a compact set contained in an affine chart. □

A a consequence of the previous argument is that Λ is contained in an affine chart whose complement in $\mathbb{RP}^3$ is a projective plane containing a spacelike plane of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. (Indeed Λ is contained in the closure of $\Omega(\Lambda)$, which is contained in an affine chart, unless Λ is the boundary of a spacelike plane, in which case the statement is trivial.) Hence it makes sense to give the following definition:

Definition 15.2.1.6.2 Given a proper achronal meridian Λ in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, we define $C(\Lambda)$ to be the convex hull of Λ, which can be taken in an affine chart containing Λ.

Observe that we have proved implicitly that if Λ is an achronal meridian in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, then $C(\Lambda)$ is contained in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, which is not obvious as $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is not convex in $\mathbb{R}\mathrm{P}^3$.

Remark 15.2.1.6.3 Since $\Omega(\Lambda)$ is convex, $C(\Lambda)$ is contained in $\Omega(\Lambda)$. Moreover if K is any convex set contained in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cup \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ and containing Λ, then $C(\Lambda) \subset K \subset \overline{\Omega(\Lambda)}$.

To see this, let V be an affine chart such that $\Lambda \subset V$ is obtained by removing a spacelike projective plane. Now, if $z \in \Lambda$ then for any $x \in \mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cap V$ the segment connecting z and x in V is contained in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ if and only if $x \in U(z)$, the half-space containing Λ and bounded by the tangent space of Λ at z, as defined in the proof of Proposition 15.2.1.6.1.

Hence by the characterization of $\overline{\Omega(\Lambda)}$ as the intersection of the $U(z)$ given in Proposition 15.2.1.6.1, if x is not in $\overline{\Omega(\Lambda)}$, then it cannot be in K. This shows that $\overline{\Omega(\Lambda)}$ is the biggest convex subset of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ containing Λ.

Assume now that Λ is not the boundary of a spacelike plane. Then the topological frontiers in $\mathbb{R}\mathrm{P}^3$ of $\Omega(\Lambda)$ and of $C(\Lambda)$ are Lipschitz surfaces homeomorphic to a sphere. This sphere is disconnected by Λ into two regions, homeomorphic to disks, which form the boundary of $\Omega(\Lambda)$ and of $C(\Lambda)$ in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. For $\Omega(\Lambda)$ those components are the image of $S_\pm(\widetilde{\Lambda})$ and will be denoted by $S_\pm(\Lambda)$.

Let us now focus on $C(\Lambda)$. Let $C(\widetilde{\Lambda})$ be a lifting of $C(\Lambda)$, which is contained in a Dirichlet region, say R. Let P be a support plane for $C(\Lambda)$, which is necessarily either spacelike or lightlike, and let $\widetilde{P}$ be its lift which touches $C(\widetilde{\Lambda})$. Hence either $\widetilde{\Lambda}$ is in $\mathrm{I}^+(\widetilde{P}) \cup \widetilde{P}$ or in $\mathrm{I}^-(\widetilde{P}) \cup \widetilde{P}$. This permits to distinguish the components of $\partial C(\Lambda) \setminus \Lambda$: the *past boundary component* $\partial_- C(\Lambda)$ has the property that $\widetilde{\Lambda}$ is contained in $\mathrm{I}^+(\widetilde{P}) \cup \widetilde{P}$ for all support planes which touch $\partial_- C(\Lambda)$, and analogously we define the *future boundary component* $\partial_+ C(\Lambda)$ by replacing I^+ with I^-. The following proposition explains that the boundary components $\partial_\pm C(\Lambda)$ and $S_\pm(\Lambda)$ have a kind of duality.

Proposition 15.2.1.6.4 *Let Λ be a proper achronal meridian in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, $x \in \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ and P_x the dual plane. Then*

- $x \in \Omega(\Lambda)$ *if and only if* $P_x \cap C(\Lambda) = \emptyset$.
- $x \in C(\Lambda)$ *if and only if* $P_x \cap \Omega(\Lambda) = \emptyset$.

In particular if Λ is not the boundary of a spacelike plane, then

- $x \in \partial_\pm \Omega(\Lambda)$ *if and only if P_x is a support plane for* $\partial_\mp C(\Lambda)$.
- $x \in \partial_\pm C(\Lambda)$ *if and only if P_x is a support plane for* $S_\mp(\Lambda)$.

Proof From Proposition 15.2.1.5.3, points in $\Omega(\Lambda)$ are dual to planes disjoint from Λ, which are precisely those which do not intersect $C(\Lambda)$, by the definition of convex hull. For the second statement, fix x and observe that $z \in P_x$ if and only if $x \in P_z$. Hence there exists a point z in the intersection $P_x \cap \Omega(\Lambda)$ if and only if x is in a plane P_z which is disjoint from Λ, namely when x is not in $C(\Lambda)$.

As a consequence $\partial C(\Lambda)$ consists of points dual to support planes of $\Omega(\Lambda)$. Take a support plane P_x of $S_+(\Lambda)$ (hence dual to a point x) which meets $S_+(\Lambda)$ at z. If $\tilde{z}$ denotes the corresponding point on $S_+(\widetilde{\Lambda})$, then $\widetilde{\Lambda} \subset \mathrm{I}^+(P^-_{\tilde{z}})$, and $P^-_{\tilde{z}} \cap \tilde{\Lambda} \neq \emptyset$. Thus P_z, which is the projection of $P^-_{\tilde{z}}$, is a support plane of $C(\Lambda)$ touching the past boundary. As $x \in P_z$, we conclude that x lies in the past boundary. Similarly points of the future boundary of $C(\Lambda)$ correspond to support planes for $S_-(\Lambda)$. □

Remark 15.2.1.6.5 It may happen that a boundary component of $C(\Lambda)$ meets the boundary of $\Omega(\Lambda)$. This exactly happens when the curve Λ contains a *sawtooth*, namely two consecutive lightlike segments in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ one past directed and the other future-directed. In this case the lightlike plane $L(z)$ tangent to $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ at the vertex z of the sawtooth contains the two consecutive lightlike segments of Λ, while the convex hull of Λ contains a lightlike triangle contained in $L(z)$. This is however not contained in $\Omega(\Lambda)$. If the curve Λ does not contain any sawtooth, then $C(\Lambda) \setminus \Lambda$ is entirely contained in $\Omega(\Lambda)$.

The fundamental example is given in Fig. 15.8, where the yellow region represents at the same time the convex hull of the proper achronal meridian Λ in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ composed of four lightlike segments, two past-directed and two future-directed, and the closure of $\Omega(\Lambda)$. See also Remark 15.3.2.1.3 and Fig. 15.14 below.

Proposition 15.2.1.6.6 *The past and future boundary components of $C(\Lambda)$ are achronal surfaces.*

Proof Let us give the proof for $\partial_+ C(\Lambda)$. Take $x, y \in \partial_+ C(\Lambda)$ and consider the segment joining x to y in an affine chart containing Λ. If this segment was timelike then the dual planes P_x and P_y would be disjoint. Then up to switching x to y we may assume that, in the universal cover, $P^1_{\tilde{x}} \subset \mathrm{I}^+(P^1_{\tilde{y}})$, where $\tilde{x}$ and $\tilde{y}$ are the lifting of x and y in the same Dirichlet region mapping to the fixed affine chart. But then $S_+(\widetilde{\Lambda})$ would be contained in $\mathrm{I}^-(P^1_{\tilde{y}})$ and could not meet $P^1_{\tilde{x}}$, thus contradicting Proposition 15.2.1.6.4. □

Remark 15.2.1.6.7 The past and future boundary components of $C(\Lambda)$ are not smooth, but only Lipschitz surfaces. Indeed the complement of Λ and of the lightlike triangles (as described in Remark 15.2.1.6.5) is locally connected by acausal Lipschitz arcs, and one can define a pseudo-distance, that in fact turns out to be a distance and makes $C(\Lambda)$ locally isometric to the hyperbolic plane.

The situation is very similar to the counterpart in hyperbolic three-space. The intersection of a spacelike support plane with $C(\Lambda)$ is either a geodesic or a straight convex subset of $\mathbb{H}^2$, i.e. a subset bounded by geodesics. Thus $\partial C(\Lambda) \setminus \Lambda$ is intrinsically a hyperbolic surface pleated along a measured geodesic lamination. A remarkable difference with respect to the hyperbolic case is that in general those surfaces may be non complete, but they are always isometric to straight convex subsets of $\mathbb{H}^2$. See [32] for more details.

15.2.2 Globally Hyperbolic Three-Manifolds

The aim of this section is to study maximal globally hyperbolic (MGH) Anti-de Sitter spacetimes containing a compact Cauchy surface of genus r (we briefly say that the globally hyperbolic spacetimes have genus r). We first prove that there are no examples for $r = 0$. We will then briefly consider the torus case, and finally we will deepen the study for $r \geq 2$, first by introducing examples, and then by giving a complete classification.

15.2.2.1 General Facts

We begin by some general results which will be used both in the genus one and in the higher genus case. Recall that an immersion $\sigma : S \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is spacelike if the pull-back metric (also called first fundamental form) is a Riemannian metric. We will provide more details on the theory of spacelike immersions in Sect. 15.2.3.1 below, which can be read independently.

Lemma 15.2.2.1.1 *Let $\sigma : S \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ be a spacelike immersion. If $\sigma^*(g_{\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}})$ is a complete Riemannian metric, then σ is a proper embedding and S is diffeomorphic to $\mathbb{R}^2$.*

Proof By Proposition 15.2.1.5.5 it is sufficient to prove that σ is a proper immersion. In the notation of Proposition 15.2.1.5.5, consider a lift $\widehat{\sigma} : \widehat{S} \to \mathbb{H}^{2,1}$. It is clearly sufficient to prove that $\widehat{\sigma}$ is proper. We will prove that if $\gamma : [0, 1) \to \widehat{S}$ is a path such that the limit $\lim_{t \to 1} \widehat{\sigma}(\gamma(t))$ exists, then also $\lim_{t \to 1} \gamma(t)$ exists.

Using the expression (15.1.10) for the metric on $\mathbb{H}^{2,1}$ under the identification with $\mathbb{H}^2 \times \mathbb{S}^1$ given by (15.1.8), we see that the length of γ for the pull-back metric is smaller that the length of the projection of γ to the $\mathbb{H}^2$ factor, with respect to the hyperbolic metric on $\mathbb{H}^2$. The latter hyperbolic length is finite by the assumption, hence γ has finite length for the pull-back metric. The assumption on the completeness of the pull-back metric implies the existence of the limit point for $\gamma(t)$. □

As an immediate consequence, there can be no globally hyperbolic AdS spacetime whose Cauchy surfaces are diffeomorphic to the sphere. In fact, supposing such a spacetime exists and Σ is a Cauchy surface, the developing map restricted to Σ would be a spacelike immersion, and the pull-back metric would be complete by compactness. But this contradicts Lemma 15.2.2.1.1. Hence we proved:

Corollary 15.2.2.1.2 *There exists no globally hyperbolic Anti-de Sitter spacetime of genus zero.*

The following is a fundamental result on the structure of globally hyperbolic AdS spacetimes.

Proposition 15.2.2.1.3 *Let M be a globally hyperbolic Anti-de Sitter spacetime of genus $r \geq 1$. Then*

(1) The developing map dev : $\widetilde{M} \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is injective.
(2) If Σ is a Cauchy surface of M, then the image of dev is contained in $\Omega(\Lambda)$, where Λ is the boundary at infinity of dev$(\widetilde{\Sigma})$.
(3) If $\rho : \pi_1(M) \to \mathrm{Isom}(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1})$ is the holonomy representation, $\rho(\pi_1(M))$ acts freely and properly discontinuously on $\Omega(\Lambda)$, and $\Omega(\Lambda)/\rho(\pi_1(M))$ is a globally hyperbolic spacetime containing M.

Proof Let $\widetilde{\mathrm{dev}} : \widetilde{M} \to \widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ be a lift of dev to the universal cover. By Theorem 15.2.1.4.2, the spacetime M admits a foliation by smooth spacelike surfaces $(\Sigma_t)_{t\in\mathbb{R}}$ of genus $r \geq 1$, such that $\Sigma_t \subset \mathrm{I}^+(\Sigma_{t'})$ for $t > t'$. Let $\widetilde{\Sigma}_t$ the lift of the foliation on $\widetilde{M}$. Since Σ_t is closed, the induced metric on Σ_t is complete, and so is the induced metric on $\widetilde{\Sigma}_t$. As $\widetilde{\mathrm{dev}}$ is a local isometry, we deduce by Lemma 15.2.2.1.1 that the restriction of $\widetilde{\mathrm{dev}}$ to $\widetilde{\Sigma}_t$ is a proper embedding.

Assume now by contradiction that $\widetilde{\mathrm{dev}}(\widetilde{\Sigma}_t) \cap \widetilde{\mathrm{dev}}(\widetilde{\Sigma}_{t'}) \neq \emptyset$ for some $t \geq t'$. Then there is a point $x \in \widetilde{\Sigma}_t$ such that $\widetilde{\mathrm{dev}}(x) \in \widetilde{\mathrm{dev}}(\widetilde{\Sigma}_{t'})$. By the assumption x is connected to $\widetilde{\Sigma}_{t'}$ by a timelike arc η in $\widetilde{M}$. Then $\widetilde{\mathrm{dev}}(\eta)$ is a timelike arc in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ with end-points in $\widetilde{\mathrm{dev}}(\widetilde{\Sigma}_{t'})$ and this contradicts the achronality of $\widetilde{\mathrm{dev}}(\widetilde{\Sigma}_{t'})$. This shows that $\widetilde{\mathrm{dev}}$ is injective, and moreover we conclude that $\widetilde{\mathrm{dev}}(\widetilde{\Sigma}_t)$ is a Cauchy surface of $\widetilde{\mathrm{dev}}(\widetilde{M})$. It follows using Proposition 15.2.1.5.5 that $\widetilde{\mathrm{dev}}(\widetilde{M}) \subset \mathcal{D}(\widetilde{\mathrm{dev}}(\widetilde{\Sigma}_t)) = \Omega(\widetilde{\Lambda})$, where $\widetilde{\Lambda}$ is the boundary at infinity of $\widetilde{\mathrm{dev}}(\widetilde{\Sigma}_t)$, which proves the second point.

Now, the map $\widetilde{\mathrm{dev}}$ is $\widetilde{\rho}$-equivariant, for a representation $\widetilde{\rho} : \pi_1(M) \to \mathrm{Isom}(\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1})$ which is a lift of the holonomy of M. As $\widetilde{\mathrm{dev}}(\widetilde{\Sigma}_t)$ is $\widetilde{\rho}$-invariant, then so are $\widetilde{\Lambda}$ and $\Omega(\widetilde{\Lambda})$. We shall prove that the action of $\pi_1(M)$ on $\Omega(\widetilde{\Lambda})$ given by $\widetilde{\rho}$ is proper. This will also show that the action is free, since $\pi_1(M)$ is isomorphic to $\pi_1(\Sigma_r)$ and therefore has no torsion.

For this purpose, let us notice that if K is relatively compact in $\Omega(\widetilde{\Lambda})$ then

$$X_K := (\mathrm{I}^+(K) \cup \mathrm{I}^-(K)) \cap \widetilde{\mathrm{dev}}(\widetilde{\Sigma}_t)$$

is relatively compact as well. As the action of $\pi_1(M)$ on $\widetilde{\Sigma}_t$, and thus on $\widetilde{\mathrm{dev}}(\widetilde{\Sigma}_t)$, is proper and $X_{\gamma K} = \gamma(X_K)$, we deduce that the set of γ such that $X_{\gamma K} \cap X_K \neq \varnothing$ is finite. On the other hand if $K \cap \gamma K \neq \emptyset$ then $X_K \cap X_{\gamma K} \neq \emptyset$. We thus conclude that the action is proper. By applying the path lifting property, one sees that the quotient $\widetilde{\mathrm{dev}}(\widetilde{\Sigma}_t)/\pi_1(M)$ is a Cauchy surface of $\Omega(\widetilde{\Lambda})/\pi_1(M)$, which is therefore globally hyperbolic.

The proof of the statement is then accomplished since by Proposition 15.2.1.5.3 the restriction of the covering map $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1} \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ to $\Omega(\widetilde{\Lambda}) \cup \Lambda$ is injective. □

A remarkable difference between Lorentzian and Riemannian geometry is that in Lorentzian geometry geodesic completeness is a very strong assumption, and in fact interesting classification results are obtained without such an assumption. However, it is necessary to impose some maximality condition to compensate for

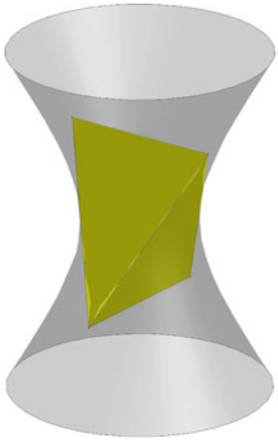


Fig. 15.8 The lightlike tetrahedron $\mathscr{T}$: two of its edges are spacelike lines of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, dual to one another (on the top and bottom), and the other four are lightlike segments contained in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$

non-completeness. Among several approaches, one of the most common is the classification of a maximal globally hyperbolic spacetimes. We give a definition here in our special setting, although one can give more general definitions in the larger class of Einstein spacetimes.

Definition 15.2.2.1.4 A globally hyperbolic Anti-de Sitter manifold (M, g) is *maximal* if any isometric embedding of (M, g) into a globally hyperbolic Anti-de Sitter manifold (M', g'), which sends a Cauchy surface of (M, g) to a Cauchy surface of (M', g'), is surjective.

The following corollary is a direct consequence of Proposition 15.2.2.1.3 and Definition 15.2.2.1.4.

Corollary 15.2.2.1.5 *An Anti de-Sitter globally hyperbolic spacetime M is maximal if and only if $\widetilde{M}$ is isometric to the invisible domain of a proper achronal meridian in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$.*

15.2.2.2 Genus $r = 1$: Examples

Our first objective is the classification of MGH AdS spacetimes of genus 1. This case has not been considered in the paper of Mess. However it has been studied in the physics literature, for instance in [59] and [36]. We start by constructing a family of examples, which will later be shown to be all examples of genus 1 up to isometry, thus providing a full classification.

Recall from Definition 15.1.2.5.4 the construction of dual spacelike lines, as in Fig. 15.8. In the $\mathrm{PSL}(2, \mathbb{R})$ model, up to isometry the two dual spacelike lines $\mathcal{L}$ and $\mathcal{L}'$ can be taken of the form $\mathcal{L} = L_{\ell,\ell}$ where ℓ is an oriented spacelike geodesic in $\mathbb{H}^2$, and $\mathcal{L}' = L_{\ell,\ell'}$ where ℓ' is ℓ endowed with the opposite orientation. This means that $\mathcal{L}$ consists of the hyperbolic isometries of $\mathbb{H}^2$ which translate along the geodesic ℓ, while $\mathcal{L}'$ consists of order-two elliptic elements with fixed point in ℓ.

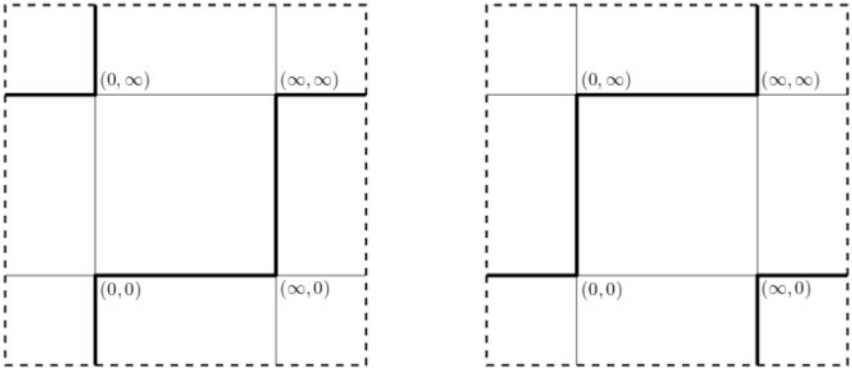


Fig. 15.9 A schematic picture of the two curves of Lemma 15.2.2.2.1 in the torus $\mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$, represented as a square with sides identified by translations

By Proposition 15.1.2.5.3, the endpoints of $\mathcal{L}$ are of the form (x, y) and (y, x) in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cong \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$, for x and y the endpoints of ℓ in $\mathbb{R}\mathrm{P}^1$, while the endpoints of $\mathcal{L}'$ are of the form (x, x) and (y, y).

The following lemma exhibits proper achronal meridians in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ containing these four points, each of which, together with the two dual lines $\mathcal{L}$ and $\mathcal{L}'$, constitute the 1-skeleton of the affine tetrahedron as in Fig. 15.8.

Lemma 15.2.2.2.1 *Let x, y be different points in $\mathbb{R}\mathrm{P}^1$. Then there exist exactly two proper achronal meridians in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ containing the points (x, x), (y, x), (y, y), (x, y).*

Proof Since (x, x) and (y, x) are in the same leaf λ_x of left ruling of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \cong \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$, a proper achronal meridian must necessily contain one of the two segments connecting (x, x) and (y, x) in λ_x, thus giving two possible choices. Once this choice is made, the same argument applies for the leaf μ_y of the right ruling containing (y, x) and (y, y), but there is only one possible choice so as to give, concatenated with the previously chosen segment in λ_x, a locally achronal curve.

More precisely, if we choose an affine chart which contains the four points (x, x), (y, x), (y, y) and (x, y) and assume the segment chosen in the first step from (x, x) to (y, x) is future-directed in this affine chart, then the segment connecting (y, x) to (y, y) must necessarily be past directed. One then iterates this argument and obtains precisely two proper achronal meridians: if we assume for simplicity that $x = 0$ and $y = \infty$ in $\mathbb{R}\mathrm{P}^1 \cong \mathbb{R} \cup \{\infty\}$, the first is the concatenation of $[0, \infty] \times \{0\}$, $\{\infty\} \times [0, \infty]$, $[\infty, 0] \times \{\infty\}$ and $\{0\} \times [\infty, 0]$; the second the concatenation of $[\infty, 0] \times \{0\}$, $\{\infty\} \times [\infty, 0]$, $[0, \infty] \times \{\infty\}$ and $\{0\} \times [0, \infty]$. See Fig. 15.9. □

Let us call Λ_1 and Λ_2 the two proper achronal meridians described in Lemma 15.2.2.2.1. Their lifts on the universal cover $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ are easily described. For this purpose, let us fix $x, y \in \mathbb{R}\mathrm{P}^1$ and let us choose a lift $\widetilde{\mathcal{L}}$ to $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ of the spacelike geodesic in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ connecting $p_1 = (x, y)$ and $p_2 = (y, x)$. Say $\widetilde{p}_1 = (\xi_1, \mathsf{t}_1)$ and $\widetilde{p}_2 = (\xi_2, \mathsf{t}_2)$ are the endpoints of $\widetilde{\mathcal{L}}$ in the boundary $\partial\mathbb{D} \times \mathbb{R}$ of

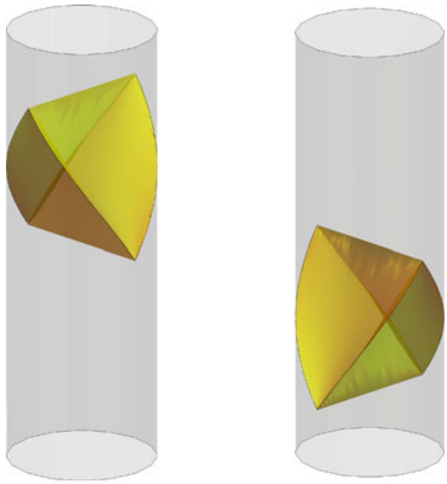


Fig. 15.10 The invisible domains of the two achronal meridians $\widetilde{\Lambda}_1$ and $\widetilde{\Lambda}_2$ composed of four lightlike segments in $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$. The 1-skeleton of the two tetrahedra contains four lightlike segments together with two dual spacelike lines. The left and right tetrahedra actually differ by rotating on $\mathbb{D}$ and translating vertically

the Poincaré model of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$. (Up to isometries, we could in fact assume that ξ_1 and ξ_2 are antipodal on $\mathbb{S}^1$ and $\mathsf{t}_1 = \mathsf{t}_2 = 0$.)

Then $\widetilde{\Lambda}_1$ and $\widetilde{\Lambda}_2$ can be expressed as the graphs of $\mathsf{f}^{\Lambda_i} : \partial\mathbb{D} \to \mathbb{R}$ defined by:

$$\mathsf{f}^{\Lambda_1}(\xi) = \min\{d_{\mathbb{S}^2}(\xi, \xi_1) + \mathsf{t}_1, d_{\mathbb{S}^2}(\xi, \xi_2) + \mathsf{t}_2\}\,, \tag{15.2.3}$$

and

$$\mathsf{f}^{\Lambda_2}(\xi) = \max\{\mathsf{t}_1 - d_{\mathbb{S}^2}(\xi, \xi_1), \mathsf{t}_2 - d_{\mathbb{S}^2}(\xi, \xi_2)\}\,. \tag{15.2.4}$$

Indeed, for f^{Λ_1}, since (ξ_1, t_1) and (ξ_2, t_2) are connected by a spacelike line, $\mathsf{f}^{\Lambda_1}(\xi_1) = \mathsf{t}_1$ and $\mathsf{f}^{\Lambda_1}(\xi_2) = \mathsf{t}_2$; moreover there are two points $\widetilde{q}_1 = (\eta_1, \mathsf{s}_1)$ and $\widetilde{q}_2 = (\eta_2, \mathsf{s}_2)$ at which the expressions $d_{\mathbb{S}^2}(\xi, \xi_1) + \mathsf{t}_1$ and $d_{\mathbb{S}^2}(\xi, \xi_2) + \mathsf{t}_2$ are equal, which are the endpoints of one lift of the dual line $\mathcal{L}'$. Hence the graph of f^{Λ_1} consists of four lightlike segments, two future-directed and two past-directed. By the way, observe that f^{Λ_1} could be written by the equivalent expression:

$$\mathsf{f}^{\Lambda_1}(\xi) = \max\{\mathsf{s}_1 - d_{\mathbb{S}^2}(\xi, \eta_1), \mathsf{s}_2 - d_{\mathbb{S}^2}(\xi, \eta_2)\}\,. \tag{15.2.5}$$

This analysis turns out to be extremely useful for the description of the invisible domain and the convex hull of Λ_1 and Λ_2. These are pictured in Fig. 15.10 below.

Proposition 15.2.2.2.2 *Let x, y be distinct points in $\mathbb{R}\mathrm{P}^1$ and let Λ_0 be a proper achronal meridian in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ containing the points $(x, x), (y, x), (y, y), (x, y)$. Then $\overline{\Omega(\Lambda_0)} = C(\Lambda_0)$ is a tetrahedron bounded by four lightlike planes.*

Proof Let us first consider the picture in the universal cover $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, and consider the lift $\widetilde{\Lambda}_1$ defined as the graph of f^{Λ_1} as in Eq. (15.2.3). As a simple consequence of the triangular inequality for the hemispherical metric, one sees that the functions f_-^X

and f_{+}^{X} we introduced in Sect. 15.2.1.2 and Lemma 15.2.1.2.2 (where now $X = \widetilde{\Lambda}_1$) are given themselves by the expressions of Eqs. (15.2.3) and (15.2.5) respectively, except that the point ξ is now allowed to vary in $\mathbb{D} \cup \partial\mathbb{D}$.

Using the description of lightlike planes we gave in Sect. 15.1.1.8, see also Fig. 15.4, the surfaces $S_{\pm}(\widetilde{\Lambda}_1)$ (which we recall are the graph of $\mathsf{f}_{\pm}^{X}$) consist of two lightlike half-planes meeting in a spacelike geodesic: the geodesic with endpoints $\widetilde{q}_1$ and $\widetilde{q}_2$ for $S_{+}(\widetilde{\Lambda}_1)$; the geodesic with endpoints $\widetilde{p}_1$ and $\widetilde{p}_2$ for $S_{-}(\widetilde{\Lambda}_1)$. Projecting down to $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, the same description holds for $S_{\pm}(\Lambda_0)$. Hence $\Omega(\Lambda_0)$ is the interior of a tetrahedron with lightlike faces. Its closure, which is the tetrahedron itself, clearly coincides with the convex hull of Λ_0 in an affine chart, which is also the convex hull of $\mathcal{L} \cup \mathcal{L}'$. □

Remark 15.2.2.2.3 The region $\Omega(\Lambda_0)$ is, up to isometries, insensitive to the choice of Λ_0 as in Lemma 15.2.2.2.1. Namely, there is an orientation-preserving, time-preserving isometry of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ which maps one proper achronal meridian as in Lemma 15.2.2.2.1 to the other. The isometry is achieved simply by mapping the spacelike line $\mathcal{L}$ to its dual $\mathcal{L}'$, and therefore $\mathcal{L}'$ is mapped to $\mathcal{L}$.

In the universal cover $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$, this isometry is easily expressed if we normalize $\widetilde{\mathcal{L}}$ so that its endpoints in $\partial\mathbb{D} \times \mathbb{R}$ are of the form $\widetilde{p}_1 = (\xi_1, \mathsf{t}_1)$ and $\widetilde{p}_2 = (\xi_2, \mathsf{t}_2)$ with $\mathsf{t}_1 = \mathsf{t}_2$ and ξ_1, ξ_2 antipodal points in the sphere. Then the isometry we are looking for is induced by the isometry of $\widetilde{\mathbb{A}\mathrm{d}\mathbb{S}}^{2,1}$ which acts as a rotation of angle $\pi/2$ on $\mathbb{D}$ and a vertical translation of $\pi/2$ on $\mathbb{R}$. See again Fig. 15.10.

In what follows, we will refer to the region $\Omega(\Lambda_0)$, which is uniquely determined up to isometries, as the *lightlike tetrahedron* $\mathscr{T}$. To give a concrete description of the MGH spacetimes of genus one, the following geometric description of the tetrahedron, from an intrinsic point of view, will be useful.

Lemma 15.2.2.2.4 *The lightlike tetrahedron $\mathscr{T}$ is isometric to $\mathbb{R}^2 \times (0, \pi/2)$ endowed with the Lorentzian metric*

$$\cos^2(z)dx^2 + \sin^2(z)dy^2 - dz^2 \,. \tag{15.2.6}$$

Proof The easiest way to perform this computation is in the quadric model $\mathbb{H}^{2,1}$. Let us consider two lifts $\widehat{\mathcal{L}}$ and $\widehat{\mathcal{L}'}$ of the spacelike dual geodesics $\mathcal{L}$ and $\mathcal{L}'$ of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. It follows from the discussion of the duality in Sect. 15.1.1.9 that points in $\mathcal{L}'$ are the midpoints of the closed timelike geodesics leaving from $\mathcal{L}$ orthogonally. Hence in the double cover we have a timelike geodesic of length $\pi/2$ connecting every point of $\widehat{\mathcal{L}}$ to every point of $\widehat{\mathcal{L}'}$. Clearly these geodesics, projected to $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, foliate the interior of the convex hull of Λ_0, namely the lightlike tetrahedron $\mathscr{T}$.

Let $\gamma : \mathbb{R} \to \widehat{\mathcal{L}}$ and $\eta : \mathbb{R} \to \widehat{\mathcal{L}}$ be arclength parameterizations of the chosen spacelike geodesics in $\mathbb{H}^{2,1}$. By virtue of the above description, and using the expression (15.1.13) for the geodesics in the quadric model, we have the following diffeomorphism Φ between $\mathbb{R}^2 \times (0, \pi/2)$ and a lift of $\mathscr{T}$ in $\mathbb{H}^{2,1}$:

$$\Phi(x, y, z) = \cos(z)\gamma(x) + \sin(z)\eta(y) \,.$$

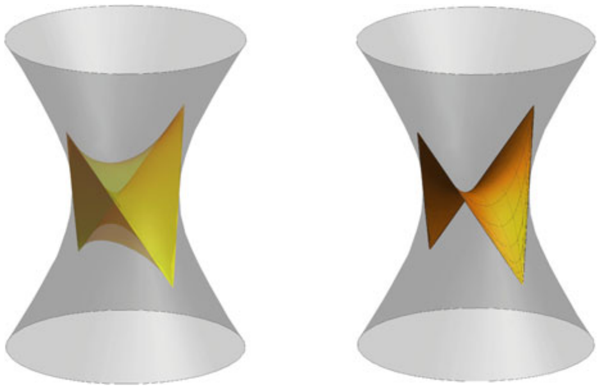


Fig. 15.11 The foliation of the lightlike tetrahedron $\mathscr{T} = \Omega(\Lambda_0)$ by flat CMC surfaces with constant values of z, in the coordinate system Φ. On the right the maximal surface corresponding to $z = \pi/4$ is highlighted

A direct computation, using that $\gamma(x)$ and $\eta(y)$ are orthogonal in $\mathbb{R}^{2,2}$ for every x, y, shows that the pull-back of the ambient metric $\langle \cdot, \cdot \rangle_{2,2}$ of $\mathbb{H}^{2,1}$ equals the metric (15.2.6). □

It is worth remarking that the surfaces given by $z = c$ under the diffeomorphism Φ are intrinsically flat and complete, hence properly embedded by Lemma 15.2.2.1.1. They are Cauchy surfaces for $\mathscr{T}$ by Proposition 15.2.1.5.5. See Fig. 15.11.

To conclude the construction of the examples, it only remains to study the stabilizer of the lines $\mathcal{L}$ and $\mathcal{L}'$. In light of the naturality of the construction of the dual line, the stabilizer of $\mathcal{L}$ actually coincides with the stabilizer of $\mathcal{L}'$. In the $\mathrm{PSL}(2,\mathbb{R})$-model, recall that we defined $\mathcal{L}$ as the one-parameter subgroup of $\mathrm{PSL}(2,\mathbb{R})$ of hyperbolic transformations which fix a geodesic ℓ in $\mathbb{H}^2$. The dual line consists of elliptic order-two isometries with fixed point on ℓ.

Let us denote by α_d the hyperbolic isometry which translates along ℓ of signed distance d. One then easily checks that the stabilizer of $\mathcal{L}$ which preserves an orientation of $\mathcal{L}$ is:

$$\mathrm{Stab}_+(\mathcal{L}) = \{(\alpha_l, \alpha_m) \mid l, m \in \mathbb{R}\} \subset \mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R}) , \qquad (15.2.7)$$

which is therefore isomorphic to $\mathbb{R}^2$. In fact, recalling the isometric action of $\mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$ on $\mathrm{PSL}(2,\mathbb{R})$ from Eq. (15.1.16), and the isometric identification of the dual plane $P_{\mathbb{1}}$ with $\mathbb{H}^2$ (Lemma 15.1.2.5.1), the isometries of the form (α_d, α_d) fix $\mathcal{L}$ pointwise and act on $\mathcal{L}'$ as a translation of length d. Conversely, the isometries of the form (α_d, α_{-d}) fix $\mathcal{L}'$ setwise and act on $\mathcal{L}$ as a translation of length d.

The orientation-preserving, time-preserving stabilizer consists of the normal subgroup $\mathrm{Stab}_+(\mathcal{L})$ and on another single coset, which consists of the rotations of angle π along each of the timelike geodesics leaving $\mathcal{L}$ orthogonally and connecting $\mathcal{L}$ to the dual geodesic $\mathcal{L}'$. In conclusion, we have the following:

Lemma 15.2.2.2.5 *The orientation-preserving, time-preserving stabilizer of $\mathscr{T}$ is isomorphic to the semidirect product $\mathbb{R}^2 \rtimes \mathbb{Z}/2\mathbb{Z}$. The normal subgroup $\mathbb{R}^2$ acts, in the coordinates given by Lemma 15.2.2.2.4, as*

$$(l, m) \cdot (x, y, z) = \left(x + \frac{l-m}{2}, y + \frac{l+m}{2}, z \right) ,$$

while a generator of the $\mathbb{Z}/2\mathbb{Z}$-factor acts as $(x, y, z) \mapsto (-x, -y, z)$.

The full stabilizer of $\mathscr{T}$ contains also orientation-reversing and time-reversing isometries, which can be easily figured out. Maximal globally hyperbolic spacetimes of genus 1 are then obtained as quotients of $\mathscr{T}$ by an action of $\mathbb{Z}^2$.

Proposition 15.2.2.2.6 *Given two linearly independent vectors (l, m) and (l', m'), the group $\mathbb{Z}^2$ generated by $\alpha = (\alpha_l, \alpha_m)$ and $\alpha' = (\alpha_{l'}, \alpha_{m'})$ acts freely and properly discontinuously on $\mathscr{T}$ and the quotient is a MGH spacetime of genus* 1.

Proof The vectors $((l-m)/2, (l+m)/2)$ and $((l'-m')/2, (l'+m'/2)$ are linearly independent if and only if (l, m) and (l', m') are linearly independent. It is then clear from Lemma 15.2.2.2.5, using the coordinates of Lemma 15.2.2.2.4, that the action on $\mathscr{T}$ is free and properly discontinuous. Since any surface $\{z = c\}$ is a Cauchy surface in $\mathscr{T}$, they project to Cauchy surfaces in the quotient, which is therefore globally hyperbolic, and maximal by Proposition 15.2.2.1.3. □

15.2.2.3 Genus $r = 1$: Classification

In this section we will prove that any MGH spacetime of genus $r = 1$ is isometric to one of those constructed in Proposition 15.2.2.2.6. The key step in the argument is the following proposition.

Proposition 15.2.2.3.1 *Let M be a globally hyperbolic spacetime of genus $r =$ 1 and let $\rho = (\rho_l, \rho_r) : \pi_1(T^2) \to \mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$ be the holonomy representation. Then ρ is discrete and faithful. Moreover ρ_l and ρ_r are elementary representations with two fixed points in $\mathbb{R}\mathrm{P}^1$.*

The last property in the statement is equivalent to the fact that $\rho_l(\gamma)$ and $\rho_r(\gamma)$ are hyperbolic transformations for any $\gamma \in \pi_1(T^2)$.

Proof By Proposition 15.2.2.1.3 the developing map $\mathrm{dev} : \widetilde{M} \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is injective, which implies that the holonomy representation is faithful. Moreover $\mathrm{dev}(\widetilde{M})$ is a domain in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ on which $\rho(\pi_1(T^2))$ acts properly. It follows that $\rho(\pi_1(T^2))$ is a discrete subgroup of $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$. The fact that ρ_l and ρ_r are elementary representations is a simple consequence of the fact that $\pi_1(T^2)$ is abelian. In order to prove that ρ_l and ρ_r fix two points on $\mathbb{R}\mathrm{P}^1$ we will show that no other possibility can hold.

First assume that both ρ_l and ρ_r have a fixed point in $\mathbb{H}^2$. Then ρ is conjugate to a representation in $\mathrm{PSO}(2) \times \mathrm{PSO}(2)$. But there is no faithful and discrete representation of $\pi_1(T^2)$ into a compact group.

To exclude the other possibilities we will use that, by Proposition 15.2.2.1.3, there is a proper achronal meridian Λ in $\partial \mathbb{A}\mathrm{d}\mathbb{S}^{2,1} = \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$ invariant under the representation ρ.

Assume first that ρ_l fixes a point in $\mathbb{H}^2$, and ρ_r fixes (at least) a point $y_0 \in \mathbb{R}\mathrm{P}^1$. For homological reasons the curve Λ must intersect the leaf λ_{y_0} at a point, say $p_0 = (x_0, y_0)$. Let γ be a non-trivial element of $\pi_1(T^2)$, and set $p_1 := \rho(\gamma)(x_0, y_0) = (\rho_l(\gamma)x_0, y_0)$. So Λ meets λ_{y_0} also at p_1. By Lemma 15.2.1.1.6, Λ contains a lightlike segment I in λ_{y_0} with end-points p_0 and p_1. Since $\bigcup_n \rho(\gamma)^n(I) = \lambda_{y_0}$ we deduce that Λ contains the entire leaf λ_{y_0} but this is a contradiction with Λ being a proper achronal meridian.

Let us now consider the case that $\rho_l(\gamma)$ and $\rho_r(\gamma)$ are parabolic transformations for all $c \in \pi_1(T)$. Up to conjugation we may assume that the fixed points of ρ_l and ρ_r are both $\infty \in \mathbb{R}\mathrm{P}^1$, hence ρ takes values into the subgroup G_∞ of $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$ fixing $p_\infty = (\infty, \infty)$. Notice that G_∞ acts by translations on the domain

$$U_0 = \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1 \setminus (\lambda_\infty \cup \mu_\infty) = (\mathbb{R}\mathrm{P}^1 \setminus \{\infty\}) \times (\mathbb{R}\mathrm{P}^1 \setminus \{\infty\}) = \mathbb{R}^2 ,$$

and such an action provides an isomorphism $G_\infty \cong \mathbb{R}^2$. Since the holonomy is discrete and faithful, $\rho(\pi_1(T^2))$ is identified to a lattice of G_∞. This implies that for every $p = (x_0, y_0) \in U_0$, the orbit of p is the set of vertices of a tessellation of $\mathbb{R}^2$ by parallelograms. In particular such an orbit must contain points of $\mathrm{I}^+_{U_0}(p) = \{(x, y) \,|\, x - x_0 > 0,\ y - y_0 < 0\}$, which shows that in U_0 there is no achronal orbit for the action of $\pi_1(T^2)$. It follows that Λ cannot meet U_0, so it is contained in $\lambda_\infty \cup \mu_\infty$. On the other hand, arguing as above we see that if Λ intersects the leaf λ_∞ (resp. μ_∞) at a point different from p_∞, then it must contain the whole leaf λ_∞ (resp. μ_∞), and this gives a contradiction.

Finally consider the case where for all $\gamma \in \pi_1(T^2)$ we have $\rho_l(\gamma)$ parabolic, and $\rho_r(\gamma)$ hyperbolic. We can assume that ∞ is the fixed point of ρ_l, and $0, \infty$ are the fixed points of ρ_r. We consider the partition of $\mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$ into ρ-invariant subsets μ_∞, λ_0, λ_∞, $U_+ = \mathbb{R} \times \mathbb{R}_+$, and $U_- = \mathbb{R} \times \mathbb{R}_-$. We will prove that no $\pi_1(T^2)$-orbit of $U_\pm$ is achronal, showing that $\Lambda \subset \mu_\infty \cup \lambda_0 \cup \lambda_\infty$. Let G be the subgroup of $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$ made of elements whose left factor is a parabolic transformation with fixed point at ∞ and whose right factor hyperbolic transformation with fixed points $0, \infty$. Let us consider the diffeomorphism

$$\Phi : \mathbb{R}^2 \to U_+ \qquad \Phi(x, y) = (x, e^y) ,$$

which conjugates the action of G_∞ on $\mathbb{R}^2$ and of G on U_+. In particular $\Phi^{-1} \circ \rho(\pi_1(T^2)) \circ \Phi$ is a lattice in G_∞. Thus as before no $\Phi^{-1} \circ \rho(\pi_1(T^2)) \circ \Phi$-orbit in

$\mathbb{R}^2$ is achronal. But Φ is conformal with respect to the Lorentzian metric $dxdy$ on $\mathbb{R}^2$ and the conformal Lorentzian structure of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ restricted to U_+. We deduce that no $\rho(\pi_1(T^2))$-orbit is achronal in U_+. A similar proof works for U_-.

This shows that Λ is contained in $\mu_\infty \cup \lambda_0 \cup \lambda_\infty$. Hence $\Lambda \cap \lambda_0$ is either one point or an arc. In the latter case the end-points of the arc should lie in ρ_∞, but the intersection of λ_0 and μ_∞ is only at the point $(\infty, 0)$, which contradicts that Λ is a proper achronal meridian. So Λ intersects λ_0 only at $(\infty, 0)$. Similarly Λ intersects λ_∞ only at (∞, ∞). This implies that $\Lambda \subset \mu_\infty$ which is a contradiction. □

Now, given a pair of elementary representations $\rho_l, \rho_r : \mathbb{Z}^2 \to \mathrm{PSL}(2, \mathbb{R})$ which map every non-trivial element to a hyperbolic transformation, assume for simplicity that the fixed points of ρ_l and ρ_r coincide, and let us call them x and y. Recall from Lemma 15.2.2.2.1 that there are two proper achronal meridians containing the four points (x, x), (y, x), (y, y), (x, y) in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. Each of them is clearly invariant under the $\mathbb{Z}^2$-action induced by ρ. The next step consists in showing that these are the only invariant proper achronal meridians.

Proposition 15.2.2.3.2 *Let $\rho : \pi_1(T^2) \to \mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$ be a representation such that ρ_l and ρ_r are elementary representations with two fixed points in $\mathbb{R}\mathrm{P}^1$. Then there are exactly two proper achronal meridians in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ which are invariant under the action of $\pi_1(T^2)$ induced by ρ.*

Proof Up to conjugation we may assume that both ρ_l and ρ_r fix 0 and ∞. It will be sufficient to show that any ρ-invariant proper achronal meridian Λ necessarily contains the four points $(0, 0)$, $(0, \infty)$, $(\infty, 0)$, (∞, ∞). Indeed by Lemma 15.2.2.2.1 this will imply that Λ is either Λ_1 or Λ_2.

We claim that Λ must be contained in the union of leaves

$$X = \lambda_0 \cup \lambda_\infty \cup \mu_0 \cup \mu_\infty \,.$$

First let us show how to conclude assuming the claim. Notice that the leaves λ_j and μ_i meet at points $p_{i,j} = (i, j)$ for $i, j = 0, \infty$. If Λ is an achronal meridian contained in X, then it must be a concatenation of arcs on the leaves λ_0, λ_∞, μ_0, μ_∞ with end-points in $\{p_{i,j} \mid i, j = 0, \infty\}$. Notice that

- If Λ contains an arc on λ_j (resp. μ_i) then it contains both $p_{0,j}$, and $p_{\infty,j}$ (resp. $p_{i,0}$ and $p_{i,\infty}$).
- If p_{ij} is contained in Λ, then Λ contains an arc on both λ_i and μ_j (otherwise Λ should contain a leaf).

In particular we easily deduce that Λ must contain all points $p_{i,j}$ and we conclude.

In order to prove the claim we will check that no point in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \setminus X$ has an achronal orbit. Notice that $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \setminus X$ has the following four connected components:

$$U_{+,+} = \mathbb{R}_+ \times \mathbb{R}_+,\; U_{+,-} = \mathbb{R}_+ \times \mathbb{R}_-,\; U_{-,+} = \mathbb{R}_- \times \mathbb{R}_+,\; U_{-,-} = \mathbb{R}_- \times \mathbb{R}_- \,.$$

Each of these components is preserved by ρ. Let us focus on $U_{+,+}$. Using the notation of Proposition 15.2.2.3.1 consider the diffeomorphism

$$\Phi_{+,+} : \mathbb{R}^2 \to U_{+,+} \qquad \Phi_{+,+}(x, y) = (e^x, e^y) ,$$

which is conformal, similarly to the last part of the proof of Proposition 15.2.2.3.1. Let $\widehat{G} = \mathrm{Stab}_+(\mathcal{L})$ be the stabilizer of the geodesic $\mathcal{L} = L_{\ell,\ell}$ preserving an orientation, as in (15.2.7), where ℓ is the oriented geodesic of $\mathbb{H}^2$ with endpoints 0 and ∞. Namely $\widehat{G}$ is the subgroup of $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$ of pairs of hyperbolic transformations with fixed points at $0, \infty$. Then $\Phi_{+,+}$ conjugates G_∞ and $\widehat{G}$. As in Proposition 15.2.2.3.1, we deduce that $\Phi_{+,+}^{-1} \circ \rho(\pi_1(M)) \circ \Phi_{+,+}$ is a lattice in $G_\infty = \mathbb{R}^2$ and therefore the action cannot have achronal orbits in $\mathbb{R}^2$. Since $\Phi_{+,+}$ is conformal, then the action of $\rho(\pi_1(M))$ cannot have achronal orbits in $U_{+,+}$. The proof for the other connected components $U_{\pm,\pm}$ is completely analogous. □

A consequence of Proposition 15.2.2.3.2 is the following. Recall from Sect. 15.2.2.2 that $\mathscr{T}$ denotes a lightlike tetrahedron whose boundary in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is a proper achronal meridian consisting of the concatenation of four lightlike segments. In Lemma 15.2.2.2.5 we showed that $\widehat{G} = \mathrm{Stab}_+(\mathcal{L})$, which is the orientation-preserving, time-preserving stabilizer of $\mathscr{T}$, is isomorphic to the semi-direct product $\mathbb{R}^2 \rtimes \mathbb{Z}/2\mathbb{Z}$.

Corollary 15.2.2.3.3 *Any MGH spacetime of genus one is isometric to a quotient of $\mathscr{T}$ by a subgroup of $\widehat{G}$ acting freely and properly discontinuously on $\mathscr{T}$.*

Proof By Proposition 15.2.2.1.3, any MGH spacetime M of genus one is isometric to the quotient of the invisible domain of a proper achronal meridian invariant under the action of $\rho(\pi_1(T^2))$, where $\rho : \pi_1(T^2) \to \mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$ is the holonomy representation. By Proposition 15.2.2.3.1, ρ maps every non-trivial element to a pair of hyperbolic transformations, and by Proposition 15.2.2.3.2 there are exactly two proper achronal meridians invariant under such a ρ, namely those described in Lemma 15.2.2.2.1. However, by Remark 15.2.2.2.3, there is an orientation-preserving, time-preserving isometry of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ sending one invariant proper achronal meridian to the other. Hence, up to composing with an isometry, we see that M is isometric to a quotient of $\mathscr{T}$, which is the invisible domain of the proper achronal meridian Λ_0 as in Proposition 15.2.2.2.2. □

Let us conclude this section by a discussion on the classification of MGH spacetimes of genus one. For this purpose, we introduce the *deformation space*

$$\mathcal{MGH}(T^2) = \{g \text{ MGH AdS metric on } T^2 \times \mathbb{R}\}/\mathrm{Diff}_0(T^2 \times \mathbb{R}) ,$$

where the group of diffeomorphisms isotopic to the identity acts by pull-back. It is a well-known fact from the theory of (G, X)-structures that the holonomy map, which is well-defined with image in the space of representations of the fundamental group into G up to conjugacy (in this case $G = \mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$), descends to the quotient $\mathcal{MGH}(T^2)$.

Now, Corollary 15.2.2.3.3 tells us that MGH spacetimes of genus 1 are determined by the holonomy representations of $\mathbb{Z}^2$ which take value in $\widehat{G}$ and act freely and properly discontinuously on $\mathscr{T}$.

Two MGH spacetimes $\mathscr{T}/\rho_1(\mathbb{Z}^2)$ and $\mathscr{T}/\rho_2(\mathbb{Z}^2)$ represent the same point in $\mathcal{MGH}(T^2)$ if and only if ρ_1 and ρ_2 are conjugate in $\mathrm{Isom}(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1})$, but in fact in this case they are necessarily conjugate in $\widehat{G}$. Hence the deformation space $\mathcal{MGH}(T^2)$ is identified to the space of $\mathbb{Z}^2$-representations in $\widehat{G}$ acting freely and properly discontinuously on $\mathscr{T}$ up to conjugacy in $\widehat{G}$.

By the proof of Proposition 15.2.2.2.6 we see that $\rho(\pi_1(T^2))$ acts freely and properly discontinuously on $\mathscr{T}$ if and only if, under the isomorphism between $\widehat{G}$ and the semi-direct product $\mathbb{R}^2 \rtimes \mathbb{Z}/2\mathbb{Z}$, its acts freely and properly discontinuously on $\mathbb{R}^2$. Under this isomorphism, conjugacy by elements in the normal subgroup $\mathbb{R}^2$ do not change ρ, while conjugacy by the generator of $\mathbb{Z}/2\mathbb{Z}$ acts by minus the identity. In conclusion, we have the following classification result:

Theorem 15.2.2.3.4 *The deformation space $\mathcal{MGH}(T^2)$ is homeomorphic to the space of discrete and faithful representations of $\pi_1(T^2)$ into $\mathbb{R}^2$ up to sign.*

As a final comment, the space of discrete and faithful representations of $\pi_1(T^2)$ into $\mathbb{R}^2$ coincides with the space of translation structures on the torus. Since they are considered up to sign change, $\mathcal{MGH}(T^2)$ is identified to the deformation space of *semi-translation structures* on the torus.

15.2.2.4 Genus $r \geq 2$: Examples

Let us now consider Σ_r an oriented surface of genus $r \geq 2$. Let us recall the definition of Fuchsian representations.

Definition 15.2.2.4.1 A representation $\rho : \pi_1(S) \to \mathrm{PSL}(2,\mathbb{R})$ is *positive Fuchsian* if there is a ρ-equivariant orientation-preserving homeomorphism $\delta : \widetilde{\Sigma}_r \to \mathbb{H}^2$.

The definition is invariant under conjugation in $\mathrm{PSL}(2,\mathbb{R}) \cong \mathrm{Isom}_0(\mathbb{H}^2)$, but not under conjugation in $\mathrm{Isom}(\mathbb{H}^2)$. By a celebrated result by Goldman [71], a representation ρ is positive Fuchsian if and only if the associated flat $\mathbb{R}\mathrm{P}^1$ bundle E_ρ, constructed as the quotient of $\widetilde{\Sigma}_r \times \mathbb{R}P^1$ by the diagonal action of $\pi_1(S)$ given by the obvious action by deck transformation on the first factor, and by ρ on the second factor, has Euler class $2 - 2r$. This is also equivalent to the existence of an orientation-preserving fiber bundle isomorphism between E_ρ and the unit tangent bundle of Σ_r.

The following classical fact in Teichmüller theory, see for instance [67], is essential for the construction of MGH spacetimes of genus $r \geq 2$.

Lemma 15.2.2.4.2 *Given two positive Fuchsian representations $\rho_l, \rho_r : \pi_1(\Sigma_r) \to \mathrm{PSL}(2,\mathbb{R})$, any (ρ_l, ρ_r)-equivariant orientation-preserving homeomorphism of $\mathbb{H}^2$, which exist as a consequence of Definition 15.2.2.4.1, extends continuously to an orientation-preserving homeomorphism of $\mathbb{H}^2 \cup \mathbb{R}\mathrm{P}^1$. Moreover, its extension $\varphi : \mathbb{R}\mathrm{P}^1 \to \mathbb{R}\mathrm{P}^1$ is the unique (ρ_l, ρ_r)-equivariant orientation preserving homeomorphism of $\mathbb{R}\mathrm{P}^1$.*

By (ρ_l, ρ_r)-equivariance of φ we mean the condition that for every $\gamma \in \pi_1(S)$:

$$\varphi \circ \rho_l(\gamma) = \rho_r(\gamma) \circ \varphi \,. \tag{15.2.8}$$

Now let $\rho_l, \rho_r : \pi_1(\Sigma_r) \to \mathrm{PSL}(2,\mathbb{R})$ be two positive Fuchsian representations. We will consider the representation

$$\rho = (\rho_l, \rho_r) : \pi_1(S) \to \mathrm{Isom}_0(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}) \cong \mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R}) \cong \mathrm{Isom}(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}) \,.$$

Definition 15.2.2.4.3 Given a pair of positive Fuchsian representations $\rho_l, \rho_r : \pi_1(\Sigma_r) \to \mathrm{PSL}(2,\mathbb{R})$, we define $\Lambda(\rho)$ to be the graph in $\mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$ of the unique (ρ_l, ρ_r)-equivariant orientation-preserving homeomorphism of $\mathbb{R}\mathrm{P}^1$, and $\Omega_\rho := \Omega(\Lambda(\rho))$ its invisible domain in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$.

Using the above construction, we can build examples of MGH spacetimes having holonomy any $\rho = (\rho_l, \rho_r)$ of this form.

Proposition 15.2.2.4.4 *The domain Ω_ρ is invariant under the isometric action of $\pi_1(\Sigma_r)$ on $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ induced by ρ. Moreover $\pi_1(\Sigma_r)$ acts freely and properly discontinuously on Ω_ρ and the quotient is a MGH spacetime of genus r and holonomy ρ.*

Proof By the definition of φ and the action of $\mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$ on it is clear that $\Lambda(\rho)$ is invariant by the action of $(\rho_l(\gamma), \rho_r(\gamma))$, for every $\gamma \in \pi_1(\Sigma_r)$. Recalling from Corollary 15.2.1.5.4 that Ω_ρ is the set of elements $x \in \mathrm{PSL}(2,\mathbb{R})$ such that $x \circ \varphi$ have no fixed point on $\mathbb{R}\mathrm{P}^1$, the invariance of Ω_ρ also follows immediately: indeed

$$(\rho_l(\gamma) \circ x \circ \rho_r(\gamma)^{-1}) \circ \varphi = \rho_l(\gamma) \circ (x \circ \varphi) \circ \rho_l(\gamma)^{-1}$$

acts freely on $\mathbb{R}\mathrm{P}^1$ if $x \circ \varphi$ does.

Let us show that for a compact set K in Ω_ρ, $\rho(\gamma) \cdot K$ stays in a compact region of Ω_ρ only for finitely many $\gamma \in \pi_1(\Sigma_r)$. This will also show that the action is free, since $\pi_1(\Sigma_r)$ has no torsion. For this purpose, take a sequence $x_n \in K$ and a sequence $\gamma_n \in \pi_1(\Sigma_r)$ not definitively constant. We claim that up to a subsequence, $(\rho(\gamma_n) \cdot x_n)$ converges to some $(\xi_+, \varphi(\xi_+))$ in $\Lambda(\rho)$. We will apply the criterion of convergence to $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ given in Lemma 15.1.2.2.2.

Since Fuchsian representations act cocompactly on $\mathbb{H}^2$, the sequence $\rho_l(\gamma_n)$ has no converging subsequences in $\mathrm{PSL}(2,\mathbb{R})$. By a well-known dynamical property of $\mathrm{PSL}(2,\mathbb{R})$ (see for instance [18]), up to taking a subsequence, there exist ξ_-, ξ_+ on

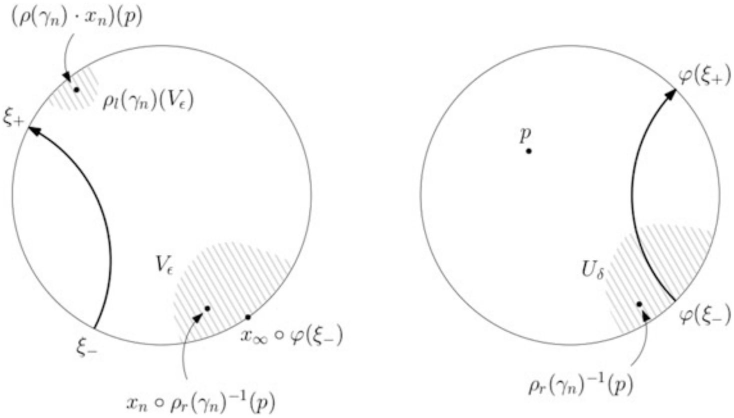


Fig. 15.12 The proof of Proposition 15.2.2.4.4, and in particular the fact that the image of $p \in \mathbb{H}^2$ under $\rho(\gamma_n)\cdot x_n = \rho_l(\gamma_n)\circ x_n \circ \rho_r(\gamma_n)^{-1}$ converges to ξ_+ as $n \to +\infty$. A completely analogous argument shows that the image of p under the inverse converges to $\varphi(\xi_+)$

$\mathbb{R}P^1$ such that $\rho_l(\gamma_n)^{\pm 1}(\xi) \to \xi_\pm$ for all $\xi \neq \xi_\mp$ and that the convergence is uniform on compact sets of $(\mathbb{H}^2 \cup \mathbb{R}P^1) \setminus \{\xi_\mp\}$. By the equivariance (15.2.8), the same holds for $\rho_r(\gamma_n)$ where now $\xi_\pm$ are replaced by $\varphi(\xi_\pm)$.

To apply the criterion of Lemma 15.1.2.2.2, pick any $p \in \mathbb{H}^2$, and recall that $\rho(\gamma_n)\cdot x_n = \rho_l(\gamma_n)\circ x_n \circ \rho_r(\gamma_n)^{-1}$. By the dynamical property above, for any $\delta > 0$ one can find n_0 such that $\rho_r(\gamma_n)^{-1}(p)$ is in the δ-neighborhood of $\varphi(\xi_-)$ (for the Euclidean metric on the closed disc), say U_δ. Since x_n lies in a compact region of Ω_ρ, we can assume that it converges to $x_\infty \in \Omega_\rho$, hence $x_\infty \circ \varphi$ has no fixed point, and in particular $x_\infty \circ \varphi(\xi_-) \neq \xi_-$.

Up to taking δ sufficiently small and n_0 large, $x_n(U_\delta)$ lies in a neighborhood V_ϵ of $x_\infty \circ \varphi(\xi_-)$ such that the closure of V_ϵ is disjoint from ξ_-. By construction $x_n \circ \rho_r(\gamma_n)^{-1}(p) \in V_\epsilon$ and by the uniform convergence on compact sets of the complement of ξ_-, $\rho_l(\gamma_n)\circ x_n \circ \rho_r(\gamma_n)^{-1}(p)$ converges to ξ_+ for n large. The very same argument then shows that $(\rho(\gamma_n)\cdot x_n)^{-1}(p) = \rho_r(\gamma_n)\circ x_n^{-1} \circ \rho_l(\gamma_n)^{-1}(p)$ converges to $\varphi(\xi_+)$. This concludes the claim. See Fig. 15.12.

Finally, the past and future boundary components $\partial_\pm C(\Lambda(\rho))$ are contained in Ω_ρ, since $\Lambda(\rho)$ is the graph of an orientation-preserving homeomorphism (see Remark 15.2.1.6.5). Hence they are ρ-invariant properly embedded Cauchy surfaces in Ω_ρ and project to Cauchy surfaces of the quotient by the action of $\rho(\pi_1(\Sigma_r))$, which are homeomorphic to Σ_r. This shows that the quotient is a globally hyperbolic spacetime of genus r, which is maximal by Proposition 15.2.2.1.3. □

15.2.2.5 Genus $r \geq 2$: Classification

In this section we will conclude the classification result, by showing essentially that the examples of Proposition 15.2.2.4.4 are *all* the MGH spacetimes of genus r.

Lemma 15.2.2.5.1 *Let $\rho = (\rho_l, \rho_r)$ be a pair of positive Fuchsian representations, and $\varphi : \mathbb{R}P^1 \to \mathbb{R}P^1$ be the unique (ρ_l, ρ_r)-equivariant orientation-preserving homeomorphism of $\mathbb{R}P^1$. Then $\Lambda(\rho)$ is the unique proper achronal meridian in $\partial\mathbb{A}d\mathbb{S}^{2,1}$ invariant under the action of $\pi_1(\Sigma_r)$ induced by ρ.*

Proof Let Λ be a proper achronal meridian invariant under the action of $\pi_1(\Sigma_r)$. We claim that the intersection $\Lambda \cap \Lambda_\varphi$ is not empty. Once the claim will be showed, the proof is concluded in the following way. If $(\xi_0, \varphi(\xi_0)) \in \Lambda$, then

$$(\rho_l(\gamma) \cdot \xi_0, \varphi(\rho_l(\gamma) \cdot \xi_0)) = (\rho_l(\gamma) \cdot \xi_0, \rho_r(\gamma) \cdot \varphi(\xi_0)) \in \Lambda \ .$$

However the $\rho_l(\pi_1(\Sigma_r))$-orbit of ξ_0 is dense in $\mathbb{R}P^1$, hence we deduce that Λ contains Λ_φ. But both Λ_φ and Λ are homeomorphic to S^1, which necessarily implies $\Lambda_\varphi = \Lambda$.

Let us prove the claim. Let γ be a non-trivial element in $\pi_1(\Sigma_r)$. It is known that $\rho_l(\gamma)$ and $\rho_r(\gamma)$ are necessarily hyperbolic elements in $\mathrm{PSL}(2, \mathbb{R})$, hence we denote by $\xi_l^+(\gamma)$, and $\xi_r^+(\gamma)$ their attractive fixed points respectively. Notice that $\xi_r^+(\gamma) = \varphi(\xi_l^+(\gamma))$, hence

$$(\xi_l^+(\gamma), \xi_r^+(\gamma)) \in \Lambda_\varphi \ .$$

By homological reasons the curve Λ must meet the leaf of the left ruling of $\partial\mathbb{A}d\mathbb{S}^{2,1}$:

$$\lambda_{\xi_r^+(\gamma)} = \{(\eta, \xi_r^+(\gamma)) \mid \eta \in \mathbb{R}P^1\} \ .$$

That is, there exists $\eta_0 \in \mathbb{R}P^1$ such that $(\eta_0, \xi_r^+(\gamma))$ lies in Λ. But then $(\rho_l(\gamma)^k \eta_0, \xi_r^+(\gamma))$ lies in Λ for $k > 0$. If $\eta_0 \neq \xi_l^-(\gamma)$ we can pass to the limit on k and deduce that $(\xi_l^+(\gamma), \xi_r^+(\gamma))$ lies in Λ.

So far, the choice of γ was arbitrary. To conclude, assume now by contradiction that for every $\gamma \in \pi_1(\Sigma_r)$ the point $(\xi_l^-(\gamma), \xi_r^+(\gamma))$ lies in Λ. Take $\alpha, \beta \in \pi_1(S)$ so that the axes of $\rho_l(\alpha)$ and $\rho_l(\beta)$ do not intersect. We may assume that the cyclic order of end-points of those axes is

$$\xi_l^+(\alpha) < \xi_l^+(\beta) < \xi_l^-(\beta) < \xi_l^-(\alpha) \ . \tag{15.2.9}$$

Since $\xi_r^\pm(\alpha) = \varphi(\xi_l^\pm(\alpha))$ and $\xi_r^\pm(\beta) = \varphi(\xi_l^\pm(\beta))$, we also have that

$$\xi_r^+(\alpha) < \xi_r^+(\beta) < \xi_r^-(\beta) < \xi_r^-(\alpha) \ . \tag{15.2.10}$$

On the other hand by assumption (applied to α, β and their inverses) the curve Λ contains $(\xi_l^+(\alpha), \xi_r^-(\alpha))$, $(\xi_l^+(\beta), \xi_r^-(\beta))$, $(\xi_l^-(\beta), \xi_r^+(\beta))$, $(\xi_l^-(\alpha), \xi_r^+(\alpha))$. By achronality of Λ, the cyclic order of the second components must be the same as that of the first components (although not necessarily strict), hence from (15.2.9) we obtain

$$\xi_r^-(\alpha) \leq \xi_r^-(\beta) \leq \xi_r^+(\beta) \leq \xi_r^+(\alpha) \ ,$$

which contradicts (15.2.10). □

Given a pair $\rho = (\rho_l, \rho_r)$ of positive Fuchsian representations of $\pi_1(\Sigma_r)$, let us denote by M_ρ the MGH spacetime $\Omega_\rho/\rho(\pi_1(\Sigma_r))$ of Proposition 15.2.2.4.4.

Corollary 15.2.2.5.2 *For any pair $\rho = (\rho_l, \rho_r)$ of positive Fuchsian representations of $\pi_1(\Sigma_r)$, M_ρ is the unique MGH spacetime with holonomy ρ.*

The last step for the classification result is that the left and right holonomies are necessarily positive Fuchsian.

Proposition 15.2.2.5.3 *Let M be an oriented, time-oriented, globally hyperbolic spacetime of genus $r \geq 2$ and let us endow a Cauchy surface Σ with the orientation induced by the future normal vector. Then the left and right components of the holonomy $\rho = (\rho_l, \rho_r) : \pi_1(\Sigma) \to \mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$ are positive Fuchsian representations.*

In the statement, we refer to the holonomy ρ with respect to an orientation-preserving developing map. Therefore ρ is well-defined up to conjugacy in $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$.

Proof We will prove that the $\mathbb{R}\mathrm{P}^1$-flat bundles with holonomy ρ_l and ρ_r are isomorphic to the unit tangent bundle of Σ. For the sake of definiteness, let us focus on ρ_l. We will construct an isomorphism

$$\Phi_l : T^1\widetilde{\Sigma} \to \widetilde{\Sigma} \times \mathbb{R}\mathrm{P}^1$$

equivariant with respect to the action on $T^1\widetilde{\Sigma}$ by the actions by deck transformation, and the diagonal action given by ρ_l on $\widetilde{\Sigma} \times \mathbb{R}\mathrm{P}^1$.

The map Φ_l is defined in the following way. For an element $(x, v) \in T^1\widetilde{\Sigma}$, let

$$\xi(x, v) = (\xi^l(x, v), \xi^r(x, v)) \in \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$$

be the end-point of the spacelike geodesic ray $\exp_x(tv)$ in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, for positive t. Then we define $\Phi_l(x, v) = (x, \xi^l(x, v))$. This map is clearly continuous, proper, equivariant and fiber preserving.

In order to prove that it is bijective it is sufficient to notice that for any $x \in \widetilde{\Sigma}$ the map $\xi_x : T_x^1(\widetilde{\Sigma}) \to \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$ is an embedding with image the boundary of the totally geodesic plane tangent to $\widetilde{\Sigma}$ at x. This boundary is the graph of an orientation-preserving map of $\mathbb{R}\mathrm{P}^1$, so the projection $v \to \xi^l(x, v)$

is bijective. Moreover, by our choice of the orientation on Σ, the orientation on $T^1_x\widetilde{\Sigma}$ corresponds to the orientation induced on $\xi_x(T^1_x\widetilde{\Sigma})$ as graph of an orientation-preserving homeomorphism. The proof for ρ_r is completely analogous. □

We conclude by stating the classification result. Let us denote the *deformation space* of MGH spacetimes of genus r by:

$$\mathcal{MGH}(\Sigma_r) = \{g \text{ MGH AdS metric on } \Sigma_r \times \mathbb{R}\}/\mathrm{Diff}_0(\Sigma_r \times \mathbb{R}) ,$$

where the group of diffeomorphisms isotopic to the identity acts by pull-back. Again the holonomy map takes value in the space of representations of $\pi_1(\Sigma_r)$ into $\mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$ and is well-defined on the quotient $\mathcal{MGH}(\Sigma_r)$.

By Proposition 15.2.2.5.3, the left and right components of the holonomy of elements of $\mathcal{MGH}(\Sigma_r)$ are positive Fuchsian representations. The space of these representations up to conjugacy is identified with the Teichmüller space of Σ_r by the aforementioned work of Goldman [71]:

$$\mathcal{T}(\Sigma_r) \cong \{\rho : \pi_1(\Sigma_r) \to \mathrm{PSL}(2,\mathbb{R}) \text{ positive Fuchsian representations}\}/\mathrm{PSL}(2,\mathbb{R}) .$$

Therefore the holonomy map can be considered as a map from $\mathcal{MGH}(\Sigma_r)$ with values in $\mathcal{T}(\Sigma_r) \times \mathcal{T}(\Sigma_r)$. We can summarize Proposition 15.2.2.4.4 and Corollary 15.2.2.5.2 with the following theorem of Mess.

Theorem 15.2.2.5.4 *The holonomy map*

$$\rho : \mathcal{MGH}(\Sigma_r) \to \mathcal{T}(\Sigma_r) \times \mathcal{T}(\Sigma_r)$$

is a homeomorphism.

15.2.3 Gauss Map of Spacelike Surfaces

In this section we will introduce the *Gauss map* associated to a spacelike surface in Anti-de Sitter space, study its properties, and deduce some results which will further highlight the deep relation of Anti-de Sitter geometry with Teichmüller theory and hyperbolic surfaces.

15.2.3.1 Spacelike Surfaces and Immersion Data

Let us start by recalling some generalities of (immersed) spacelike surfaces in Anti-de Sitter geometry. For the moment, we shall assume that all our immersed surfaces are of class C^1.

Let us therefore assume that $\sigma : \Sigma \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is a C^1 immersion, and recall that σ is *spacelike* if the pull-back $\sigma^* g_{\mathbb{A}\mathrm{d}\mathbb{S}}$ of the ambient Lorentzian metric $g_{\mathbb{A}\mathrm{d}\mathbb{S}}$

is positive definite for every point of Σ. In this case the Riemannian metric $I := \sigma^* g_{\mathbb{A}\mathrm{d}\mathbb{S}}$ is called *first fundamental form* of σ.

The tangent bundle TS is then naturally identified to a subbundle of the pull-back bundle $\sigma^*(TM)$ by means of the differential $d\sigma$. The normal bundle N_σ of σ is defined as the $g_{\mathbb{A}\mathrm{d}\mathbb{S}}$-orthogonal complement of TS in $\sigma^*(TM)$, and the restriction of $g_{\mathbb{A}\mathrm{d}\mathbb{S}}$ to N_σ is negative definite. Using the decomposition

$$\sigma^*(TM) = TS \oplus N_\sigma \,,$$

the pull-back of the ambient Levi-Civita connection ∇, restricted to sections tangent to S splits as the sum of the Levi Civita connection ∇^I of the first fundamental form I, and a symmetric 2-form with value in N_σ. As a consequence of time-orientability of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, the normal bundle admits a natural trivialization, which is the same as a choice of a continuous unit normal vector field for σ. We will denote by $\nu : S \to N_\sigma$ the future-directed unit normal vector field, and consider the decomposition for all vector field X, Y tangent to S:

$$\nabla_V W = \nabla^I_V W + I\!I(V, W)\nu \,,$$

where the symmetric $(2, 0)$-tensor $I\!I$ is called *second fundamental form*. It will be convenient to consider the I-symmetric $(1, 1)$-tensor $B \in (TS)^* \otimes TS$ defined by $I\!I(V, W) = I(B(V), W)$, which is called *shape operator* of σ. Similarly to the Riemannian case, it turns out that $\sigma_*(B(v)) = \nabla_v \nu$.

The first and the second fundamental form of an immersion σ satisfy constraint equations, known as the *Gauss–Codazzi equations*. More precisely the Gauss equation consists of the identity

$$K_I = -1 - \det_I I\!I \tag{15.2.11}$$

where K_I is the curvature of I and $\det_I I\!I := \det B$ by definition. On the other hand the Codazzi equation states that $\nabla^I I\!I$ is a totally symmetric 3-form. In other words we have

$$(\nabla^I_V I\!I)(W, U) = (\nabla^I_W I\!I)(V, U) \tag{15.2.12}$$

which sometimes is also written in the equivalent form $d^{\nabla^I} B = 0$ where

$$d^{\nabla^I} B(V, W) = (\nabla^I_V B)(W) - (\nabla^I_W B)(V) = \nabla^I_V(B(W)) - \nabla^I_W(B(V)) - B([V, W]) \,. \tag{15.2.13}$$

The following classical result states that the Gauss–Codazzi equations are the only constraints to be satisfied by the first and second fundamental forms.

Theorem 15.2.3.1.1 (Fundamental Theorem of Immersed Surfaces) *Let S be a simply connected surface, let I be a Riemannian metric on S and $I\!I$ be a symmetric*

$(2, 0)$-*tensor on* S. *If* I *and* $I\!I$ *satisfy the Gauss–Codazzi equations* (15.2.11) *and* (15.2.12)*, then there exists a spacelike immersion* $\sigma : S \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ *having* I *and* $I\!I$ *as first and second fundamental form. Moreover if* σ *and* σ' *are two such immersions, then there exists a time-preserving isometry* f *such that* $\sigma' = f \circ \sigma$.

15.2.3.2 Germs of Spacelike Immersions in AdS Manifolds

Let us now consider the case of an oriented surface Σ, not necessarily simply connected. Given a spacelike immersion $\sigma : \Sigma \to (M, g)$, where (M, g) is an oriented Anti-de Sitter manifold, we can associate to σ the pair $(I, I\!I)$ of first and second fundamental form exactly as in the previous section, where $I\!I$ is computed with respect to the future unit normal vector ν of σ. Moreover, in this section we shall always assume that the orientation of Σ and ν are compatible with the orientation of M.

The pair $(I, I\!I)$ only depends on the *germ* of σ, which we introduce in the following definition:

Definition 15.2.3.2.1 A *germ* of a spacelike immersion of Σ into an Anti-de Sitter three-manifold is an equivalence class of spacelike immersions $\sigma : \Sigma \to (M, g)$, where the time-oriented Lorentzian manifold (M, g) has constant sectional curvature -1, by the following relation: $\sigma : \Sigma \to (M, g)$ and $\sigma' : \Sigma \to (M', g')$ are equivalent if there exist open subsets U in M and U' in M' and an orientation-preserving, time-preserving isometry $f : (U, g) \to (U', g')$ such that $\sigma' = f \circ \sigma$.

Observe that in the definition, $\sigma' = f \circ \sigma$ implies that U is an open neighbourhood of the image of σ, and similarly for U'. It is a simple exercise to check that the above definition gives an equivalence relation.

Now, given a pair $(I, I\!I)$ on a surface Σ, one can perform the following construction. If $\pi : \widetilde{\Sigma} \to \Sigma$ is a universal cover, the pair $(\pi^* I, \pi^* I\!I)$ clearly satisfy the Gauss–Codazzi equations on $\widetilde{\Sigma}$, hence by the existence part of Theorem 15.2.3.1.1 there exists a spacelike immersion $\widetilde{\sigma} : \widetilde{\Sigma} \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ having immersion data $(\pi^* I, \pi^* I\!I)$. The uniqueness part of Theorem 15.2.3.1.1 then has two consequences:

- Any two such immersions differ by post-composition by a global isometry of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$.
- Given any such $\widetilde{\sigma}$, there exists a map $\rho : \pi_1(\Sigma) \to \mathrm{Isom}_0(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1})$ such that, for every $\gamma \in \pi_1(\Sigma)$, $f \circ \gamma = \rho(\gamma) \circ f$.

It is easily checked that ρ is in fact a group representation. Moreover changing $\widetilde{\sigma}$ by post-composition with an isometry f has the effect of conjugating ρ by f. The immersion σ can then be extended to an immersion of U, an open neighbourhood of $\Sigma \times \{0\}$ in $\Sigma \times \mathbb{R}$, into $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, by mapping (x, t) to the point $\gamma(t)$ on the timelike geodesic γ such that $\gamma(0) = \sigma(p)$ and $\gamma'(0)$ is the future normal vector of σ at x. We collect here the expression of the Anti-de Sitter metric in such a tubular neighborhood of σ, which is in fact a local computation and will be useful for future reference:

Lemma 15.2.3.2.2 *Given a spacelike immersion $\sigma : \Sigma \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, the pull-back of the ambient metric by means of the map $(p,t) \mapsto \exp_{\sigma(x)}(t\nu(x))$ has the expression:*

$$-dt^2 + \cos^2(t) I + 2\cos(t)\sin(t) I\!I + \sin^2(t) I\!I\!I\ , \qquad (15.2.14)$$

where I, $I\!I$ and $I\!I\!I$ are the first, second and third fundamental form of σ.

Recall that the third fundamental form is defined as $I\!I\!I(\cdot,\cdot) = I(B(\cdot), B(\cdot))$ where B is the shape operator. Conversely observe that, by a simple computation, the immersion data of $x \mapsto (x, 0)$ in (15.2.14) are $(I, I\!I)$.

Proof We may use the quadric model, introduced in Sect. 15.1.1.5. By Eq. (15.1.13), we have $\exp_{\sigma(x)}(t\nu(x)) = \cos(t)\sigma(x) + \sin(t)\nu(x)$. The differential in t gives the vector $-\sin(t)\sigma(x) + \cos(t)\nu(x)$, while the differential in the spatial direction $(V, 0)$ gives the vector $\cos(t)d\sigma_x(v) + \sin(t)d_x\nu(v)$. The two vectors are orthogonal. Recalling that $I(\cdot,\cdot) = \langle d\sigma(\cdot), d\sigma(\cdot)\rangle$ and that the differential of σ identifies $B(v)$ and $\nabla_v \nu$, namely the tangential component of $d\nu(v)$, the expression of the pull-back metric follows immediately. □

Therefore, given a pair $(I, I\!I)$, the expression (15.2.14) provides a Lorentzian metric of constant curvature -1 on an open set U in $\Sigma \times \mathbb{R}$ containing the slice $\Sigma \times \{0\}$, and thus a germ of immersion of Σ into an Anti-de Sitter three-manifold with immersion data $(I, I\!I)$. The conclusion of the above discussion is summarized in the following statement:

Proposition 15.2.3.2.3 *Given a surface Σ, there are natural identifications between the following spaces:*

(1) The space of pairs $(I, I\!I)$ on Σ which are solutions of the Gauss–Codazzi equations.
(2) The space of germs of spacelike immersions of Σ into Anti-de Sitter manifolds.
(3) The space of spacelike immersions of $\widetilde{\Sigma}$ into $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, equivariant with respect to a representation $\rho : \pi_1\Sigma \to \mathrm{Isom}_0(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1})$, up to the action of $\mathrm{Isom}_0(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1})$ by post-composition.

The identifications are equivariant with respect to the actions of $\mathrm{Diff}(\Sigma)$, by pull-back in item (1) and by pre-composition in items (2) and (3).

Let us now consider the case where Σ is a closed surface. By the arguments of the previous section, the equivariant immersion $\widetilde{\sigma}$ in item (3) of Proposition 15.2.3.2.3 is necessarily an embedding, which can be extended to an embedding of $\widetilde{\Sigma} \times \mathbb{R}$ onto a domain of dependence in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. The representation $\rho : \pi_1(\Sigma) \to \mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$ coincides with the holonomy of a maximal globally hyperbolic Anti-de Sitter manifold (M, g) (after identifying $\pi_1(\Sigma)$ with $\pi_1(M)$ using the embedding of Σ into $M \cong \Sigma \times \mathbb{R}$), and therefore ρ consists of a pair of positive Fuchsian representations by Proposition 15.2.2.5.3.

Quite remarkably, the embedding data (I, II) permit to recover explicitly the pair of elements in the space $\mathcal{T}(S) \times \mathcal{T}(S)$ which parameterizes maximal globally hyperbolic Anti-de Sitter manifolds with compact Cauchy surfaces—recall Theorem 15.2.2.5.4. Such an explicit formula is the content of Proposition 15.2.3.3.7 in the next section.

15.2.3.3 Gauss Map and Projections

We are now ready to define the Gauss map for spacelike surfaces in $\mathbb{A}d\mathbb{S}^{2,1}$, see [9]. Recall from Proposition 15.1.2.5.2 that the space of timelike geodesics of $\mathbb{A}d\mathbb{S}^{2,1}$ is naturally identified with $\mathbb{H}^2 \times \mathbb{H}^2$, where the identification maps a geodesic of the form

$$L_{p,q} = \{X \in \mathrm{PSL}(2, \mathbb{R}) \mid X \cdot q = p\}$$

to the pair $(p, q) \in \mathbb{H}^2 \times \mathbb{H}^2$. We still suppose that our spacelike immersions are C^1 here, and will discuss certain cases of weaker regularity in the next section.

Definition 15.2.3.3.1 Let $\sigma : S \to \mathbb{A}d\mathbb{S}^{2,1}$ a spacelike immersion. The *Gauss map* $G_\sigma : S \to \mathbb{H}^2 \times \mathbb{H}^2$ is defined as $G_\sigma(x) = (p, q)$ such that $L_{p,q}$ is the timelike geodesic orthogonal to $\mathrm{Im}(d_x\sigma)$ at $\sigma(x)$.

As a consequence of the equivariance property given in Proposition 15.1.2.5.2, the Gauss map G_σ is natural with respect to the action of the isometry group, meaning that

$$G_{f\circ\sigma} = f \cdot G_\sigma$$

for every $f \in \mathrm{Isom}_0(\mathbb{A}d\mathbb{S}^{2,1}) = \mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$.

Example 15.2.3.3.2 *Recall that in Lemma 15.1.2.5.1 we gave an isometric embedding of $\mathbb{H}^2$ in $\mathbb{A}d\mathbb{S}^{2,1}$ with image the plane $P_{\mathbb{1}}$ dual to the identity. This isometric embedding is defined by sending $p \in \mathbb{H}^2$ to the unique order-two element in* $\mathrm{PSL}(2, \mathbb{R})$ *fixing p, which by definition lies on the geodesic $L_{p,p}$. Moreover the geodesic $L_{p,p}$ is orthogonal to $P_{\mathbb{1}}$. Hence the Gauss map associated to this isometric embedding of $\mathbb{H}^2$ is simply $p \mapsto (p, p)$.*

By construction, the Gauss map of a spacelike immersion σ is invariant by reparametrization, in the sense that $G_{\sigma\circ\phi} = G_\sigma \circ \phi$ for a diffeomorphism $\phi : S' \to S$. Hence it makes sense to talk about the Gauss map of a spacelike surface in $\mathbb{A}d\mathbb{S}^{2,1}$. For example, for the plane $P_{\mathbb{1}}$ dual to the identity as in Example 15.2.3.3.2, the Gauss map of $P_{\mathbb{1}}$ is sends order-two element of $\mathrm{PSL}(2, \mathbb{R})$ to the pair (p, p) where p is its fixed point.

Lemma 15.2.3.3.3 *Given a spacelike immersion $\sigma : S \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ with future unit normal vector field ν, if $\sigma(p) = \mathbb{1}$, then*

$$G_\sigma(p) = G_{P_{\mathbb{1}}}\left(\exp\left(\frac{\pi}{2}\nu(p)\right)\right) . \tag{15.2.15}$$

Proof The proof follows from Example 15.2.3.3.2 and the observation that the geodesic leaving from $\mathbb{1}$ with velocity $\nu(p)$ meets orthogonally $P_{\mathbb{1}}$ at $\exp((\pi/2)\nu(p))$. □

Let us now introduce the map

$$\mathrm{Fix} : T^{1,+}_{\mathbb{1}}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \to \mathbb{H}^2$$

where $T^{1,+}_{\mathbb{1}}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ denotes the hyperboloid of future unit timelike vectors in $T_{\mathbb{1}}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, such that $\mathrm{Fix}(\nu)$ is the fixed point of the one-parameter elliptic group $\{\exp(t\nu) \mid t \in \mathbb{R}\}$. This map is equivariant for the action of $\mathrm{PSL}(2,\mathbb{R})$, which acts on the hyperboloid $T^{1,+}_{\mathbb{1}}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ by the adjoint representation and on $\mathbb{H}^2$ by the obvious action. Since both $T^{1,+}_{\mathbb{1}}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ and $\mathbb{H}^2$ have constant curvature -1, it follows immediately from the equivariance that Fix is an isometry.

In terms of the map Fix, Eq. (15.2.15) reads

$$G_\sigma(p) = (\mathrm{Fix}(\nu(p)), \mathrm{Fix}(\nu(p))) , \tag{15.2.16}$$

provided $\sigma(p) = \mathbb{1}$. Using Lemma 15.2.3.3.3 and the naturality, we get the following description of the Gauss map.

Lemma 15.2.3.3.4 *Given a spacelike immersion $\sigma : S \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ with future unit normal vector field ν,*

$$G_\sigma(p) = \left(\mathrm{Fix}((R_{\sigma(p)^{-1}})_*(\nu(p))), \mathrm{Fix}((L_{\sigma(p)^{-1}})_*(\nu(p)))\right) .$$

Proof Let us first observe that, if $\sigma(p) = \mathbb{1}$, then the equality holds true by Eq. (15.2.16). In the general case, the immersion $\sigma' = (\mathbb{1}, \sigma(p)) \circ \sigma$ has the property that $\sigma'(p) = \mathbb{1}$, and the future normal vector at $\sigma'(p)$ equals $\nu'(p) = (R_{\sigma(p)^{-1}})_*(\nu(p))$. Therefore

$$G_{\sigma'}(p) = \left(\mathrm{Fix}((R_{\sigma(p)^{-1}})_*(\nu(p))), \mathrm{Fix}((R_{\sigma(p)^{-1}})_*(\nu(p)))\right) .$$

By the naturality of the Gauss map,

$$\begin{aligned} G_\sigma(p) &= (\mathbb{1}, \sigma(p)^{-1}) \cdot G_{\sigma'}(p) \\ &= \left(\mathrm{Fix}((R_{\sigma(p)^{-1}})_*(\nu(p))), \sigma(p)^{-1} \circ \mathrm{Fix}((R_{\sigma(p)^{-1}})_*(\nu(p)))\right) \\ &= \left(\mathrm{Fix}((R_{\sigma(p)^{-1}})_*(\nu(p))), \mathrm{Fix}((L_{\sigma(p)^{-1}})_*(\nu(p)))\right) , \end{aligned}$$

where in the last line we used the fact that Fix is equivariant with respect to the adjoint action on the hyperboloid $T^{1,+}_{\mathbb{1}}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. □

The components of the Gauss map are called *left* and *right projections*, and will be denoted by $\Pi_l, \Pi_r : S \to \mathbb{H}^2$.

Remark 15.2.3.3.5 Under the identification given by Fix, the left and right projections can be interpreted in the following way. Given $p \in S$, $\Pi_l(p)$ is the parallel transport in $\mathbb{1}$ of the future unit vector $\nu(p)$ at $\sigma(p)$ with respect to the right-invariant connection D^r we introduced in Sect. 15.1.2.3. The right projection $\Pi_r(p)$ is instead obtained by parallel transport with respect to the left-invariant connection.

Remark 15.2.3.3.6 Another interpretation of the Gauss map, which was originally given in the work of Mess, is the following. Given $p \in S$ one can find a unique left isometry $f_l(p)$, and a unique right isometry $f_r(p)$, sending the tangent plane P to the image of σ at $\sigma(p)$ to $P_{\mathbb{1}}$. Indeed the isometries $f_l(p)$ and $f_r(p)$ are simply obtained by left and right multiplication by the inverse of dual point of the tangent plane P, namely $\varsigma(p) = \exp_{\sigma(p)}((\pi/2)\nu(p))$. Using the identification of the dual plane $P_{\mathbb{1}}$ with $\mathbb{H}^2$ provided by Lemma 15.1.2.5.1, $\Pi_l(p)$ and $\Pi_r(p)$ are the image of $\sigma(p)$ under the right and left isometries respectively:

$$\Pi_l(p) = f_r(p) \circ \sigma(p) = (\mathbb{1}, \varsigma(p)) \cdot \sigma(p) \qquad \text{and}$$
$$\Pi_r(p) = f_l(p) \circ \sigma(p) = (\varsigma(p)^{-1}, \mathbb{1}) \cdot \sigma(p)\ .$$

We are now ready to prove the formulae which express the pull-back of the hyperbolic metrics by the left and right projections. When applying these formulae to the embedding data of a surface in an MGH Cauchy compact Anti-de Sitter spacetime (M, g), we obtain a pair of hyperbolic metrics whose isotopy classes give are the parameters of (M, g) in $\mathcal{T}(S) \times \mathcal{T}(S)$. (See also Proposition 15.2.3.2.3 and the following paragraph.)

Proposition 15.2.3.3.7 *Let $\sigma : S \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ be a spacelike immersion, let $\Pi_l, \Pi_r : S \to \mathbb{H}^2$ be the left and right projections and let $g_{\mathbb{H}^2}$ be the hyperbolic metric. Then*

$$\Pi_l^* g_{\mathbb{H}^2} = I((\mathrm{id}-\mathcal{J}B)\cdot, (\mathrm{id}-\mathcal{J}B)\cdot) \qquad \textit{and} \qquad \Pi_r^* g_{\mathbb{H}^2} = I((\mathrm{id}+\mathcal{J}B)\cdot, (\mathrm{id}+\mathcal{J}B)\cdot)\ ,$$

where I is the first fundamental form of σ, $\mathcal{J}$ its associated almost-complex structure, and B the shape operator.

These formulae appeared in [79, Lemma 3.16], and are proved also in [9, Section 6.2]. Here we provide a self-contained proof.

Proof Let us check the formula for the pull-back of Π_r. By Lemma 15.2.3.3.4,

$$\Pi_r(x) = \mathrm{Fix}((L_{\sigma(x)^{-1}})_*(\nu(x)))\ .$$

Since Fix : $T_{\mathbb{1}}^{1,+}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \to \mathbb{H}^2$ is an isometry, $\Pi_r^* g_{\mathbb{H}^2}$ equals the pull-back of the Anti-de Sitter metric through the map $\widehat{\Pi}_r : S \to T_{\mathbb{1}}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ defined by $\widehat{\Pi}_r(x) = (L_{\sigma(x)^{-1}})_*(\nu(x))$.

Let us fix a orthonormal basis of left-invariant vector fields $E_1, \dots, E_n$ on $T\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. Then we can express the unit normal vector as $\nu(x) = \sum_i \nu_i(x)E_i(\sigma(x))$, for some functions $\nu_i : S \to \mathbb{R}$. By Remark 15.2.3.3.5,

$$\widehat{\Pi}_r(x) = \sum_i \nu_i(x)E_i(\mathbb{1}) .$$

By differentiating we obtain

$$d\widehat{\Pi}_r(v) = \sum_i d\nu_i(v)E_i(\mathbb{1}) . \tag{15.2.17}$$

On the other hand, since left-invariant vector fields are parallel for the left-invariant connection D^l, we have

$$D^l_v \nu = \sum_i d\nu_i(v)E_i(\sigma(x)) . \tag{15.2.18}$$

The identities (15.2.17) and (15.2.18) together show that $\Pi_r^* g_{\mathbb{H}^2}(v, w)$, which equals the Anti-de Sitter metric $g_{\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}}$ at the identity evaluated on the tangent vectors $d\widehat{\Pi}_r(v)$ and $d\widehat{\Pi}_r(w)$, equals $g_{\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}}(D^l_v \nu, D^l_w \nu)$.

Using Eq. (15.1.26) and Lemma 15.1.2.4.1,

$$D^l_v \nu = \nabla_v \nu + v \boxtimes \nu = B(v) - \nu \boxtimes v = (B - \mathcal{J})v .$$

We conclude that

$$\Pi_r^* g_{\mathbb{H}^2}(v, w) = I((B - \mathcal{J})v, (B - \mathcal{J})w) = I((\mathrm{id} + \mathcal{J}B)v, (\mathrm{id} + \mathcal{J}B)w)$$

as claimed. The proof for the left projection is exactly the same, using right-invariant vector fields and the right-invariant connection, and one gets a difference in sign when applying Eq. (15.1.26). □

15.2.3.4 Consequences and Comments

We collect here several consequences and remarks around Proposition 15.2.3.3.7.

- A first simple remark is that if σ is a totally geodesic immersion, which means that B vanishes identically, then the projections are local isometries. Even without using Proposition 15.2.3.3.7, we have already observed this fact in Example 15.2.3.3.2 for the totally geodesic plane $P_{\mathbb{1}}$, and it is therefore true by the naturality property for every totally geodesic immersion.

- Proposition 15.2.3.3.7 shows that the differential of the left and right projections essentially has the expression

$$d\sigma_x \circ (B \pm \mathcal{J}) ,$$

up to post-composing with an isometry sending the image of $d\sigma_x$ to a fixed copy of $\mathbb{H}^2$. Since B is I-symmetric, $\mathcal{J} \circ B$ is traceless, and therefore

$$\det(B \pm \mathcal{J}) = 1 + \det B = -K_I . \tag{15.2.19}$$

This shows that Π_l is a local diffeomorphism at a point x if and only if Π_r is, which is the case if and only if the intrinsic curvature of I at x is different from 0.
- Since the trace of $B \pm \mathcal{J}$ equals 2, the differentials of Π_l and Π_r have either rank 2 or rank 1. (In fact, by (15.2.19), when the differential of Π_l has rank 1, the same holds for the differential of Π_r.) Hence the differential of the Gauss map $G : S \to \mathbb{H}^2 \times \mathbb{H}^2$ is always non-singular. Moreover, we have the following dichotomy: for every point x, either the image of G is locally a graph of a map between (open subsets of) $\mathbb{H}^2$, or it is tangent at $G(x)$ to a maximal flat of $\mathbb{H}^2 \times \mathbb{H}^2$, that is, to the product of two geodesics.
- If an immersed surface has the property that the curvature of the first fundamental form never vanishes, and if moreover Π_l and Π_r are globally injective, then the image of G is the graph of a diffeomorphism F_σ between two subsets of $\mathbb{H}^2$, called the *associated map*. From Eq. (15.2.19), the Jacobians of Π_l and Π_r are equal, hence the associated map is area-preserving. When Π_l and Π_r are only locally injective, but not globally, we still obtain an area-preserving local diffeomorphism F_σ which is now defined between two hyperbolic surfaces, not globally isometric to subsets of $\mathbb{H}^2$.
- More generally, as a consequence of the previous points, the image of G is always a *Lagrangian submanifold* in $\mathbb{H}^2 \times \mathbb{H}^2$ with respect to the symplectic form

$$\Omega = \pi_l^* \omega_{\mathbb{H}^2} - \pi_r^* \omega_{\mathbb{H}^2} , \tag{15.2.20}$$

where $\omega_{\mathbb{H}^2}$ is the hyperbolic area form. This result has been proved in several works with different methods: see [9, 30, 101]. Moreover the Lagrangian condition is *locally* the only obstruction to inverting this construction, that is, to realizing an immersed surfaces in $\mathbb{H}^2 \times \mathbb{H}^2$ locally as the image of the Gauss map of a spacelike immersion in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$.
- Finally, given a spacelike immersion σ, the normal evolution of σ is defined as

$$\sigma_t(x) = \exp_{\sigma(x)}(t\nu(x)) ,$$

where ν is the future unit normal vector field. In general σ_t may fail to be an immersion for $|t|$ large. (We will come back to this point in Sect. 15.3.1.1, in particular Remark 15.3.1.1.3). When it is an immersion, the computation of the metric in Lemma 15.2.3.2.2 shows that the image of σ_t at x is orthogonal to the geodesic $\gamma(t) = \exp_{\sigma(x)}(t\nu(x))$. In other words, the Gauss map of σ_t is equal to the Gauss map of σ. Hence with respect to the previous point, given a spacelike immersion, there is actually a one-parameter family of immersions, which differ from one another by the normal evolution, which have the same Lagrangian submanifold of $\mathbb{H}^2 \times \mathbb{H}^2$ as Gauss map image. This phenomenon can be explained in a more transparent way in terms of the unit tangent bundle, see Sect. 15.2.3.5 below.

Remark 15.2.3.4.1 Given a Riemannian metric I, suppose A a $(1, 1)$-tensor which is I-symmetric and I-Codazzi, namely satisfying $d^{\nabla^I}A = 0$ (recall Eq. (15.2.13) for the definition of the exterior derivative). Then the curvature of the metric $g = I(A\cdot, A\cdot)$ is expressed by the formula

$$K_g = \frac{K_I}{\det(A)} ,$$

see [82]. The tensors $\mathrm{id} \pm \mathcal{J} \circ B$ appearing in Proposition 15.2.3.3.7 are I-Codazzi, since id, $\mathcal{J}$ and B are all I-Codazzi. Then using the Gauss equation and Eq. (15.2.19), one verifies directly that, when non-degenerate, the pull-back metrics of Proposition 15.2.3.3.7 are hyperbolic.

15.2.3.5 Future Unit Tangent Bundle Perspective

The Gauss map of an embedded surface can be described concisely in terms of the *future timelike unit tangent bundle* $T^{1,+}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, namely the bundle whose fiber over $x \in \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is the subset of $T_x\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ consisting of future-directed timelike vectors of square norm -1 (which is therefore a copy of $\mathbb{H}^2$). The total space of $T^{1,+}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ has also a structure of principal $\mathbb{S}^1$-bundle over $\mathbb{H}^2 \times \mathbb{H}^2$: the projection $\pi : T^{1,+}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \to \mathbb{H}^2 \times \mathbb{H}^2$ maps (x, v) to the equivalence class (up to reparametrization) of the timelike geodesic γ with $\gamma(0) = x$ and $\gamma'(0) = v$, and the $\mathbb{S}^1$-action is by the geodesic flow, namely the action of e^{it} maps (x, v) to $(\gamma(2t), \gamma'(2t))$ (recalling that timelike geodesics have length π).

One can then define the Gauss map for any spacelike immersed surface $\sigma : S \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$: this is simply obtained by first lifting σ to a map $\widetilde{\sigma} : S \to T^{1,+}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, where $\widetilde{\sigma} = (\sigma, \nu)$, and then projecting $\widetilde{\sigma}$ to $\mathbb{H}^2 \times \mathbb{H}^2$ by composition with π. Hence one clearly recovers the fact that the Gauss map is invariant by the normal evolution, since normal evolution corresponds to the action of the geodesic flow on $T^{1,+}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$.

We will not pursue this point of view very far here, and we refer to [30] for the interested reader and to [77] for the necessary background. However, it is worth mentioning that there is a natural fiber bundle connection on the principal bundle $\pi : T^{1,+}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \to \mathbb{H}^2 \times \mathbb{H}^2$, which has interesting consequences. To define the connection, recall that the Levi-Civita connection of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ induces a decomposition of $T_{(x,v)}T\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} = H \oplus V$, where V is the tangent space to the fiber, hence naturally isomorphic to $T_x\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, while H consists of vectors tangent to the lifts to $T\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ of geodesics of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, and therefore the differential of the projection $T\mathbb{A}\mathrm{d}\mathbb{S}^{2,1} \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ identifies H with $T_x\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. Then the pseudo-Riemannian Sasaki metric g_S is defined on $T\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ by declaring that H and V are orthogonal, and that g_S restricted to H and V coincides with the metric of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ under the above isomorphisms. It turns out that the Sasaki metric is invariant both under the action of the isometry group $\mathrm{Isom}(\mathbb{A}\mathrm{d}\mathbb{S}^{2,1})$ and by the geodesic flow. Then the connection, which in this case is simply a real-valued 1-form on $T^{1,+}\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, is defined as

$$\omega(\cdot) = g_S(\chi, \cdot) ,$$

where χ is the infinitesimal generator of the geodesic flow: at a point (x, v), the component of χ along V vanishes, while its component along H is v. One can then prove (see [30, Proposition 3.9]) that the curvature of the connection ω, which is defined as $d\omega$ (in general there is an additional term $\omega \wedge \omega$, which vanished automatically here), coincides up to a factor with the pull-back $\pi^*\Omega$, where Ω is the symplectic form of $\mathbb{H}^2 \times \mathbb{H}^2$ defined in (15.2.20). This permits to recover once more the fact that the Lagrangian condition is the only local obstruction to realize a submanifold of $\mathbb{H}^2 \times \mathbb{H}^2$ as the Gauss map image of a spacelike surface in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, a fact we mentioned already in Sect. 15.2.3.4. In [30] this technology is applied to determine a global obstruction to the reversibility of the Gauss map construction for MGH Cauchy compact manifold, namely in presence of the action of a pair of Fuchsian representations $\rho = (\rho_l, \rho_r) : \pi_1\Sigma \to \mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$, in terms of Hamiltonian orbits.

15.2.3.6 Non-smooth Surfaces

The construction of the Gauss map can be extended in the non-smooth setting, for instance for convex spacelike surfaces S in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, which means that every support plane of S is spacelike. Then one defines the set-valued Gauss map as the map sending each point x of S to the set of future unit vectors in $T^{1,+}_x\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ orthogonal to support planes of S at x. Hence the image of a point $x \in S$ is a convex subset of $T_x\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, and it is reduced to a single point if and only if S is differentiable at x. The image of G in $T^{1,+}_x\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is a $C^{1,1}$ surface.

In this pioneering work, Mess highlighted the relation between pleated surfaces and earthquake maps. Recall that, given an achronal meridian $\Lambda \subset \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, the upper and lower boundary components $\partial_\pm C(\Lambda)$ of the convex hull of Λ are a convex and a concave pleated surface, see Proposition 15.2.1.6.6 and Remark 15.2.1.6.7.

In general the pleated surfaces $\partial_\pm C(\Lambda)$ may contain lightlike triangles, which happens exactly in correspondence of a sawtooth, see Remark 15.2.1.6.5 and also Remark 15.3.2.1.3 below. In this case, the Gauss map is of course not defined on these lightlike triangles. The fundamental claim is then the following, see also [9] for more details:

Lemma 15.2.3.6.1 *Given an achronal meridian $\Lambda \subset \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, the image of the Gauss map of $\partial_+ C(\Lambda)$ and $\partial_- C(\Lambda)$, which are defined only on the spacelike parts, are the graph of (left and right respectively) earthquake maps between straight convex sets of $\mathbb{H}^2$.*

More precisely, what happens is that the left and right projections from $\partial_+ C(\Lambda)$ to $\mathbb{H}^2$ are (right and left respectively) earthquake maps with image a straight convex set, and the earthquake measured laminations coincide with the bending measured laminations. Hence the composition $\Pi_r \circ \Pi_l^{-1}$ gives a left earthquake map, which is in fact defined in the complement of the simplicial leaves of the lamination, and its earthquake measured lamination is identified to the bending measured lamination of $\partial_+ C(\Lambda)$. The same holds for $\partial_- C(\Lambda)$, by reversing the roles of left and right.

We will not give a full proof of Lemma 15.2.3.6.1, but in the next section we will explain the case of a surface pleated along a single geodesic, which is the essential step. The full lemma can then be proved by an approximation argument as in [21]. For more details on the part of the statement about straight convex sets, see [32].

Now, when the curve Λ is the graph of an orientation-preserving homeomorphism, one obtains as a result earthquake maps of $\mathbb{H}^2$. When moreover φ is the homeomorphism which conjugates the left and right representations $\rho_l, \rho_r : \pi_1\Sigma \to \mathrm{PSL}(2,\mathbb{R})$ of the holonomy of a MGH Cauchy compact manifold, the naturality of the construction implies that the earthquake map descends to an earthquake map from the left to the right hyperbolic surfaces, namely $\mathbb{H}^2/\rho_l(\pi_1\Sigma)$ and $\mathbb{H}^2/\rho_r(\pi_1\Sigma)$. (By Lemma 15.3.2.2.1 which will be discussed below, one actually sees directly that the earthquake maps of $\mathbb{H}^2$ extends continuously to φ on the boundary at infinity.)

Let us denote, for a measured geodesic lamination μ on Σ, the left and right earthquake maps by

$$E^l_\mu : \mathcal{T}(\Sigma) \to \mathcal{T}(\Sigma) \qquad \text{and} \qquad E^r_\mu : \mathcal{T}(\Sigma) \to \mathcal{T}(\Sigma)\ ,$$

seen as maps of the Teichmüller space of Σ to itself. As a consequence of the previous discussion, and the example to be explained in the next section, Mess recovered Thurston's Earthquake Theorem:

Theorem 15.2.3.6.2 (Earthquake Theorem) *Given any two hyperbolic metrics h, h' on Σ, there exists a unique pair of measured geodesic laminations (μ_l, μ_r) on Σ such that $[h'] = E^l_{\mu_l}([h]) = E^r_{\mu_r}([h])$.*

We mention here that in [61] the Gauss map is considered for convex polyhedral surfaces Σ in MGH Cauchy compact manifolds M. These convex polyhedral surfaces are therefore contained in the complement of the convex core, and their Gauss map will be again set-valued. The bending locus of Σ, which replaces the bending lamination, induces two geodesic graphs on the left and right hyperbolic surfaces with different combinatorics, called *flippable tilings*. Roughly speaking, this is because the image of the vertices of Σ are mapped to hyperbolic polygons under the left and right projections, and the associated map "flips" these polygons with respect to the adjacent components of the complement of the geodesic graph. As a result of this construction, in [61] the authors prove the existence of (many) left and right "flip" maps between any two closed hyperbolic surfaces of the same genus.

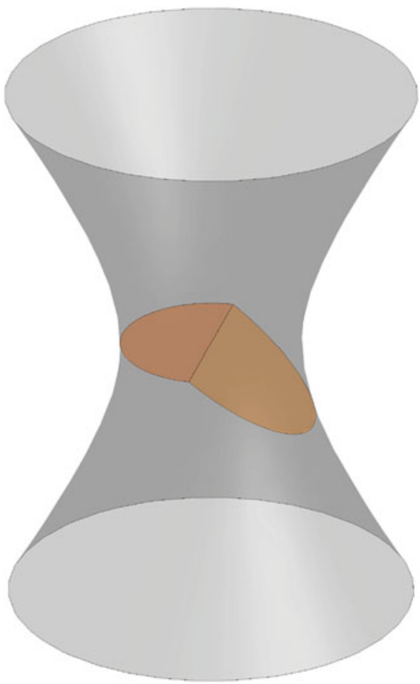
15.2.3.7 The Fundamental Example

Finally, in order to understand how earthquake maps are associated to pleated surfaces, let us now consider the fundamental example. Let S be a piecewise totally geodesic surface consisting of the union of two half-planes in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ meeting along a spacelike geodesic, see Fig. 15.13.

Our aim is to understand the left and right projection for this surface S. Observe that these are well-defined in the complement of the spacelike geodesic which constitutes the bending locus of S. As already observed above (see the first point in Sect. 15.2.3.4), the projections Π_l and Π_r are isometric on each (totally geodesic) connected component of the complement of such bending geodesic in S. Let us call these two components S_1 and S_2.

We may assume that S_1 is contained in the plane $P_{\mathbb{1}}$, composed of order-two elliptic elements in $\mathrm{PSL}(2, \mathbb{R})$. Therefore the bending locus is a spacelike geodesic contained in $P_{\mathbb{1}}$, namely the set of order-two elliptic elements having fixed point in

Fig. 15.13 A pleated surface with bending locus a single geodesic, in an affine chart for $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$



a geodesic ℓ of $\mathbb{H}^2$. From the notation of Sect. 15.1.2.5, it has the form

$$L_{\ell,\ell'} = \{X \in \mathrm{PSL}(2,\mathbb{R}) \mid X \cdot \ell' = \ell \text{ as oriented geodesics}\} ,$$

where ℓ' is the same geodesic but endowed with the opposite orientation.

To understand the behaviour of the projections, the key point is to understand the stabilizer of the spacelike geodesic $L_{\ell,\ell'}$. This is a group isomorphic to $\mathbb{R}^2$, consisting of pairs $(A, B) \in \mathrm{PSL}(2,\mathbb{R}) \times \mathrm{PSL}(2,\mathbb{R})$ where both A and B are hyperbolic transformations preserving ℓ. The stabilizer of $L_{\ell,\ell'}$ fixes setwise also the dual geodesic, namely $L_{\ell,\ell}$ (Proposition 15.1.2.5.5).

In fact, by the definition of dual geodesic (Definition 15.1.2.5.4), the dual point of the spacelike plane S_2 lies in the dual geodesic, and is therefore a hyperbolic transformation σ_0 with axis ℓ. Now, from the discussion of Sect. 15.2.3.3 (see in particular Remark 15.2.3.3.6), the left projection $\Pi_l : S_1 \sqcup S_2 \to \mathbb{H}^2$ is obviously the identity on S_1 (where we identify as usual $P_{\mathbb{1}}$ with a copy of $\mathbb{H}^2$), while on S_2 it is given by multiplication on the right by σ_0^{-1}. Similarly, the right projection is the identity when restricted to S_1, and left multiplication by σ_0^{-1} when restricted to S_2.

In conclusion, the composition $\Pi_r \circ \Pi_l^{-1}$ acts on $P_{\mathbb{1}}$ as the identity on one connected component of the complement of $L_{\ell,\ell'}$, and conjugates by σ_0 on the other connected component—which simply means acting by the hyperbolic transformation σ_0 under the identification between $P_{\mathbb{1}}$ and $\mathbb{H}^2$ (Lemma 15.1.2.5.1). This is exactly an earthquake map with associated earthquake lamination the geodesic ℓ. Since the angle between the spacelike planes containing S_1 and S_2 equals the distance in the dual geodesic $L_{\ell,\ell}$ between the corresponding dual points, we also conclude that the bending measure equals the measure associated with the earthquake map. In short, the bending and earthquake measured laminations are identified.

15.3 Part 3: Further Results

In this part we will explain various results which have been obtained after the work of Mess. Of course we do not aim at an exhaustive treatment here; as mentioned already in the introduction, our choice is to underline mostly the relations between Anti-de Sitter geometry and Teichmüller theory.

15.3.1 More on MGH Cauchy Compact AdS Manifolds

In this section we first consider MGH Cauchy compact manifolds, which have been studied in Chap. 15.2.2, with the purpose of describing more deeply their structure, their deformation space, and the applications to Teichmüller theory of closed hyperbolic surfaces.

15.3.1.1 Foliations

A smoothly embedded spacelike surface S in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ has *constant mean curvature*(CMC) if its mean curvature, namely the trace of the shape operator B, is constant. We will mostly denote by H the constant value of the mean curvature, whose sign actually depends on the choice of a normal unit vector (we will implicitly consider the *future* unit normal vector here). A particular case are *maximal surfaces*, for which $H = \mathrm{tr}\,(B) = 0$. We will implicitly assume that surfaces are spacelike.

Theorem 15.3.1.1.1 ([12]) *Every maximal globally hyperbolic Anti-de Sitter manifold with compact Cauchy surface is uniquely foliated by closed CMC surfaces, where the mean curvature H varies in* $(-\infty, +\infty)$.

In fact, for every H the CMC surface Σ_H is unique, as an application of the maximum principle. Moreover, the *CMC function* $\tau : M \to \mathbb{R}$ which associates to p the unique H such that the CMC surface Σ_H contains p is a *time function*, namely it is strictly increasing along future-directed causal curves.

The embedded surface S has *constant Gaussian curvature*(CGC) if the determinant $\det B$ is constant. In this case the value of the constant is well-defined, and we will consider here the case of positive Gaussian curvature, which will be denoted by $K \in (0, +\infty)$. Hence a CGC surface is either locally convex or locally concave, where the distinction between convex and concave is relative to the time orientation.

Theorem 15.3.1.1.2 ([14]) *Let M be a maximal globally hyperbolic Anti-de Sitter manifold with compact Cauchy surface. Then each connected component of $M \setminus C(M)$ is uniquely foliated by closed CGC surfaces, where the Gaussian curvature K varies in* $(0, +\infty)$.

Again, each surface with constant Gaussian curvature K is unique in its connected component. On each connected component the function which associates to every point the corresponding value of K is again a time function (up to changing the sign if necessary).

There is a remarkable relation between Theorems 15.3.1.1.1 and 15.3.1.1.2, which is given by the normal evolution, and which will appear again in the generalizations of these foliation results to the setting of universal Teichmüller space. Let us introduce this relation here.

Recall that, given a spacelike immersion $\sigma : S \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ with future unit normal vector field ν, the normal evolution of σ is defined as

$$\sigma_t(x) = \exp_{\sigma(x)}(t\nu(x)) \ .$$

(See also the final item in Sect. 15.2.3.4.) The computation of Lemma 15.2.3.2.2 shows that the pull-back of the ambient metric by σ_t has the form:

$$\sigma_t^* g_{\mathbb{A}\mathrm{d}\mathbb{S}} = I((\cos(t)\mathrm{id} + \sin(t)B)\cdot, (\cos(t)\mathrm{id} + \sin(t)B)\cdot) \ , \qquad (15.3.1)$$

where as usual I is the first fundamental form of σ and B its shape operator.

Remark 15.3.1.1.3 The pull-back $\sigma_t^* g_{\mathbb{A}\mathrm{d}\mathbb{S}}$ might in general be degenerate, corresponding to the fact that the differential of σ_t might be singular for some t. Under the identification between the tangent space of the image of σ and σ_t at x, $d\sigma$ and $d\sigma_t$ differ by pre-composition with $\cos(t)\mathrm{id} + \sin(t)B$, whose eigenvalues are $\cos(t) + \sin(t)\lambda_1$ and $\cos(t) + \sin(t)\lambda_2$, λ_i being the principal curvatures.

Under certain conditions relating B and t, however, one can make sure that the map σ_t is an immersion. For instance, by compactness there exists $\epsilon > 0$ (which depends on the norm of B) such that σ_t is an immersion for $t \in (-\epsilon, \epsilon)$. A more significant condition is the following: if σ is a convex immersion (meaning that B is positive definite with respect to the future unit normal vector), then σ_t is an immersion for positive times t, and of course the same holds for a concave immersion and negative times.

The relation between CMC surfaces and CGC surfaces is then contained in the following statement, which also appears in [40].

Proposition 15.3.1.1.4 *Let $\sigma : S \to \mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ be an immersion of constant Gaussian curvature $K > 0$. Then the normal evolution σ_{t_K} on the convex side of σ, for time $t_K = \arctan(K^{1/2})$ is an immersion of constant mean curvature $H = K^{-1/2}(K-1)$.*

Proof By Remark 15.3.1.1.3 the normal evolution σ_t is an immersion. Let us denote by I_t its first fundamental form, and by B_t its shape operator. To check the relation between mean and Gaussian curvature, it is smarter to apply Eq. (15.3.1) to σ_t for negative times and the expression of the curvature given in Remark 15.2.3.4.1. Then one obtains

$$K_I = \frac{K_{I_t}}{\det(\cos(t)\mathrm{id} - \sin(t)B_t)} = -\frac{1 + \det B_t}{\cos^2(t) + \sin^2(t)\det B_t - \cos(t)\sin(t)\mathrm{tr}\, B_t} \ .$$

Hence one can check that K_I (which equals $-1 - \det B$) is constant if and only if $\mathrm{tr}\, B_t = 2/\tan(2t)$, in which case an explicit computation shows that $\tan(t) = 1/\sqrt{K}$. The result follows. □

We remark here that *a priori*, the construction of Proposition 15.3.1.1.4 cannot be reversed, since the normal evolution σ_t obtained from a constant mean curvature immersion σ might be singular at some points. From the proof of Proposition 15.3.1.1.4, one sees that in fact this occurs if and only if the intrinsic curvature of the CMC immersion σ vanishes, which is equivalent to the determinant of the shape operator being equal to -1.

But *a posteriori* the correspondence is indeed bijective when applied to the closed surfaces of constant mean and Gaussian curvature of Theorems 15.3.1.1.1

and 15.3.1.1.2, as a consequence of the uniqueness statements. Indeed if a closed surface of constant mean curvature H in a MGH spacetime (M, g) were not obtained by the "reversed" normal evolution construction with respect to Proposition 15.3.1.1.4 (either in the future or in the past), then applying Proposition 15.3.1.1.4 to an actual surface of the expected constant Gaussian curvature (given by Theorem 15.3.1.1.2) one would find a new surface of constant mean curvature in (M, g), thus contradicting the uniqueness of Theorem 15.3.1.1.1. This means that each surface Σ_H of constant mean curvature H has two equidistant surfaces of constant Gaussian curvature K_+ and K_- (which only depend on H), one convex in the past of Σ_H, the other concave in its future.

15.3.1.2 Minimal Lagrangian Maps and Landslides

Using the results of the previous section, we can recover the existence of special maps between closed hyperbolic surfaces, as maps associated to surfaces with constant mean or Gaussian curvature.

Definition 15.3.1.2.1 Given two hyperbolic metrics h and h' on a surface S, a smooth map $f : (S, h) \to (S, h')$ is *minimal Lagrangian* if its graph is a minimal Lagrangian submanifold of $S \times S$ with respect to the Riemannian product metric $h \oplus h'$ and the symplectic form $\pi_l^*\omega_h - \pi_r^*\omega_{h'}$.

Given a maximal surface Σ_0 in a maximal globally hyperbolic spacetime (M, g), with compact Cauchy surface Σ, we claim that the associated map f_0 is a minimal Lagrangian map from (Σ, h) to (Σ, h'), where h and h' are the quotient metrics induced in $\mathbb{H}^2/\rho_l(\pi_1\Sigma)$ and $\mathbb{H}^2/\rho_l(\pi_1\Sigma)$. We have already discussed the Lagrangian condition, which amounts to f_0 being area-preserving and is always verified by the Gauss map image (Sect. 15.2.3.4). The fact that the graph of f_0 is minimal in $(\Sigma \times \Sigma, h \oplus h')$ is a consequence of Proposition 15.2.3.3.7.

In fact, we shall show that the Gauss map is conformal and harmonic, which implies that its image is a minimal surface. By Proposition 15.2.3.3.7 the pull-back of the product Riemannian metric has the expression $2(I + I\!I\!I)$. When the trace of B vanishes identically, by the Cayley–Hamilton theorem $B^2 + (\det B)\mathrm{id} = 0$, which implies $I\!I\!I = -(\det B)I$ showing conformality. Also, observe that the projections are local diffeomorphisms since by the previous section Σ_0 is obtained as the equidistant surface from a convex surface (of intrinsic curvature -2), the projections are always local diffeomorphisms on convex surfaces (Sect. 15.2.3.4), and its Gauss map coincides with that of Σ_0 (Sect. 15.2.3.4, last item). By a topological argument, the projections are then global diffeomorphisms from Σ_0 to (Σ, h) and (Σ, h').

The harmonicity of the Gauss map is equivalent to the harmonicity of each projection. Since the notion of harmonic map between Riemannian surfaces only depends on the conformal structure on the source, it suffices to show that

$$\Pi_l : (\Sigma_0, I) \to (\Sigma_0, I((\mathrm{id} - \mathcal{J} \circ B)\cdot, (\mathrm{id} - \mathcal{J} \circ B)\cdot)$$

is harmonic, and the same for Π_r. To see this, we can rewrite the target metric as $(I + I\!I\!I) - 2I((\mathcal{J} \circ B)\cdot, \cdot)$. We have used that B is traceless and thus $\mathcal{J} \circ B$ is I-symmetric. Together with the Codazzi property of $\mathcal{J} \circ B$, this also implies that $2I((\mathcal{J} \circ B)\cdot, \cdot)$ is the real part of a holomorphic quadratic differential, in light of the following well-known fact, see [75, 79, 107].

Lemma 15.3.1.2.2 *Given a Riemannian metric g on a surface and a $(1, 1)$-tensor A, A is traceless if and only if $g(A\cdot, \cdot)$ is the real part of a quadratic differential for the conformal structure of g. Moreover the quadratic differential is holomorphic if and only if A is g-Codazzi.*

Therefore Π_l is harmonic. The same proof clearly holds for Π_r. This construction can actually be reversed, in the sense that every minimal Lagrangian map can be realized as the map associated with a maximal surface. This permits to reprove the following theorem of existence and uniqueness of minimal Lagrangian diffeomorphisms in a given isotopy class:

Theorem 15.3.1.2.3 ([82, 100]) *Given a closed surface Σ and two hyperbolic metrics h and h' on Σ, there exists a unique minimal Lagrangian diffeomorphism $f_0 : (\Sigma, h) \to (\Sigma, h')$ isotopic to the identity.*

Let us briefly turn our attention to *landslides*, a natural generalization of minimal Lagrangian maps introduced in [33], which turn out to be precisely the maps associated to constant mean curvature and constant Gaussian curvature surfaces.

Given two hyperbolic metrics h and h' on a surface S, and $\theta \in (0, \pi)$ a θ-landslide $f_\theta : (S, h) \to (S, h')$ is a smooth map which satisfies one of the equivalent conditions (see [29, Section 4.3] for more details and for the equivalence):

1. There exists a smooth $(1, 1)$-tensor A such that (if $\mathcal{J}_h$ is the almost-complex structure of h):
$$f_\theta^* h' = h(((\cos\theta)\mathrm{id} + (\sin\theta)\mathcal{J}_h \circ A)\cdot, ((\cos\theta)\mathrm{id} + (\sin\theta)\mathcal{J}_h \circ A)\cdot)$$
which is positive-definite, h-symmetric, h-Codazzi and has unit determinant.
2. There exist harmonic maps $f : (S, X) \to (S, h)$ and $f' : (S, X) \to (S, h')$, where X is a conformal structure on S, such that $f_\theta = f' \circ f^{-1}$ whose Hopf differentials satisfy
$$\mathrm{Hopf}(f) = e^{2i\theta}\mathrm{Hopf}(f') .$$
Moreover, in the non-compact case, one has to further impose that f and f' have the same holomorphic energy density.

When $\theta = \pi/2$ we recover minimal Lagrangian maps, as the above two conditions are in fact equivalent to Definition 15.3.1.2.1. It then turns out that θ-landslides are precisely the maps associated to surfaces of constant mean curvature $H = 2/\tan\theta$, and therefore also to the two equidistant surfaces of constant Gaussian curvature $\tan^2(\theta/2)$ and $1/\tan^2(\theta/2)$. Hence the following result is a consequence of Theorem 15.3.1.1.1 (or Theorem 15.3.1.1.2):

Theorem 15.3.1.2.4 *Given a closed surface Σ and two hyperbolic metrics h and h' on Σ, and $\theta \in (0, \pi)$, there exists a unique diffeomorphism $f_\theta : (\Sigma, h) \to (\Sigma, h')$ isotopic to the identity which is a θ-landslide.*

It is worth remarking that, when θ approaches 0, then one of the two associated surfaces of constant Gaussian curvature (namely the one having Gaussian curvature $\tan^2(\theta/2)$) approaches a boundary component of the convex core of the ambient manifold (M, g), while the other escapes at infinity in the other end of (M, g). When θ approaches π instead, the roles are switched. Hence the landslide maps f_θ converge to the left and right earthquake maps between (Σ, h) and (Σ, h') as θ diverges in its interval of definition $(0, \pi)$. Morally, θ-landslides are a natural one-parameter family of smooth extensions which interpolate between left earthquake, minimal Lagrangian maps, and right earthquakes.

As a final remark for this section, recall from Sect. 15.2.3.4 that the area-preserving condition for maps from (Σ, h) to (Σ, h'), or more generally the Lagrangian condition for submanifolds of $\Sigma \times \Sigma$ (endowed with the symplectic form induced in the quotient by (15.2.20)), are the only *local* obstructions to reversing the Gauss map construction.

Roughly speaking, this means that any Lagrangian immersion of a simply connected surface in $\mathbb{H}^2 \times \mathbb{H}^2$ can be realized as the Gauss map image of a spacelike immersion in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. However, if the Lagrangian immersion in $\mathbb{H}^2 \times \mathbb{H}^2$ is equivariant with respect to a pair of Fuchsian representations $\rho = (\rho_l, \rho_r)$, the corresponding immersion in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is not necessarily ρ-equivariant.

There is indeed an additional obstruction to reversing the Gauss map construction *globally*. As mentioned already in Sect. 15.2.3.5, this obstruction has been studied in [30] and [101] in terms of an orbit of the group of Hamiltonian symplectomorphisms. This obstruction translates to a cohomological vanishing condition by means of the *flux homomorphism*, a tool from symplectic geometry.

15.3.1.3 Cotangent Bundle of Teichmüller Space

The existence and uniqueness of maximal and constant mean curvature surfaces can be remarkably applied to provide new parameterizations of the deformation space of MGH Cauchy compact Anti-de Sitter three-manifolds. The fundamental observation is that, from Lemma 15.3.1.2.2, given a maximal surface Σ of constant mean curvature H, the second fundamental form of Σ uniquely determines a holomorphic quadratic differential α for the conformal structure associated to I. Hence the conformal class of g, together with α, determines an element of the cotangent bundle $T^*\mathcal{T}(\Sigma)$.

By virtue of the following theorem, the construction can be perfectly reversed. This approach is known in the physics literature as *ADM reduction* based on the article [5], see also [91] and [35, Chapter 2].

Theorem 15.3.1.3.1 ([79, Lemma 3.10]) *Given a Riemannian metric g on a closed surface Σ of genus at least 2 and a holomorphic quadratic differential α for g, there exists a unique germ of maximal spacelike embedding in a MGH Anti-de Sitter manifold having first fundamental form conformal to g and second fundamental form the real part of α.*

The proof roughly consists in applying PDE methods to find a function f such that the Riemannian metric $g' = e^{2f} g$, together with the real part of α, solves the Gauss equation. The Codazzi equation is still automatically satisfied as a consequence of Lemma 15.3.1.2.2, and therefore the pair $(g', \mathrm{Re}(\alpha))$ determines the embedding data of a maximal surface using Theorem 15.2.3.1.1.

Remark 15.3.1.3.2 In [79, Lemma 3.10] the proof is actually given in the more general case of surfaces of constant mean curvature $H \in (-1, 1)$, where now the traceless part of the second fundamental form, namely $I\!I_0 = I\!I - (H/2)I$, is the real part of a holomorphic quadratic differential.

Theorem 15.3.1.3.1, together with Theorem 15.3.1.1.1 and the parameterization of Mess, then provides a homeomorphism $F : T^*\mathcal{T}(S) \to \mathcal{T}(S) \times \mathcal{T}(S)$. It was asked in [16, Question 8.1] whether this map is a symplectomorphism with respect to the natural symplectic structures Ω_{COT} on the cotangent bundle, and $\pi_1^*\Omega_{\mathrm{WP}} - \pi_2^*\Omega_{\mathrm{WP}}$ on $\mathcal{T}(S) \to \mathcal{T}(S)$, where Ω_{WP} is the Weil–Petersson symplectic form on $\mathcal{T}(S)$. The answer is affirmative up to a multiplicative factor.

Theorem 15.3.1.3.3 ([98, Theorem 1.11]) *The map $F : T^*\mathcal{T}(S) \to \mathcal{T}(S) \times \mathcal{T}(S)$ satisfies*

$$F^*(\pi_1^*\Omega_{\mathrm{WP}} - \pi_2^*\Omega_{\mathrm{WP}}) = -2\Omega_{\mathrm{COT}} .$$

15.3.1.4 Volume

In this section we will briefly mention the results of [31] on the volume of MGH Cauchy compact Anti-de Sitter manifolds, which is an interesting invariant on the deformation space $\mathcal{T}(S) \times \mathcal{T}(S)$. Here we are mostly interested in the coarse behaviour of the volume function.

A first foundational fact is that the volume of a MGH Cauchy compact manifold is finite, and is coarsely comparable to the volume of the convex core, up to a constant which only depends on the topology.

Proposition 15.3.1.4.1 ([31, Proposition G]) *Given a MGH Anti-de Sitter manifold M with compact Cauchy surface homeomorphic to Σ, the volume of $M \setminus C(M)$ is at most $\pi^2|\chi(\Sigma)|/2$, with equality if and only if M is Fuchsian.*

Special surfaces in M can then be used to obtain bounds on the volume in terms of certain quantities defined on the deformation space $\mathcal{T}(S) \times \mathcal{T}(S)$, related to energies of L^1-type, by means of their associated maps (for instance, earthquake maps from pleated surfaces, and minimal Lagrangian maps from maximal surfaces).

More concretely, given a C^1-map $f : (\Sigma, h) \to (\Sigma, h')$, where h and h' are hyperbolic metrics on Σ, the L^1-*energy* of f is defined as:

$$E_1(f) = \int_\Sigma ||df||\mathrm{dA}_h \ ,$$

where the norm $||df||$ of the differential is computed with respect to the metrics h and h', and dA_h denotes the area form of h. Unlike the more studied L^2-energy, which is the integral of $||df||^2$, the L^1-energy is not conformally invariant on the source, but fully depends on the Riemannian metrics both on the source and on the target.

Remark 15.3.1.4.2 The L^1-energy can be defined under weaker regularity assumptions on f, and in fact it coincides with the notion of total variation for BV maps. For earthquake maps, the total variation is essentially the length of the earthquake lamination, up to a constant which only depends on the genus. The latter is defined for simple closed curves as the product of the length of the h-geodesic realization of the simple closed curve and its weight, and is then extended to general measured laminations by a continuity argument.

Another important energy of L^1-type is the *holomorphic energy*, which is defined as

$$E_\partial(f) = \int_\Sigma ||\partial f||\mathrm{dA}_h \ ,$$

and was studied in [111]. In particular, it was shown that minimal Lagrangian diffeomorphisms are minima of $E_\partial(f)$ on the space of diffeomorphisms from (Σ, h) to (Σ, h') isotopic to the identity.

Let us now summarize the results of [31], although we omit some of the details here. We say that two quantities f, g, defined here on $\mathcal{T}(\Sigma) \times \mathcal{T}(\Sigma)$, are *coarsely equivalent* if there exist positive constants M_1, M_2, A_1, A_1 (M for multiplicative and A for additive) such that

$$M_1 f - A_1 \le g \le M_2 f + A_2 \ .$$

Theorem 15.3.1.4.3 ([31]) *The following quantities are coarsely equivalent over $\mathcal{T}(\Sigma) \times \mathcal{T}(\Sigma)$, with explicit multiplicative constants (universal) and additive constants (which instead depend only on the genus of Σ):*

- *the volume of the MGH Cauchy compact manifold with left and right metrics h and h';*
- *the volume of its convex core;*

- *the infimum of the* L^1*-energy over* C^1 *maps from* (Σ, h) *to* (Σ, h') *homotopic to the identity;*
- *the length of the left earthquake lamination from* (Σ, h) *to* (Σ, h')*;*
- *the length of the right earthquake lamination from* (Σ, h) *to* (Σ, h')*;*
- *the holomorphic energy density of the minimal Lagrangian diffeomorphism* $f_0 : (\Sigma, h) \to (\Sigma, h')$ *isotopic to the identity.*

It is also worth mentioning that the multiplicative constants from above and below for the length of the left and right earthquake maps all agree, hence one obtains as a corollary that, given two points in Teichmüller space, the length of the left and right earthquake laminations differ by at most a constant which only depends on the topology (explicitly, the constant is $2\pi^2|\chi(\Sigma)|$).

Bounds on the volume are obtained also in terms of the exponential of the Thurston distance between the corresponding points in $\mathcal{T}(\Sigma)$ (either of the two asymmetric distances), from above, and in terms of the exponential of the Weil–Petersson distance, from below, always up to additive and multiplicative constants, both depending on the topology in this case. These results answer to some extent Question 4.1 of [16].

15.3.1.5 Realization of Metrics and Laminations

A consequence of the pleated surface construction of Sect. 15.2.3.6 is that the geometry of a MGH Anti-de Sitter manifold with compact Cauchy surface homeomorphic to Σ is determined by the pair of a hyperbolic metric h on Σ and a measured geodesic lamination. In fact, lifting the measured geodesic lamination μ on the universal cover $\mathbb{H}^2$ of (Σ, h), one can realize a pleated surface having bending lamination μ, which will be equivariant for some representation $\rho : \pi_1\Sigma \to \mathrm{PSL}(2, \mathbb{R}) \times \mathrm{PSL}(2, \mathbb{R})$. Such pleated surface maps in the quotient to a boundary component of the convex core of a MGH Anti-de Sitter manifold.

Theorem 15.3.1.3.1 is, to some extent, a smooth analogue, where pleated surfaces are replaced by maximal surfaces, and the data of a holomorphic quadratic differential is a measure of the curvature of the surface. It is then a natural question to ask, if the geometry of the MGH Cauchy compact manifold is uniquely determined by other pairs, for instance the two bending laminations on the boundary components of the convex core, or the hyperbolic metrics induced, or even by the induced metrics on a pair of smooth surfaces. These questions can therefore provide new parameterizations of the deformation space, and are of course motivated also by their counterparts for quasi-Fuchsian hyperbolic manifolds. We briefly collect here the state-of-the-art on these questions. Most of these questions were asked in [16, Section 3].

Theorem 15.3.1.5.1 ([28, Theorem 1.4]) *Given a compact surface* Σ *and two measured laminations* μ_-, μ_+ *which fill* Σ*, there exists a MGH Anti-de Sitter manifold homeomorphic to* $\Sigma \times \mathbb{R}$ *such that the bending measured laminations of* *the lower and upper boundary components of the convex core are isotopic to* μ_- *and* μ_+ *respectively.*

The statement can be equivalently reformulated as the fact that, given any two measured laminations which fill Σ, the composition of the corresponding left earthquakes, seen as maps from $\mathcal{T}(\Sigma)$ to itself, has a fixed point. See [28, Theorem 1.1]. Observe that, similarly to the quasi-Fuchsian case, the hypothesis that μ_- and μ_+ fill Σ is a necessary condition. Uniqueness of the MGH manifold (up to isotopy of course) is still open.

The following result of Diallo concerns the prescriptions of the induced hyperbolic metric on the boundary of the convex core.

Theorem 15.3.1.5.2 ([56]) *Given a compact surface* Σ *and two hyperbolic metric* h_-, h_+*, there exists a MGH Anti-de Sitter manifold homeomorphic to* $\Sigma \times \mathbb{R}$ *such that induced metrics on the lower and upper boundary components of the convex core are isotopic to* h_- *and* h_+ *respectively.*

The following result is a smooth analogue, for convex/concave surfaces lying outside the convex core.

Theorem 15.3.1.5.3 ([103]) *Given a compact surface* Σ *and two Riemannian metrics* g_-, g_+ *of curvature* < -1*, there exists a MGH Anti-de Sitter manifold homeomorphic to* $\Sigma \times \mathbb{R}$ *containing a convex surface* Σ_- *and concave surface* Σ_+ *whose first fundamental forms are isotopic to* g_- *and* g_+ *respectively.*

Again, uniqueness is not known in general. When the curvature of g_- and g_+ is equal to -2, uniqueness is proved in [79, Theorem 3.21] by proving that g_- and g_+ can be uniquely realized as the metrics on the equidistant surfaces from the maximal surface (in the sense of Proposition 15.3.1.1.4 for $H = 0$, $K = -1$). This gives a new parameterization of the deformation space by $\mathcal{T}(S) \times \mathcal{T}(S)$.

When the two metrics g_- and g_+ coincide, then Theorem 15.3.1.5.3 had been already obtained in [83], by showing that there exists a Fuchsian realization. This has been recently generalized by Labeni in [81], showing that one can realize any locally CAT(-1) distance on Σ as the induced distance on a convex surface in a Fuchsian MGH Anti-de Sitter spacetime. The result of Labeni generalizes also [60], which concerns the realizability of a hyperbolic metric with cone singularities. It is natural to expect that any two locally CAT(-1) distances can be realized, probably uniquely, in a (non-Fuchsian, in general) MGH Anti-de Sitter manifold, but this is still an open question.

15.3.2 Non-closed Surfaces

In this last chapter we will survey other results where the topology of spacelike surfaces is not that of a closed surface. We will first discuss a number of *universal* constructions, meaning that they generalize the situation one sees in the universal covering of a MGH Cauchy compact Anti-de Sitter manifold, which was explained in Chap. 15.2.2. This will have applications for the theory of universal Teichmüller space. Then we briefly discuss the state-of-the-art for manifolds with conical singularities of timelike type, called *particles* and corresponding to Cauchy surfaces with cone points, and the so-called *multi-black holes* which instead correspond to surfaces with boundary.

15.3.2.1 Foliations with Asymptotic Boundary

Recall from Chap. 15.2.2 and in particular Proposition 15.2.2.1.3 that, given a MGH Cauchy compact manifold M, any Cauchy surface in M lifts to a spacelike embedded surface in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ having asymptotic boundary a proper achronal meridian Λ which is the graph of the unique homeomorphism conjugating the left and right representations of $\pi_1\Sigma$ in $\mathrm{PSL}(2,\mathbb{R})$.

Some of the constructions we discussed above can be generalized to the setting of any proper achronal meridian. Recall that $\Omega(\Lambda)$ denotes the invisible domain of Λ, see Sect. 15.2.1.2.

Theorem 15.3.2.1.1 *Let $\Lambda \subset \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ be the graph of an orientation-preserving homeomorphism. Then there exists a unique foliation of the invisible domain $\Omega(\Lambda)$ by properly embedded spacelike surfaces of constant mean curvature H, as $H \in (-\infty, +\infty)$.*

The CMC function associated to the foliation is a time function, similarly to the compact case (see Theorem 15.3.1.1.1 and the following discussion).

In the literature Theorem 15.3.2.1.1 does not appear as it is stated here. The existence for maximal surfaces (that is $H = 0$) is proved in [27, Theorem 1.6] (where the statement is indeed given in any dimension). See also [84]. The result for any value of H, including uniqueness and the foliation statement, appears in [104, Theorem 3.1] again under the assumption that Λ is the graph of a quasi-symmetric homeomorphism.

Moreover, as in the compact case, for every H the surface of constant mean curvature H and asymptotic boundary Λ is unique. This is proved in [27, Theorem 1.10] for Λ the graph of a quasi-symmetric homeomorphism (Definition 15.3.2.2.4 below) and under the additional assumption of bounded second fundamental form. Using the foliation result, uniqueness for the CMC surfaces is showed in [104, Theorem 5.2] again assuming that Λ the graph of a quasi-symmetric homeomorphism.

The proofs, however, can be extended without further difficulty to the general case of any orientation-preserving homeomorphism. We believe that, by a refinement of the arguments, the statement can also be proved for Λ any proper achronal meridian.

When Λ is the graph of an orientation-preserving homeomorphism, the existence part of Theorem 15.3.2.1.1 can be actually be obtained as a straightforward consequence of the following result for constant Gaussian curvature, by applying the normal evolution described in Sect. 15.3.1.1.

Theorem 15.3.2.1.2 ([29, Theorem 1.3]) *Let Λ be any proper achronal meridian in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ which is not a two-step curve. Then there exists a foliation of each connected component of the complement of the convex hull $C(\Lambda)$ in the invisible domain $\Omega(\Lambda)$ by spacelike surfaces of constant Gaussian curvature K, as $K \in (0, +\infty)$.*

Here we say that a proper achronal meridian $\Lambda \subset \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ is a two-step curve if it is the union of four lightlike segments, two horizontal and two vertical in an alternate fashion, under the natural identification of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ with $\mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$. Up to isometry, the configuration of a two-step curve is the one drawn in Fig. 15.8.

In other words, every domain of dependence in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ which is not the invisible domain of a two-step curve admits a foliation of the complement of the convex hull by surfaces of constant Gaussian curvature. The Gaussian curvature function is a time function, up to changing the sign as for the closed case (Theorem 15.3.1.1.1). Moreover in [29, Theorem 1.4] uniqueness of the surfaces of constant Gaussian curvature K in each connected component is proved under the assumption that Λ is the graph of a quasisymmetric homeomorphism.

Theorems 15.3.2.1.1 and 15.3.2.1.2 provide affirmative answers to Questions 8.3 and 8.4 of [16] respectively.

Remark 15.3.2.1.3 Unlike Theorem 15.3.2.1.1, the surfaces $S_K^{\pm}$ of constant Gaussian curvature $K \in (0, +\infty)$ are not always properly embedded in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. Their boundary, however, is explicitely described: it coincides with Λ in the complement of all sawteeth of Λ, where by a *sawtooth* we mean two consecutive lightlike segments in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, one horizontal and one vertical, which are maximal (meaning that they cannot be extended to longer lightlike segments). In correspondence of each sawtooth, the boundary of $S_K^{\pm}$ has an interior spacelike geodesic, namely the geodesic connecting the two endpoints of the sawtooth. See Fig. 15.14, and compare also with Remark 15.2.1.6.5. In conclusion, the surfaces $S_K^{\pm}$ are actually properly embedded in $\Omega(\Lambda)$. When Λ is the graph of an orientation-preserving homeomorphism, it has no sawteeth, hence the boundary of $S_K^{\pm}$ is precisely Λ, hence in this case $S_K^{\pm}$ is properly embedded also in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$.

Remark 15.3.2.1.4 The invisible domain of a two-step curve coincides with its convex hull, hence clearly there can be no existence for surfaces of constant Gaussian curvature in $(0, +\infty)$ in this case, since such a surface would be either convex or concave and therefore live in the convex hull complement, which is empty. On the other hand, when Λ is a two-step curve, there does exists a foliation by surfaces of constant mean curvature with asymptotic boundary Λ, which is given by the surfaces with $z = c$ in the parameterization of $\Omega(\Lambda)$ of Lemma 15.2.2.2.4 and pictured in Fig. 15.11.

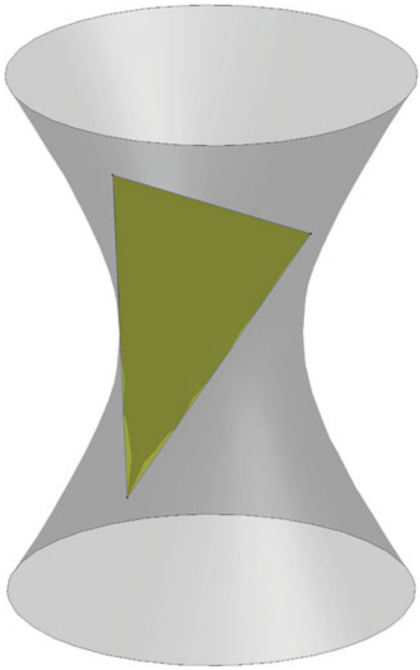


Fig. 15.14 A lightlike triangle, whose boundary consists of a sawtooth, contained in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, and a spacelike geodesic of $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$

15.3.2.2 Extensions and Universal Teichmüller Space

The results of Sect. 15.3.2.1 have applications to the extensions of circle homeomorphisms, by means of the associated map which has already been discussed in the closed case, together with the relevant definitions, in Sect. 15.3.1.2.

Extensions of Circle Homeomorphisms The essential lemma to obtain extensions of circle homeomorphisms is the following. Recall from Sect. 15.1.2.2 that we denoted by π_l, π_r the projections from $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ to $\partial\mathbb{H}^2$ which come from the left and right rulings, or in other words, the projections to the first and second factor in the identification of $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ with $\mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1$.

Lemma 15.3.2.2.1 ([27, Lemma 3.18, Remark 3.19]) *Let $\Lambda \subset \partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ the graph of an orientation-preserving homeomorphisms and S be a spacelike convex (or concave) surface in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ with boundary at infinity Λ. Then the left and right projections $\Pi_l, \Pi_r : S \to \mathbb{H}^2$ extend to the restrictions of $\pi_l, \pi_r : \mathbb{R}\mathrm{P}^1 \times \mathbb{R}\mathrm{P}^1 \to \mathbb{R}\mathrm{P}^1$ on Λ.*

We do not repeat the computations of Sect. 15.3.1.2 here, but essentially for every $\theta \in (0, \pi)$ there are three special surfaces with asymptotic boundary Λ, equidistant from one another, one with constant mean curvature $H = 2/\tan\theta$, one convex with constant Gaussian curvature $\tan^2(\theta/2)$ and one concave with constant Gaussian curvature $1/\tan^2(\theta/2)$, which all have associated map a θ-

landslide. Hence Theorems 15.3.2.1.1, or 15.3.2.1.2, together with the extension lemma (Lemma 15.3.2.2.1), imply the following result:

Corollary 15.3.2.2.2 *Given any orientation-preserving homeomorphism $\varphi : \mathbb{S}^1 \to \mathbb{S}^1$ and any $\theta \in (0, \pi)$, there exists a θ-landslide $\Phi_\theta : \mathbb{H}^2 \to \mathbb{H}^2$ whose extension to $\mathbb{S}^1$ equals φ.*

Since surfaces of constant mean curvature or Gaussian curvature are smooth, the extension Φ_θ will be smooth on $\mathbb{H}^2$, and continuous up to the boundary.

Quasiconformal Mappings We now briefly recall some notions of the theory of quasiconformal mappings. Since here we will be only interested in smooth maps, we give a simplified definition under the smoothness assumption. Here we denote by $\mathbb{D}$ the unit disc in $\mathbb{C}$, as a Riemann surface. A diffeomorphism of $\mathbb{D}$ is quasiconformal if its differential maps circles in $T_z\mathbb{D} \cong \mathbb{C}$ to ellipses of uniformly bounded eccentricity. More formally:

Definition 15.3.2.2.3 A diffeomorphism $\Phi : \mathbb{D} \to \mathbb{D}$ is *quasiconformal* if

$$K(\Phi) := \sup_{z \in \mathbb{D}} \left(\frac{\text{largest eigenvalue of } d\Phi_z^T d\Phi_z}{\text{smallest eigenvalue of } d\Phi_z^T d\Phi_z} \right)^{1/2} < +\infty \, .$$

In this case, the quantity $K(\Phi)$ is called *maximal dilatation* of Φ.

Quasiconformal mappings of $\mathbb{D}$ are precisely those which extend to homeomorphisms of the circle which are quasisymmetric, namely, they distort the cross-ratio of symmetric quadruples in a uniformly controlled way.

Definition 15.3.2.2.4 An orientation-preserving homeomorphism $\varphi : \partial\mathbb{D} \to \partial\mathbb{D}$ is *quasisymmetric* if

$$||\varphi|| := \sup_{\text{cr}(Q)=-1} |\log|\text{cr}(\varphi(Q))|| < +\infty \, .$$

In this case, the quantity $||\varphi||$ is called *cross-ratio norm* of φ.

Some explanation is required concerning Definition 15.3.2.2.4. In the definition of $||\varphi||$, the supremum is taken over all quadruples of points in $\partial\mathbb{D}$ which have cross-ratio equal to -1. We use here a definition of cross-ratio such that a quadruple (p, q, r, s) of points of $\partial\mathbb{D}$ has cross-ratio -1 if and only if it is symmetric, meaning that the two geodesics for the Poincaré metric on $\mathbb{D}$ with endpoints (p, r) and (q, s) intersect orthogonally.

Remark 15.3.2.2.5 Observe also that the cross-ratio norm is a well-defined invariant of a curve Λ in the boundary of $\mathbb{A}\text{d}\mathbb{S}^{2,1}$, since applying isometries of $\mathbb{A}\text{d}\mathbb{S}^{2,1}$, which corresponds to pre- and post-composing φ with elements of $\text{PSL}(2, \mathbb{R})$, do not change the cross-ratio norm. The cross-ratio norm is indeed well-defined on *universal Teichmüller space* $\mathcal{T}(\mathbb{D})$, which can be defined as the space of quasisymmetric homeomorphism up to post-composition with $\mathrm{PSL}(2, \mathbb{R})$.

Theorem 15.3.2.2.6 ([25]) *Any quasiconformal diffeomorphism of* $\mathbb{D}$ *extends to a quasisymmetric homeomorphism of* $\partial\mathbb{D}$. *Conversely, every quasisymmetric homeomorphism of* $\partial\mathbb{D}$ *admits a quasiconformal extension to* $\mathbb{D}$.

Back to Anti-de Sitter geometry, the following result is proved in [27, Theorem 1.4] for minimal Lagrangian extensions (i.e. $\theta = \pi/2$) using maximal surfaces, and in [29, Corollary 1.5] in full generality, using surfaces of constant Gaussian curvature.

Theorem 15.3.2.2.7 *Given any quasisymmetric homeomorphism* $\varphi : \partial\mathbb{D} \to \partial\mathbb{D}$ *and any* $\theta \in (0, \pi)$, *there exists a unique quasiconformal diffeomorphism* $\Phi_\theta : \mathbb{D} \to \mathbb{D}$ *extending* φ *which is a* θ*-landslide.*

This result, which can be seen as a version of the Schoen Conjecture ([22, 86, 100]) for minimal Lagrangian maps and θ-landslides, has essentially two additional points with respect to Corollary 15.3.2.2.2. The first is the fact that one can find an extension which is *quasiconformal*, and the second is the uniqueness *under the quasiconformality assumption*.

For the former, the proof given in [29] essentially consists in showing that the surfaces with constant Gaussian curvature with asymptotic boundary the graph of φ have bounded principal curvatures, a condition which translates in the quasiconformality of the associated map. We will discuss the ingredients of the proof given in [27] below together with other applications.

For the uniqueness instead, roughly speaking the key point is that not only the construction of the associated map can be reversed, but moreover if one starts with a quasiconformal extension (minimal Lagrangian, or landslide more generally), then this can be obtained as the associated map of a surface S (which can be taken of constant mean curvature or of constant Gaussian curvature) in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ whose first fundamental form is complete. Then S has asymptotic boundary a curve Λ in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$, which by Lemma 15.3.2.2.1 coincides necessarily with the graph of φ. The uniqueness of the extension Φ_θ then follows from the uniqueness of the surface, as mentioned in the comments after Theorems 15.3.2.1.1 and 15.3.2.1.2.

Optimality of Minimal Lagrangian Extensions An essential ingredient in the proof of Theorem 15.3.2.2.7 given in [27] for minimal Lagrangian extension is the *width* of the convex hull of Λ, which is defined for every proper achronal meridian Λ as the supremum of the length of timelike curves contained in $C(\Lambda)$. It turns out that the width is always at most $\pi/2$. One then has the following equivalent conditions:

- the proper achronal meridian Λ is the graph of a quasisymmetric homeomorphism;
- the width of the convex hull $C(\Lambda)$ is strictly less than $\pi/2$;
- the principal curvatures of the maximal surface with asymptotic boundary Λ are in $(-1+\epsilon, 1-\epsilon)$ for $\epsilon > 0$;
- the minimal Lagrangian associated map is quasiconformal.

These four points all played an essential rôle in the proof of the following theorem.

Theorem 15.3.2.2.8 ([102, Corollary 2.D]) *There exists a universal constant $C > 0$ such that, for every quasisymmetric homeomorphism $\varphi : \partial\mathbb{D} \to \partial\mathbb{D}$, the maximal dilatation of the quasiconformal minimal Lagrangian extension Φ satisfies:*

$$\log K(\Phi) \leq C||\varphi|| \,.$$

Theorem 15.3.2.2.8 thus adresses a question asked in [16, Section 4.3] about the efficiency of minimal Lagrangian extensions in terms of Teichmüller distance. Let us outline the strategy of the proof. Essentially, the equivalence of the four equivalent points mentioned above need to be quantified. One can in fact obtain quantitative inequalities between the cross-ratio norm of φ and the width of the convex hull, between the width and the principal curvatures of the maximal surface, and finally between the principal curvatures and the maximal dilatation of the minimal Lagrangian extension. Putting together all the quantitative estimates, one obtains the inequality of Theorem 15.3.2.2.8.

There are two extreme cases to be understood. The first is when φ is the graph of a transformation in $\mathrm{PSL}(2, \mathbb{R})$, which is equivalent to the maximal surface being totally geodesic, to the width of the convex hull being equal to 0 and to the minimal Lagrangian extension being isometric. In this case $\|\varphi\| = 0$ and $K(\Phi) = 1$. The qualitatively opposite case occurs for a two-step curve Λ, namely the concatenation of four lightlike segments, which we described in detail in Sect. 15.2.2.2. In this case the maximal surface is given by $\{z = \pi/4\}$ in the coordinates introduced in Lemma 15.2.2.2.4 on the lightlike tetrahedron $\mathscr{T}$. See also Fig. 15.11. Here the maximal surface is intrinsically flat, hence by the Gauss equation its principal curvatures are 1 and -1 at every point. It follows from the discussion of Sect. 15.2.3.4 that in this case there is no associated map at all between subsets of $\mathbb{H}^2$. The width of the convex hull, which is the tetrahedron $\mathscr{T}$ itself, equals $\pi/2$. The proof of Theorem 15.3.2.2.8 roughly speaking consists in showing that, as $\|\varphi\|$ approaches 0 or $+\infty$, the geometry of the corresponding maximal surface approaches that of the two extreme examples, the totally geodesic plane and the flat maximal surface with principal curvatures identically 1 and -1.

15.3.2.3 Related Results

We briefly mention here some related results. The work [97] studies a "universal" version of the correspondence between MGH Anti-de Sitter manifolds and Teichmüller space, and parameterizes a suitably defined moduli space of these structures by the product of two copies of the universal Teichmüller space $\mathcal{T}(\mathbb{D})$. (Recall Remark 15.3.2.2.5 for the definition of $\mathcal{T}(\mathbb{D})$.) Moreover, a universal version of the map F of Theorem 15.3.1.3.3 is constructed, namely a map $T^*\mathcal{T}(\mathbb{D}) \to$

$\mathcal{T}(\mathbb{D}) \times \mathcal{T}(\mathbb{D})$, using the fact that the cotangent bundle $T^*\mathcal{T}(\mathbb{D})$ is identified to the space of bounded holomorphic quadratic differentials.

A qualitatively opposite situation is described in [105], where maximal surfaces whose first fundamental form is conformal to $\mathbb{C}$, and the second fundamental form is the real part of a polynomial quadratic differential on $\mathbb{C}$, are considered. It is shown that these maximal surfaces are characterized by having asymptotic boundary a curve in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ composed of the concatenation of a finite number of lightlike sides.

A result of prescription of the induced metric on convex surfaces of constant Gaussian curvature, generalizing to the universal setting some of the results discussed in Sect. 15.3.1.5, is proved in [34]. In a similar spirit, results about the realization of metrics on the boundary of ideal polyhedra in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ are presented in [53], and a first result on the prescription of bending laminations on the boundary of the convex hull in the universal setting will appear in [88].

15.3.2.4 Cone Singularities and Manifolds with Particles

The last part of this paper will briefly survey results on Anti-de Sitter manifolds with spacelike surfaces of finite type, namely homeomorphic to the complement of a finite number of punctures in a closed surface. Depending on the geometry near the removed points, different geometric structures can arise. For instance MGH Anti-de Sitter manifolds with *particles* can be defined, where a particle is a cone singularity of timelike type. It is required in the definition that the manifold contains a locally convex Cauchy surface orthogonal to the singular locus. Hence the first fundamental form of such a Cauchy surface has a cone point in correspondence of each intersection with a particle. Many of the results we mention here are the counterpart "with particles" of the results which have been described in Sect. 15.3.1 for the closed case. Hence we will omit most of the details here.

If one fixes the number of cone points (say n) and the cone angles $\theta_1, \ldots, \theta_n$ at each cone point, which is assumed to be smaller than π, it was proved in [26, Theorem 1.4] that the deformation space of MGH Anti-de Sitter manifolds with a Cauchy surface homeomorphic to Σ and particles of cone angles $\theta_1, \ldots, \theta_n$ is homeomorphic to the product of two copies of the Teichmüller space of Σ with n cone points of angles $\theta_1, \ldots, \theta_n$. An earthquake theorem for hyperbolic surfaces with cone points of angle less than π has then been proved, see [26, Theorem 1.2]. The prescription of measured geodesic laminations on the boundary of the convex core, as an analogue of Theorem 15.3.1.5.1, has been established in [28].

The existence and uniqueness of a maximal surface orthogonal to the singular locus was proved in [110, Theorem 1.4], thus obtaining as a consequence the existence and uniqueness of a minimal Lagrangian map between two hyperbolic surfaces with the same cone angles (less than π) in a given isotopy class ([110, Theorem 1.3]). Moreover, together with [79, Theorem 5.11] which is a version "with particles" of Theorem 15.3.1.3.1, one obtains a parameterization of the deformation space of MGH AdS manifolds with particles by means of the cotangent bundle of the

Teichmüller space of Σ with cone angles $\theta_1, \ldots, \theta_n < \pi$, which is also identified to the bundle of holomorphic quadratic differentials with at most simple poles at the punctures.

The existence and uniqueness of the maximal surface orthogonal to the singular locus has then been improved in [40, Theorem 1.1] to the existence of a foliation by constant mean curvature surfaces. The proof actually relies on the results of [39], namely the existence of a foliation by surfaces of constant Gaussian curvature of each connected component of the convex core complement ([39, Theorem 1.1]). These results have of course applications for the existence of θ-landslides between hyperbolic surfaces with cone angles, see [39, Theorem 5.8]. Many of these results had been conjectured in [16, Section 6.2], see Questions 6.2–6.5.

The general study of cone singularities besides the case of particles, including the possibility of intersections between singularities ("collisions"), and introducing the notion of global hyperbolicity in this setting together with examples, has been pursued in [15] and [17].

15.3.2.5 Boundary Components and Multi-Black Holes

Teichmüller theory for hyperbolic surfaces with boundary components or cusps is instead intimately related to the geometry of Anti-de Sitter manifolds with multi-black holes. Here we only sketch the definition, and we refer to [7] and [32] for more details.

Let us now assume that $\rho_l, \rho_r : \pi_1\Sigma \to \mathrm{PSL}(2, \mathbb{R})$ are the holonomy representations of complete hyperbolic structures on Σ with cusps and geodesic boundary components. Then there is a maximal domain Ω in $\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$ on which the action of (ρ_l, ρ_r) is free and properly discontinuous, which is however not globally hyperbolic. The domain Ω can actually be described as the union of all globally hyperbolic domains on which the action is free and properly discontinuous, and each of these can be obtained by the following construction.

The limit set of the action of $\rho_l(\pi_1\Sigma)$ on $\mathbb{H}^2$ can be described as the complement of a family of open arcs in $\partial\mathbb{H}^2$, where the endpoints of each removed arc are the endpoints of a lift to $\mathbb{H}^2$ of a geodesic boundary component. The limit set for $\rho_r(\pi_1\Sigma)$ has an analogue description, and similarly to the closed case, one can find a circle homeomorphism $\varphi : \partial\mathbb{H}^2 \to \partial\mathbb{H}^2$ which is equivariant with respect to the actions of ρ_l and ρ_r. This is however not uniquely determined when geodesic boundary components are present. In fact, there are many possible choices of the equivariant map φ, and the freedom of such choice corresponds to the definition of φ on each open arc in $\partial\mathbb{H}^2$ in the complement of the limit set.

The choice of some particular equivariant map φ gives a globally hyperbolic domain, namely the invisible domain of the graph of φ, on which the action is free and properly discontinuous. The union of all such domains provides the maximal domain Ω, whose quotient is a maximal Anti-de Sitter manifold with multi-black holes.

Using the pleated surface construction in this setting, in [32] an earthquake theorem was proved for surfaces with boundary, namely given two hyperbolic structures on Σ with geodesic boundary, there exist 2^k left (or right) earthquake maps, where k is the number of boundary components of Σ. The 2^k choices correspond to the choice, for every boundary component of Σ, to the sense in which the earthquake lamination "spirals" around the boundary; in terms of Anti-de Sitter geometry, this is the choice of a future or past sawtooth in $\partial\mathbb{A}\mathrm{d}\mathbb{S}^{2,1}$. An extension of this result to crowned hyperbolic surfaces was given in [94].

The PhD thesis [94] contains a result of prescription of two filling measured geodesic laminations on a hyperbolic surface with boundary, as the bending laminations on the boundary components of the convex core of an Anti-de Sitter manifold with multi-black holes. Finally, [106] contains a study of the maximal surfaces which appear in this case, and of the associated minimal Lagrangian diffeomorphisms, in terms of holomorphic quadratic differentials with poles of order at most 2, hence extending the parameterization we discussed in Sect. 15.3.1.3 in the closed case.

Acknowledgments We would like to thank Athanase Papadopoulos for the opportunity of writing this article, for his patience during the preparation of the manuscript, and for suggesting several improvements on the exposition. Moreover, we are very grateful to Thierry Barbot and François Fillastre for useful comments on a preliminary version of this work.

References

1. D. Alessandrini, Higgs bundles and geometric structures on manifolds. SIGMA Symmetry Integr. Geom. Methods Appl. **15**, paper 039, 32 (2019)
2. D. Alessandrini, Q. Li, AdS 3-manifolds and Higgs bundles. Proc. Am. Math. Soc. **146**(2), 845–860 (2018)
3. L. Andersson, Constant mean curvature foliations of flat space-times. Commun. Anal. Geom. **10**(5), 1125–1150 (2002)
4. L. Andersson, V. Moncrief, A.J. Tromba, On the global evolution problem in $2+1$ gravity. J. Geom. Phys. **23**(3–4), 191–205 (1997)
5. R. Arnowitt, S. Deser, C.W. Misner, Dynamical structure and definition of energy in general relativity. Phys. Rev. II. Ser. **116**, 1322–1330 (1959)
6. T. Barbot, Causal properties of AdS-isometry groups. I. Causal actions and limit sets. Adv. Theor. Math. Phys. **12**(1), 1–66 (2008)
7. T. Barbot, Causal properties of AdS-isometry groups. II. BTZ multi-black-holes. Adv. Theor. Math. Phys. **12**(6), 1209–1257 (2008)
8. T. Barbot, Deformations of Fuchsian AdS representations are quasi-Fuchsian. J. Differ. Geom. **101**(1), 1–46 (2015)
9. T. Barbot, Lorentzian Kleinian groups, in *Handbook of Group Actions. Vol. III*. Advanced Lectures in Mathematics, vol. 40 (International Press, Somerville, 2018), pp. 311–358
10. T. Barbot, Q. Mérigot, Anosov AdS representations are quasi-Fuchsian. Groups Geom. Dyn. **6**(3), 441–483 (2012)
11. T. Barbot, F. Béguin, A. Zeghib, Feuilletages des espaces temps globalement hyperboliques par des hypersurfaces à courbure moyenne constante. C. R. Math. Acad. Sci. Paris **336**(3), 245–250 (2003)

12. T. Barbot, F. Béguin, A. Zeghib, Constant mean curvature foliations of globally hyperbolic spacetimes locally modelled on AdS_3. Geom. Dedicata **126**, 71–129 (2007)
13. T. Barbot, V. Charette, T. Drumm, W.M. Goldman, K. Melnick, A primer on the (2 + 1) Einstein universe, in *Recent Developments in Pseudo-Riemannian Geometry*. ESI Lectures in Mathematics and Physics (European Mathematical Society, Zürich, 2008), pp. 179–229
14. T. Barbot, F. Béguin, A. Zeghib, Prescribing Gauss curvature of surfaces in 3-dimensional spacetimes: application to the Minkowski problem in the Minkowski space. Ann. Inst. Fourier (Grenoble) **61**(2), 511–591 (2011)
15. T. Barbot, F. Bonsante, J.-M. Schlenker, Collisions of particles in locally AdS spacetimes I. Local description global examples. Commun. Math. Phys. **308**(1), 147–200 (2011)
16. T. Barbot, F. Bonsante, J. Danciger, W.M. Goldman, F. Guéritaud, F. Kassel, K. Krasnov, J.-M. Schlenker, A. Zeghib, Some open questions on anti-de sitter geometry (2012). ArXiv:1205.6103
17. T. Barbot, F. Bonsante, J.-M. Schlenker, Collisions of particles in locally AdS spacetimes II. Moduli of globally hyperbolic spaces. Commun. Math. Phys. **327**(3), 691–735 (2014)
18. A. Basmajian, M. Zeinalian, Möbius transformations of the circle form a maximal convergence group, in *The Geometry of Riemann Surfaces and Abelian Varieties*. Contemporary Mathematics, vol. 397 (American Mathematical Society, Providence, 2006), pp. 1–6
19. J.K. Beem, P.E. Ehrlich, *Global Lorentzian Geometry*. Monographs and Textbooks in Pure and Applied Mathematics, vol. 67 (Marcel Dekker, New York, 1981)
20. R. Benedetti, F. Bonsante, (2 + 1) Einstein spacetimes of finite type, in *Handbook of Teichmüller Theory. Vol. II*. IRMA Lectures in Mathematics and Theoretical Physics, vol. 13 (European Mathematical Society, Zürich, 2009), pp. 533–609
21. R. Benedetti, F. Bonsante, Canonical Wick rotations in 3-dimensional gravity. Mem. Am. Math. Soc. **198**(926), viii+164 (2009)
22. Y. Benoist, D. Hulin, Harmonic quasi-isometric maps between rank one symmetric spaces. Ann. Math. **185**(3), 895–917 (2017)
23. A.N. Bernal, M. Sánchez, On smooth Cauchy hypersurfaces and Geroch's splitting theorem. Commun. Math. Phys. **243**(3), 461–470 (2003)
24. A.N. Bernal, M. Sánchez, Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes. Commun. Math. Phys. **257**(1), 43–50 (2005)
25. A. Beurling, L. Ahlfors, The boundary correspondence under quasiconformal mappings. Acta Math. **96**, 125–142 (1956)
26. F. Bonsante, J.-M. Schlenker, AdS manifolds with particles and earthquakes on singular surfaces. Geom. Funct. Anal. **19**(1), 41–82 (2009)
27. F. Bonsante, J.-M. Schlenker, Maximal surfaces and the universal Teichmüller space. Invent. Math. **182**(2), 279–333 (2010)
28. F. Bonsante, J.-M. Schlenker, Fixed points of compositions of earthquakes. Duke Math. J. **161**(6), 1011–1054 (2012)
29. F. Bonsante, A. Seppi, Area-preserving diffeomorphisms of the hyperbolic plane and K-surfaces in anti-de Sitter space. J. Topol. **11**(2), 420–468 (2018)
30. F. Bonsante, A. Seppi, Equivariant maps into anti-de Sitter space and the symplectic geometry of $\mathbb{H}^2 \times \mathbb{H}^2$. Trans. Am. Math. Soc. **371**(8), 5433–5459 (2019)
31. F. Bonsante, A. Seppi, A. Tamburelli, On the volume of anti–de Sitter maximal globally hyperbolic three-manifolds. Geom. Funct. Anal. **27**(5), 1106–1160 (2017)
32. F. Bonsante, K. Krasnov, J.-M. Schlenker, Multi-black holes and earthquakes on Riemann surfaces with boundaries. Int. Math. Res. Not. IMRN **2011**(3), 487–552 (2011)
33. F. Bonsante, G. Mondello, J.-M. Schlenker, A cyclic extension of the earthquake flow I. Geom. Topol. **17**(1), 157–234 (2013)
34. F. Bonsante, J. Danciger, S. Maloni, J.-M. Schlenker, The induced metric on the boundary of the convex hull of a quasicircle in hyperbolic and anti de sitter geometry (2019). ArXiv:1902.04027, to appear in Geom. Topol
35. S. Carlip, *Quantum Gravity in* 2 + 1 *Dimensions* (Cambridge University Press, Cambridge, 1998)

36. S. Carlip, Quantum gravity in 2 + 1 dimensions: the case of a closed universe. Living Rev. Relat. **8**, 63 (2005). Id/No 2005-1
37. V. Charette, T.A. Drumm, W.M. Goldman, Affine deformations of a three-holed sphere. Geom. Topol. **14**(3), 1355–1382 (2010)
38. V. Charette, T.A. Drumm, W.M. Goldman, Finite-sided deformation spaces of complete affine 3-manifolds. J. Topol. **7**(1), 225–246 (2014)
39. Q. Chen, J.-M. Schlenker, Constant gauss curvature foliations of ads spacetimes with particles. Trans. Am. Math. Soc. **373**(6), 4013–4049 (2020)
40. Q. Chen, A. Tamburelli, Constant mean curvature foliation of globally hyperbolic (2 + 1)-spacetimes with particles. Geom. Dedicata **201**, 281–315 (2019)
41. S. Choi, W. Goldman, Topological tameness of Margulis spacetimes. Am. J. Math. **139**(2), 297–345 (2017)
42. Y. Choquet-Bruhat, Théorème global d'unicité pour les solutions des équations d'Einstein. C. R. Acad. Sci. Paris Sér. A-B **266**, A182–A184 (1968)
43. Y. Choquet-Bruhat, The bearings of global hyperbolicity on existence and uniqueness theorems in general relativity. Gen. Relativ. Gravit. **2**, 1–6 (1971)
44. B. Collier, N. Tholozan, J. Toulisse, The geometry of maximal representations of surface groups into $SO_0(2, n)$. Duke Math. J. **168**(15), 2873–2949 (2019)
45. D. Cooper, J. Danciger, A. Wienhard, Limits of geometries. Trans. Am. Math. Soc. **370**(9), 6585–6627 (2018)
46. J. Danciger, A geometric transition from hyperbolic to anti-de Sitter geometry. Geom. Topol. **17**(5), 3077–3134 (2013)
47. J. Danciger, Ideal triangulations and geometric transitions. J. Topol. **7**(4), 1118–1154 (2014)
48. J. Danciger, F. Guéritaud, F. Kassel, Fundamental domains for free groups acting on anti–de Sitter 3-space. Math. Res. Lett. **23**(3), 735–770 (2016)
49. J. Danciger, F. Guéritaud, F. Kassel, Geometry and topology of complete Lorentz spacetimes of constant curvature. Ann. Sci. Éc. Norm. Supér. **49**(1), 1–56 (2016)
50. J. Danciger, F. Guéritaud, F. Kassel, Geometry and topology of complete Lorentz spacetimes of constant curvature. Ann. Sci. Éc. Norm. Supér. **49**(1), 1–56 (2016)
51. J. Danciger, F. Guéritaud, F. Kassel, Margulis spacetimes via the arc complex. Invent. Math. **204**(1), 133–193 (2016)
52. J. Danciger, F. Guéritaud, F. Kassel, Convex cocompactness in pseudo-Riemannian hyperbolic spaces. Geom. Dedicata **192**, 87–126 (2018)
53. J. Danciger, S. Maloni, J.-M. Schlenker, Polyhedra inscribed in a quadric. Invent. Math. **221**(1), 237–300 (2020)
54. J. Danciger, T.A. Drumm, W.M. Goldman, I. Smilga, Proper actions of discrete groups of affine transformations, in *Dynamics, Geometry, Number Theory: The Impact of Margulis on Modern Mathematics* (2020, to appear)
55. B. Deroin, N. Tholozan, Dominating surface group representations by Fuchsian ones. Int. Math. Res. Not. IMRN **2016**(13), 4145–4166 (2016)
56. B. Diallo, Prescribing metrics on the boundary of convex cores of globally hyperbolic maximal compact AdS manifolds. Ph.D. Thesis, Université Paul Sabatier - Toulouse III (2014)
57. T.A. Drumm, Fundamental polyhedra for Margulis space-times. Topology **31**(4), 677–683 (1992)
58. T.A. Drumm, Linear holonomy of Margulis space-times. J. Differ. Geom. **38**(3), 679–690 (1993)
59. K. Ezawa, Chern-Simons quantization of (2 + 1)-anti-de Sitter gravity on a torus. Classical Quant. Gravity **12**(2), 373–391 (1995)
60. F. Fillastre, Fuchsian polyhedra in Lorentzian space-forms. Math. Ann. **350**(2), 417–453 (2011)
61. F. Fillastre, J.-M. Schlenker, Flippable tilings of constant curvature surfaces. Illinois J. Math. **56**(4), 1213–1256 (2012)

62. F. Fillastre, A. Seppi, Spherical, hyperbolic, and other projective geometries: convexity, duality, transitions, in *Eighteen Essays in Non-Euclidean Geometry*. IRMA Lectures in Mathematics and Theoretical Physics, vol. 29 (European Mathematical Society, Zürich, 2019), pp. 321–409
63. C. Frances, Géométrie et dynamique lorentziennes conformes. Ph.D. Thesis, École Normale Supérieure de Lyon (2002)
64. C. Frances, The conformal boundary of anti-de Sitter space-times, in *AdS/CFT Correspondence: Einstein Metrics and Their Conformal Boundaries*. IRMA Lectures in Mathematics and Theoretical Physics, vol. 8, pp. 205–216 (European Mathematical Society, Zürich, 2005)
65. D. Fried, W.M. Goldman, Three-dimensional affine crystallographic groups. Adv. Math. **47**(1), 1–49 (1983)
66. S. Gallot, D. Hulin, J. Lafontaine, *Riemannian Geometry*, Universitext. 3rd edn. (Springer, Berlin, 2004)
67. F.P. Gardiner, N. Lakic, *Quasiconformal Teichmüller Theory*. Mathematical Surveys and Monographs, vol. 76 (American Mathematical Society, Providence, 2000)
68. R. Geroch, Domain of dependence. J. Math. Phys. **11**, 437–449 (1970)
69. O. Glorieux, D. Monclair, Critical exponent and hausdorff dimension in Pseudo-Riemannian hyperbolic geometry. International Mathematics Research Notices. rnz098 (2019)
70. O. Glorieux, D. Monclair, N. Tholozan, Hausdorff dimension of limit sets for projective anosov representations. ArXiv: 1902.01844 (2019)
71. W.M. Goldman, Discontinuous groups and the Euler class. Ph.D. Thesis. University of California, Berkeley (1980)
72. W.M. Goldman, Crooked surfaces and anti-de Sitter geometry. Geom. Dedicata **175**, 159–187 (2015)
73. W.M. Goldman, F. Labourie, G. Margulis, Proper affine actions and geodesic flows of hyperbolic surfaces. Ann. Math. **170**(3), 1051–1083 (2009)
74. F. Guéritaud, O. Guichard, F. Kassel, A. Wienhard, Anosov representations and proper actions. Geom. Topol. **21**(1), 485–584 (2017)
75. H. Hopf, Über Flächen mit einer Relation zwischen den Hauptkrümmungen. Math. Nachr. **4**, 232–249 (1951)
76. F. Kassel, Geometric structures and representations of discrete groups, in *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. II. Invited Lectures* (World Scientific, Hackensack, 2018), pp. 1115–1151
77. K. Kobayashi, S. Nomizu, *Foundations of Differential Geometry. Vol. II. Reprint of the 1969 Original.* Wiley Classics Library. A Wiley-Interscience Publication (Wiley, New York, 1996)
78. S. Kobayashi, *Transformation Groups in Differential Geometry*. Classics in Mathematics (Springer, Berlin, 1995). Reprint of the 1972 edition
79. K. Krasnov, J.-M. Schlenker, Minimal surfaces and particles in 3-manifolds. Geom. Dedicata **126**, 187–254 (2007)
80. R.S. Kulkarni, F. Raymond, 3-dimensional Lorentz space-forms and Seifert fiber spaces. J. Differ. Geom. **21**(2), 231–268 (1985)
81. H. Labeni, Fuchsian isometric immersion in anti-de sitter space (2019). ArXiv:1912.03123
82. F. Labourie, Surfaces convexes dans l'espace hyperbolique et $\mathbb{C}P^1$-structures. J. Lond. Math. Soc. **45**(3), 549–565 (1992)
83. F. Labourie, J.-M. Schlenker, Surfaces convexes fuchsiennes dans les espaces lorentziens à courbure constante. Math. Ann. **316**(3), 465–483 (2000)
84. F. Labourie, J. Toulisse, M. Wolf, Plateau Problems for Maximal Surfaces in Pseudo-Hyperbolic Spaces. ArXiv 2006.12190
85. G.-S. Lee, L. Marquis, Anti-de Sitter strictly GHC-regular groups which are not lattices. Trans. Am. Math. Soc. **372**(1), 153–186 (2019)
86. V. Markovic, Harmonic maps and the Schoen conjecture. J. Am. Math. Soc. **30**(3), 799–817 (2017)
87. E.J. McShane, Extension of range of functions. Bull. Am. Math. Soc. **40**(12), 837–842 (1934)

88. L. Merlin, J.-M. Schlenker, Bending laminations on convex hulls of anti-de sitter quasicircles (2020). ArXiv 2006:13470
89. G. Mess, Lorentz spacetimes of constant curvature. Geom. Dedicata **126**, 3–45 (2007)
90. D. Monclair, J.-M. Schlenker, N. Tholozan, Gromov-Thurston manifolds and AdS geometry (in preparation, 2020)
91. V. Moncrief, Reduction of the Einstein equations in $2+1$ dimensions to a Hamiltonian system over Teichmüller space. J. Math. Phys. **30**(12), 2907–2914 (1989)
92. P. Piccione, D.V. Tausk, The single-leaf Frobenius theorem with applications. Resenhas **6**(4), 337–381 (2005)
93. S. Riolo, A. Seppi, Geometric transition from hyperbolic to anti-de sitter structures in dimension four (Preprint, 2019). arXiv:1908.05112
94. D. Rosmondi, Earthquakes on hyperbolic surfaces with geodesic boundary and Anti de Sitter geometry. Ph.D. Thesis, Università degli Studi di Pavia (2017)
95. F. Salein, Variétés anti-de Sitter de dimension 3, in *Séminaire de Théorie Spectrale et Géométrie, No. 15, Année 1996–1997*, vol. 15 (Univ. Grenoble I, Saint-Martin-d'Hères, 1997), pp. 37–42
96. F. Salein, Variétés anti-de Sitter de dimension 3 exotiques. Ann. Inst. Fourier (Grenoble) **50**(1), 257–284 (2000)
97. C. Scarinci, K. Krasnov, The universal phase space of AdS_3 gravity. Commun. Math. Phys. **322**(1), 167–205 (2013)
98. C. Scarinci, J.-M. Schlenker, Symplectic Wick rotations between moduli spaces of 3-manifolds. Ann. Sci. Norm. Super. Pisa Cl. Sci. **18**(3), 781–829 (2018)
99. J.-M. Schlenker, Variétés lorentziennes plates vues comme limites de variétés anti–de Sitter [d'après Danciger, Guéritaud et Kassel]. Astérisque, (380, Séminaire Bourbaki. vol. 2014/2015):Exp. No. 1103, 475–497 (2016)
100. R.M. Schoen, The role of harmonic mappings in rigidity and deformation problems, in *Complex Geometry (Osaka, 1990)*. Lecture Notes in Pure and Applied Mathematics, vol. 143 (Dekker, New York, 1993), pp. 179–200
101. A. Seppi, The flux homomorphism on closed hyperbolic surfaces and anti-de Sitter three-dimensional geometry. Complex Manifolds **4**(1), 183–199 (2017)
102. A. Seppi, Maximal surfaces in anti–de Sitter space, width of convex hulls and quasiconformal extensions of quasisymmetric homeomorphisms. J. Eur. Math. Soc. **21**(6), 1855–1913 (2019)
103. A. Tamburelli, Prescribing metrics on the boundary of anti–de Sitter 3-manifolds. Int. Math. Res. Not. IMRN **2018**(5), 1281–1313 (2018)
104. A. Tamburelli, Constant mean curvature foliation of domains of dependence in AdS_3. Trans. Am. Math. Soc. **371**(2), 1359–1378 (2019)
105. A. Tamburelli, Polynomial quadratic differentials on the complex plane and light-like polygons in the Einstein universe. Adv. Math. **352**, 483–515 (2019)
106. A. Tamburelli, Regular globally hyperbolic maximal anti-de Sitter structures. J. Topol. **13**(1), 416–439 (2020)
107. C.H. Taubes, Minimal surfaces in germs of hyperbolic 3-manifolds, in *Proceedings of the Casson Fest*. Geometry and Topology Monographs, vol. 7 (Geometry and Topology Publications, Coventry, 2004), pp. 69–100
108. N. Tholozan, Dominating surface group representations and deforming closed anti-de Sitter 3-manifolds. Geom. Topol. **21**(1), 193–214 (2017)
109. N. Tholozan, The volume of complete anti–de Sitter 3-manifolds. J. Lie Theory **28**(3), 619–642 (2018)
110. J. Toulisse, Maximal surfaces in anti–de Sitter 3-manifolds with particles. Ann. Inst. Fourier Grenoble) **66**(4), 1409–1449 (2016)

111. S. Trapani, G. Valli, One-harmonic maps on Riemann surfaces. Commun. Anal. Geom. **3**(3–4), 645–681 (1995)
112. A. Wienhard, An invitation to higher Teichmüller theory, in *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. II. Invited lectures* (World Scientific, Hackensack, 2018), pp. 1013–1039
113. E. Witten, 2 + 1-dimensional gravity as an exactly soluble system. Nuclear Phys. B **311**(1), 46–78 (1988/1989)

Chapter 16
Quasi-Fuchsian Co-Minkowski Manifolds

Thierry Barbot and François Fillastre

Abstract Since the work of W.P. Thurston, some maps from Teichmüller space into itself can be described using the extrinsic geometry of surfaces in three dimensional hyperbolic space. Similarly, since the work of G. Mess, some of these maps can be described using the extrinsic geometry of surfaces in Lorentzian space-forms. Here we will use the degenerate geometry of co-Minkowski space to prove, and generalize to any dimension, a theorem of Thurston that says that the total length of measured geodesic laminations provides an asymmetric norm on the tangent bundle of Teichmüller space.

The main part of this chapter is an introduction to the geometry of co-Minkowksi space, the space of unoriented spacelike hyperplanes of the Minkowski space. Affine deformations of cocompact lattices of hyperbolic isometries act on it, in a way similar to the way that quasi-Fuchsian groups act on hyperbolic space. In particular, there is a convex core construction. There is also a unique "mean" hypersurface, i.e. with traceless second fundamental form. The mean distance between the mean hypersurface and the lower boundary of the convex core endows the space of affine deformations of a given lattice with an asymmetric norm. The symmetrization of the asymmetric norm is simply the volume of the convex core.

In dimension $2 + 1$, the asymmetric norm is the total length of the bending lamination of the lower boundary component of the convex core. We then obtain an extrinsic proof of the theorem of Thurston mentioned above.

We also exhibit and comment on the Anosov-like character of these deformations, similar to the Anosov character of the quasi-Fuchsian representations pointed out in Guichard and Wienhard (Invent Math 190(2):357–438, 2012).

T. Barbot
Université d'Avignon, Département de Mathématiques, Avignon, France

F. Fillastre (✉)
CY Cergy Paris Université, Laboratoire AGM, Cergy, France
e-mail: francois.fillastre@u-cergy.fr

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1_16

Keywords Co-Minkowski space · Compact hyperbolic manifolds · Earthquake norm · Codazzi tensors · Convex core · Anosov representation

AMS Codes 52A55, 32G15, 37D20

16.1 Introduction

One of the main examples of the relation between the geometry of the convex core of hyperbolic quasi–Fuchsian manifolds (such as exposed for example in [26]) and transformations on Teichmüller space as pointed out by Thurston is the grafting map. Similarly, Mess noted that the geometry of the convex core in some anti-de Sitter manifolds is intimately related to the earthquake map. Here we will study the geometry of the convex core of co-Minkowski quasi-Fuchsian manifolds (to be defined below). Actually our study is not restricted to the $2+1$ dimensional case.

Action of Hyperbolic Isometries on Model Spaces Let $\mathbb{H}^d/\Gamma$ be an oriented compact hyperbolic manifold. In the Klein projective model, the hyperbolic space $\mathbb{H}^{d+1}$ is the interior of a ball, and some features of the action of Γ can be described looking at the exterior of the ball, naturally endowed with a Lorentzian structure of constant curvature one, and called de Sitter space. Using affine duality with respect to the unit sphere, de Sitter space can be seen as the space of totally geodesic hypersurfaces of $\mathbb{H}^d$.

Since the work of G. Mess [1, 51], the action of cocompact lattices of $O(d,1)$ on model spaces[1] attracted attention from geometers, see e.g. the surveys [6, 29]. Apart from de Sitter space, Anti-de Sitter space has constant curvature -1 and Minkowski space is the flat one. As we said, de Sitter space is the dual of the hyperbolic space, and Anti-de Sitter space is its own dual, see e.g. [28]. Co-Minkowski space is the dual of Minkowski space. More precisely, it is the space of spacelike hyperplanes of Minkowski space. It comes with a degenerate metric of constant curvature -1.[2]

Curvature -1 spaces	$\overset{dual}{\longleftrightarrow}$	Lorentzian spaces
Hyperbolic space	$\longleftrightarrow$	de Sitter space
co-Minkowski space	$\longleftrightarrow$	Minkowski space
Anti-de Sitter space	$\longleftrightarrow$	Anti-de Sitter space

[1] We call a d-dimensional *model space* the quotient by the antipodal map of a pseudo-sphere in $\mathbb{R}^{d+1}$, see [28].

[2] For a pseudo-Riemannian manifold, the sectional curvature is computed only for planes of the tangent space on which the metric is non-degenerate.

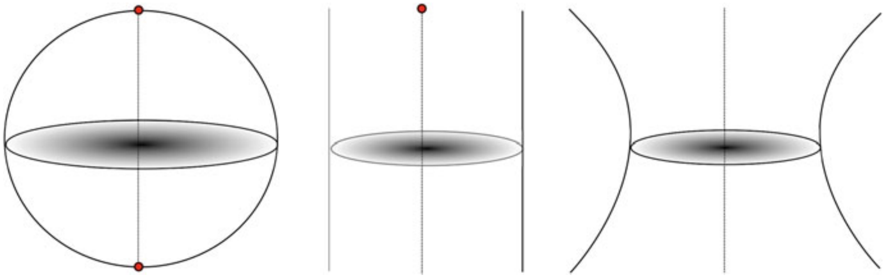


Fig. 16.1 Affine models of the three 3d model spaces of constant curvature -1. Shadowed discs are totally geodesic embedded hyperbolic planes. From left to right: hyperbolic space (Riemannian), co-Minkowski space (degenerated), anti-de Sitter space (Lorentzian)

Co-Minkowski Space The first part of this chapter is an elementary introduction to co-Minkowski space $^*\mathbb{R}^{d,1}$. This space has recently attracted attention under the name "half-pipe", as introduced by J. Dancinger in [20, 21],[3] and used in recent works [24, 58, 60], see also [28].

We will focus on a "Klein model" of co-Minkowski space as the subspace $B^d \times \mathbb{R}$ of the affine space $\mathbb{R}^{d+1}$, where B^d is an open unit ball, see Fig. 16.1. In general, the interest of an affine model is that (unparameterized) geodesics are affine segments, so for example some affine notions as convexity or convex hull are easily tractable. In the particular case of co-Minkowski space, many analogues of classical differential geometry results are easier than the original ones, for example:

- the (smooth) hypersurfaces carrying a non-degenerate induced metric are all hyperbolic, and when they are metrically complete, they are graphs of functions on the ball B^d,
- the shape operator of graph hypersurfaces gives symmetric Codazzi tensors on the hyperbolic space $\mathbb{H}^d$,
- actually, the correspondence between complete hyperbolic hypersurfaces and hyperbolic symmetric Codazzi tensors is one-to-one, that gives a simplified co-Minkowski version of the fundamental theorem of hypersurfaces (Sect. 16.2.3.1),
- complete hyperbolic hypersurfaces such that the trace of the shape operator vanishes are called *mean surfaces*; existence and uniqueness of such hypersurfaces are straightforward consequence of classical theory of elliptic PDEs on the ball (Sect. 16.2.3.2),
- the functions whose graph is a boundary of the convex hull of the graph of a continuous map $b : \partial B^d \to \mathbb{R}$ are solutions of the classical Monge–Ampère equation (Sect. 16.2.3.3).

[3]The surface $co\mathcal{M}in^{1+1}$ in Fig. 16.2 would deserve the name *half-pipe*. The name co-Minkowski space comes from the particular situation of this *co-pseudo-Euclidean space*, see the corresponding entry in the *Encyclopaedia of Mathematics*.

Another nice feature of the cylinder model of co-Minkowski space is that it allows an easy definition of *degenerations* of hyperbolic or Anti-de Sitter manifolds to a co-Minkowski manifold, as Fig. 16.1 heuristically suggests. In turn, co-Minkowski geometry as a *transitional geometry* between the hyperbolic geometry and the AdS geometry was the main motivation of [20, 21], see also [28, 57, 60]. Such considerations are out of the scope of the present survey.

The Action of $H^1(\Gamma, \mathbb{R}^{d,1})$ By duality, the group of isometries of Minkowski space, that is $O(d,1) \ltimes \mathbb{R}^{d,1}$, acts on co-Minkowski space, preserving the degenerate metric (see Remark 16.2.2). For our purpose, it will be more relevant to restrict ourselves to the action of $O_0(d,1) \ltimes \mathbb{R}^{d,1}$, where $O_0(d,1)$ is the connected component of the identity of $O(d,1)$. If Γ is a Kleinian cocompact subgroup of $O_0(d,1)$, then the representations of Γ into $O_0(d,1) \ltimes \mathbb{R}^{d,1}$ are parameterized by maps $\tau : \Gamma \to \mathbb{R}^{d,1}$ satisfying a cocycle relation. Let $Z^1(\Gamma, \mathbb{R}^{d,1})$ be the space of cocycles.

Different totally geodesic embedding of $\mathbb{H}^d$ (on which Γ acts) into co-Minkowski space will give different cocycles, related by coboundary conditions. So we are interested in the space $H^1(\Gamma, \mathbb{R}^{d,1})$, the quotient of the space of cocycles by the coboundaries. From an extrinsic point of view, the vector space $H^1(\Gamma, \mathbb{R}^{d,1})$ is the space of deformations of Γ into the group of affine isometries, up to conjugacy by translations. But $H^1(\Gamma, \mathbb{R}^{d,1})$ encodes more information:

- For $d > 2$, due to Mostow's rigidity theorem, it is not possible to non-trivially deform Γ among Kleinian subgroups of $O(d,1)$. But it is possible to look at deformations of the canonical representation of Γ into $O(d+1,1)$, that corresponds to the deformation of the flat conformal structure of $\mathbb{H}^d/\Gamma$. At an infinitesimal level, the deformations are parameterized by $H^1(\Gamma, \mathfrak{so}(d+1,1))$. Due to the well-known splitting $\mathfrak{so}(d+1,1) = \mathfrak{so}(d,1) \oplus \mathbb{R}^{d,1}$, we have

$$H^1(\Gamma, \mathfrak{so}(d+1,1)) = H^1(\Gamma, \mathfrak{so}(d,1)) \oplus H^1(\Gamma, \mathbb{R}^{d,1}) ,$$

  but due to the Calabi–Weil infinitesimal rigidity theorem, $H^1(\Gamma, \mathfrak{so}(d,1))$ reduces to 0 [42, 8.10].
- For $d = 2$, $H^1(\Gamma, \mathbb{R}^{2,1})$ is also isomorphic, as a linear space, to the tangent space of the Teichmüller space at (the conjugacy class of) Γ, when we consider the Teichmüller space as the space of discrete, faithful representations of Γ into the isometries of the hyperbolic plane up to conjugacy, see Sect. 16.3.4.
- There is a natural isomorphism between $H^1(\Gamma, \mathbb{R}^{d,1})$ and the space of traceless symmetric Codazzi tensors on $\mathbb{H}^d/\Gamma$ (see Proposition 16.3.17 for a proof using extrinsic co-Minkowski geometry), and the space of traceless symmetric Codazzi tensors parameterizes the space of infinitesimal deformations of the flat conformal structure of $\mathbb{H}^d/\Gamma$, as well as the space of infinitesimal deformations of the Riemannian metric of $\mathbb{H}^d/\Gamma$ preserving the total volume and the harmonicity of the curvature [48].
- $H^1(\Gamma, \mathbb{R}^{d,1})$ parameterizes the space of future complete flat globally hyperbolic maximal Cauchy compact spacetimes (in short, future complete flat GHMC

spacetimes), with Γ as the linear part of the holonomy, see [1, 3, 12, 51] for more details and precise definitions. The universal covers of such spacetimes isometrically embed as convex sets in Minkowski space, whose duals in co-Minkowski space define the convex cores that will be mentioned below, see Remark 16.3.24.

As a consequence of the second point, $H^1(\Gamma, \mathbb{R}^{2,1})$ is a vector space of dimension $(6g-6)$, where g is the genus of $\mathbb{H}^2/\Gamma$. For $d>2$, it is not clear whether $H^1(\Gamma, \mathbb{R}^{d,1})$ is trivial or not. A classical result is that it has dimension at least r if $\mathbb{H}^d/\Gamma$ contains r disjoint embedded totally geodesic hypersurfaces [41, 45, 48]. We give an elementary co-Minkowski proof of this fact in Sect. 16.3.1. See for example [2] and [41] for more information, and [7] for up-to-date references about this question.

The action of Γ_τ, that is Γ deformed by an element τ of $Z^1(\Gamma, \mathbb{R}^{d,1})$, on co-Minkowski space is also interesting in its own. Namely, here too, it is a baby toy model, this time comparing to the study of quasi-Fuchsian hyperbolic manifolds on the one hand, and to AdS GHMC manifolds on the other one (they are the Lorentzian analogues of quasi-Fuchsian hyperbolic manifolds). We will focus on the following aspects. Let $\tau \in Z^1(\Gamma, \mathbb{R}^{d,1})$.

- There exists a smooth hypersurface invariant under the action of Γ_τ. This is a simple illustration of the general "Ehresmann–Weil–Thurston principle", see Proposition 16.3.13.
- The group Γ_τ acts freely and properly discontinuously on co-Minkowski space, and the quotient gives a $(d+1)$-dimensional manifold homeomorphic to $\mathbb{H}^d/\Gamma \times \mathbb{R}$ (see Lemma 16.3.1).
- The co-Minkowski manifold $^*\mathbb{R}^{d,1}/\Gamma_\tau$ has a *convex core*, i.e. it contains a non-empty compact convex set. So the action of Γ_τ on co-Minkowski space is *convex cocompact* in the sense of [22, 23].
- The co-Minkowski manifold $^*\mathbb{R}^{d,1}/\Gamma_\tau$ contains a unique "mean" hypersurface, that is with vanishing mean curvature. This situation is reminiscent of *almost Fuchsian manifolds*, a particular case of quasi-Fuchsian manifolds which contain a unique minimal surface, see [46].
- Moreover, $^*\mathbb{R}^{d,1}/\Gamma_\tau$ is foliated by CMC hypersurfaces, equidistant to the mean hypersurface, see Remark 16.3.16.

We consider that co-Minkowski space is a toy model, because with a pedestrian approach, we are able to give an almost self-contained presentation of the different properties evoked above.

Multi-Dimensional Thurston Earthquake Norm Until this point, all the mentioned results were previously more or less known, at least under the form of dual statements in Minkowski space. Also, the present survey contains the following original contribution.

As we said, the quotient of co-Minkowski space by Γ_τ has a convex core, and a unique mean hypersurface, contained in the convex core. The mean distance between the lower boundary component of the convex core and this mean hypersurface gives a non-negative number, which is uniquely defined by the class in $H^1(\Gamma, \mathbb{R}^{d,1})$ of τ. This gives a map from $H^1(\Gamma, \mathbb{R}^{d,1})$ to $\mathbb{R}_+$, which is actually an asymmetric norm on $H^1(\Gamma, \mathbb{R}^{d,1})$, see Sect. 16.3.3.2. We will call it the S_1 *norm* (see Remark 16.3.26 for the signification of S_1).

The symmetrization of the S_1 norm is:

- the volume of the convex core;[4]
- a "mean distance" between the future complete and the past complete flat GHMC having the same holonomy (see Remark 16.3.24).

In dimension 2, it appears that this asymmetric norm corresponds to the *earthquake norm* introduced by Thurston in [63]. In particular, we obtain a new proof of Theorem 5.2 in [63], saying that the earthquake norm is an asymmetric norm on the tangent space of Teichmüller space. The tangent space of Teichmüller space can be identified with the space of measured geodesic laminations, and the earthquake norm in the total length of the lamination, see Sect. 16.3.4.

In turn, the volume of the convex core is the sum of the total length of the bending laminations of its boundary. Here again, this result should be compared with its more involved analogues in the hyperbolic and anti-de Sitter cases [16, 17].

Using two successive identifications of the tangent space of Teichmüller space with its cotangent space and a formula of Wolpert, the earthquake norm defines another asymmetric norm on the tangent space of Teichmüller space, the *length norm*, see (16.42) for a formula. The length norm defines an asymmetric Finsler structure on Teichmüller space, that in turn defines a distance, now called the *Thurston asymmetric distance*, and introduced by Thurston in [63]. This distance recently attracted attention [54, 55, 65]. Note that the earthquake norm also induces an asymmetric distance on Teichmüller space, but, to the best of our knowledge, nothing is known about this distance.

Anosov Feature In the third and last part of the present survey, we see that co-Minkowski space is also a baby toy model for the theory of Anosov representations, which has known during the recent years, after the pioneering work of F. Labourie [47], a series of development (see [10, 19, 36, 37, 43], see also [6] for a complementary discussion on Anosov representations in the context of Lorentzian geometry, and [32] for a proof of the Anosov character of the representations considered in the present survey).

Once more, it turns out that in the context of co-Minkowski space the theory of Anosov representations reduces to a particularly simple form. Moreover, this point of view provides a proof of the fact that convergence of cocycle implies *uniform* convergence of limit curves (Lemma 16.4.11).

[4]This fact was noted to the first author by Andrea Seppi.

16.2 Co-Minkowski Geometry

Co-Minkowski space is the space of (unoriented) spacelike hyperplanes of Minkowski space. We first investigate the space of oriented spacelike hyperplanes (Sect. 16.2.1). Then we introduce a cylindrical affine model for co-Minkowski space, similar to the Klein ball model of hyperbolic space (Sect. 16.2.2). In the cylindrical model, the co-Minkowski space is the cylinder $B^d \times \mathbb{R}$, where B^d is the open unit ball of $\mathbb{R}^d$ centered at the origin. In particular, extrinsic co-Minkowski geometry of graphs of maps $h : B^d \to \mathbb{R}$ can be investigated (Sect. 16.2.3).

16.2.1 Definition of Co-Minkowski Space

16.2.1.1 Space of Spacelike Hyperplanes

Let us recall that the *Minkowski space* $\mathbb{R}^{d,1}$ of Lorentzian geometry is the affine space $\mathbb{R}^{d+1}$ endowed with the bilinear form

$$\langle x, y\rangle_{d,1} = x_1 y_1 + \cdots + x_d y_d - x_{d+1} y_{d+1} .$$

A hyperplane P of $\mathbb{R}^{d,1}$ is *spacelike* (resp. *timelike*, *lightlike*) if the restriction of $\langle\cdot,\cdot\rangle_{d,1}$ to P is positive-definite (resp. has signature $(+,\ldots,+,-)$, is degenerate). The isometry group of $\mathbb{R}^{d,1}$ is $O(d,1) \ltimes \mathbb{R}^{d,1}$: it is generated by translations and linear transformations preserving $\langle\cdot,\cdot\rangle_{d,1}$.

Linear spacelike hyperplanes are parameterized by the set of *future* unit normal vectors (for $\langle\cdot,\cdot\rangle_{d,1}$):

$$\mathcal{H}^d := \{x \in \mathbb{R}^{d,1} | \langle x, x\rangle_{d,1} = -1, x_{d+1} > 0\} .$$

Let $g_{\mathcal{H}^d}$ be the metric induced by $\langle\cdot,\cdot\rangle_{d,1}$ on the tangent spaces of $\mathcal{H}^d$. It is well known that $(\mathcal{H}^d, g_{\mathcal{H}^d})$ is a model of the d-dimensional hyperbolic space.

Let P be an affine spacelike hyperplane of $\mathbb{R}^{d,1}$. If $n \in \mathcal{H}^d \subset \mathbb{R}^{d,1}$ is the future timelike unit normal to P, then there exists $h \in \mathbb{R}$ such that

$$P = \{y \in \mathbb{R}^{d,1} | \langle y, n\rangle_{d,1} = h\} .$$

This defines a point

$$\tilde{P}^* = (n, h)$$

in $\mathbb{R}^{d+1} \times \mathbb{R} = \mathbb{R}^{d+2}$. More precisely, the point $\tilde{P}^*$ belongs to one of the connected components of the degenerate quadric

$$co\mathcal{M}in^{d+1} := \{x \in \mathbb{R}^{d+2} | \langle x, x\rangle_{d,1,0} = -1\}$$

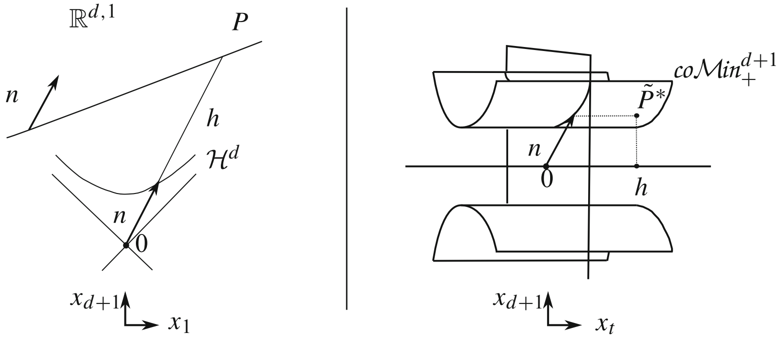


Fig. 16.2 The dual of a spacelike hyperplane of Minkowski space $\mathbb{R}^{d,1}$ in $co\mathcal{M}in^{d+1}$. In the picture, $d = 1$

where

$$\langle (x_1, \ldots, x_{d+1}, x_t), (y_1, \ldots, y_{d+1}, y_t) \rangle_{d,1,0} = x_1 y_1 + \cdots + x_d y_d - x_{d+1} y_{d+1} .$$

Later we will identify both components of the quadric, and the point $\tilde{P}^*$ will correspond to a point P^*, that explains the $\tilde{}$ for our current notation. Note that $co\mathcal{M}in^{d+1}$ is the space of oriented spacelike hyperplanes of Minkowski space. See Fig. 16.2.

We will denote by $g_{co\mathcal{M}^{d+1}}$ the degenerate $(0, 2)$-tensor induced by $\langle \cdot, \cdot \rangle_{d,1,0}$ on the tangents spaces of $co\mathcal{M}in^{d+1}$. The connected component $co\mathcal{M}in_+^{d+1} = co\mathcal{M}in^{d+1} \cap \{x_{d+1} > 0\}$ of $co\mathcal{M}in^{d+1}$ containing the point $\tilde{P}^*$ is homeomorphic to $\mathcal{H}^d \times \mathbb{R}$. This leads to a fibration:

$$\pi : co\mathcal{M}in_+^{d+1} \to \mathcal{H}^d$$

mapping $(x_1, \ldots, x_{d+1}, x_t)$ to $(x_1, \ldots, x_{d+1})$. It is a principal $\mathbb{R}$-bundle; it is an isometry, and the fibers are precisely tangent to the kernel of the degenerate metric $g_{co\mathcal{M}^{d+1}}$, and

$$g_{co\mathcal{M}^{d+1}} = \pi^* g_{\mathcal{H}^d} .$$

16.2.1.2 Isometries

As the "metric" $g_{co\mathcal{M}^{d+1}}$ is degenerate, it will be more relevant to consider a group acting on $co\mathcal{M}in^{d+1}$. As an isometry of Minkowski space sends spacelike hyperplanes onto spacelike hyperplanes, it acts naturally on $co\mathcal{M}in^{d+1}$. This is the way we define the isometry group of $co\mathcal{M}in^{d+1}$. More precisely, it is immediate that if

$$P = \{y \in \mathbb{R}^{d,1} | \langle y, n \rangle_{d,1} = h\}$$

is a spacelike hyperplane of $\mathbb{R}^{d,1}$ and $A \in O(d,1)$, so that $\tilde{P}^* = (n, h)$, then

$$\widetilde{AP}^* = (An, h)$$

and if $v \in \mathbb{R}^{d,1}$,

$$\widetilde{P+v}^* = (n, \langle v, n\rangle_{d,1} + h) \ .$$

So $\mathrm{O}(d,1) \ltimes \mathbb{R}^{d,1}$ acts linearly on $co\mathcal{M}in^{d+1}$ via the representation

$$(A, v) \mapsto \left(\begin{array}{c|c} A & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline {}^t v J A & 1 \end{array}\right) , \tag{16.1}$$

where $J = \mathrm{diag}(1, \dots, 1, -1)$ (recall that $A \in O(d,1)$ if and only if ${}^t A = J A^{-1} J$). So we define the *isometry group* of $co\mathcal{M}in^{d+1}$ as $O(d,1) \ltimes \mathbb{R}^{d,1}$ with the action on $\mathbb{R}^{d+2}$ induced by the representation (16.1). In particular, the group structure on $O(d,1) \ltimes \mathbb{R}^{d,1}$ is

$$(A_1, v_1) \cdot (A_2, v_2) = (A_1 A_2, v_1 + A_1 v_2) \ . \tag{16.2}$$

Remark 16.2.1 Let $O_+(d,1)$ be the subgroup of $O(d,1)$ preserving $\mathcal{H}^d$. Then, $O_+(d,1) \ltimes \mathbb{R}^{d,1}$ preserves the connected component $co\mathcal{M}in_+^{d+1}$, and the fibration π is $O_+(d,1) \ltimes \mathbb{R}^{d,1}$-equivariant. The elements of $O_+(d,1) \ltimes \mathbb{R}^{d,1}$ inducing the identity map on $\mathcal{H}^d$ are precisely the translations (elements of $\mathbb{R}^{d,1}$).

Every fiber of π admits a natural Euclidean structure, for which they are individually isomorphic to the real line. The action of $O_+(d,1) \ltimes \mathbb{R}^{d,1}$ preserves this Euclidean structure along the fibers. Indeed, such a fiber is a collection of parallel spacelike hyperplanes with the same direction n, and one can define the "distance" between two elements P, P' of the same fiber as the proper time of any timelike segment orthogonal to P and P', and with extremities in P, P'.

Remark 16.2.2 The isometry group of $co\mathcal{M}in^{d+1}$ is smaller than the group of transformations preserving the degenerate metric $g_{co\mathcal{M}^{d+1}}$. For example, for $c > 0$, the map $H_c : \mathbb{R}^{d+2} \to \mathbb{R}^{d+2}$, $H_c(x) = (x_1, x_2, \dots, x_{d+1}, c x_t)$, preserves $\langle \cdot, \cdot \rangle_{d,1,0}$ (hence it preserves $co\mathcal{M}in^{d+1}$ and $g_{co\mathcal{M}^{d+1}}$), but by definition it is not an isometry of $co\mathcal{M}in^{d+1}$.

There does not exist any (non-degenerate) semi-Riemannian metric on $co\mathcal{M}in^{d+1}$ invariant under the isometry group of $co\mathcal{M}in^{d+1}$ [28, Fact 2.27].

16.2.1.3 Connection, Geodesics

We have now the hypersurface $co\mathcal{M}in^{d+1}$ in $\mathbb{R}^{d+2}$ together with an "isometry group" and a degenerate metric $g_{co\mathcal{M}^{d+1}}$. As those elements arise from the degenerate form $\langle\cdot,\cdot\rangle_{d,1,0}$ on the ambient $\mathbb{R}^{d+2}$, there is no obvious metric notion of "unit normal vector" to $co\mathcal{M}in^{d+1}$. Nevertheless, we can proceed similarly to classical affine differential geometry [52]. Namely, at a point $x \in co\mathcal{M}in^{d+1}$, let us define as a "normal field" the vector field $\mathrm{N}(x) = x$. Obviously, N is transverse to $co\mathcal{M}in^{d+1}$ and invariant under the group of isometries of $co\mathcal{M}in^{d+1}$. The choice of this normal field allows one to define a connection $\nabla^{co\mathcal{M}^{d+1}}$ on $co\mathcal{M}in^{d+1}$ induced by the canonical connection D of the ambient linear space $\mathbb{R}^{d+2}$:

$$D_Y X = \nabla_Y^{co\mathcal{M}^{d+1}} X + \langle X, Y\rangle_{d,1,0}\,\mathrm{N}\ .$$

The following facts are easily checked, see [28, Section 4.2].

Fact 16.2.3 *The connection $\nabla^{co\mathcal{M}^{d+1}}$ has the following properties:*

- *it is torsion free,*
- *compatible with the degenerate metric $g_{co\mathcal{M}^{d+1}}$,*
- *invariant under isometries,*
- *its (unparameterized) geodesics are intersection of $co\mathcal{M}in^{d+1}$ with linear planes of $\mathbb{R}^{d+2}$.*

It follows from the last point that the intersection of $co\mathcal{M}in^{d+1}$ with linear k-planes of $\mathbb{R}^{d+2}$ are totally geodesic. Those intersections will play a fundamental role as the following fact shows.

Fact 16.2.4 *The intersection of $co\mathcal{M}in^{d+1}$ with a linear k-planes of $\mathbb{R}^{d+2}$ transverse to the degenerate direction is isometric (for the metric induced by $g_{co\mathcal{M}^{d+1}}$) to the hyperbolic space of dimension k.*

Moreover, $\nabla^{co\mathcal{M}^{d+1}}$ coincides with the Levi-Civita connection of the hyperbolic metric on any such subspace.

Proof Immediate as one can always find an isometry of $co\mathcal{M}in^{d+1}$ sending a linear k-plane to a linear k-plane contained in $\{x_t = 0\}$. □

16.2.1.4 Co-Minkowski Space

The co-Minkowski space is the space of unoriented spacelike hyperplanes of Minkowski space, that is, the quotient of $co\mathcal{M}in^{d+1}$ by the antipodal map.

Definition 16.2.5 The *co-Minkowski space* $^*\mathbb{R}^{d,1}$ is the following subspace of the projective space: $^*\mathbb{R}^{d,1} = co\mathcal{M}in^{d+1}/\{\pm \mathrm{Id}\}$, endowed with the push-forward of the degenerate metric $g_{co\mathcal{M}^{d+1}}$, denoted by $g_{^*\mathbb{R}^{d,1}}$.

The connection $\nabla^{co\mathcal{M}^{d+1}}$ also induces a connection $\nabla^{*\mathbb{R}^{d,1}}$ on $*\mathbb{R}^{d,1}$.

We define the *isometry group* of $*\mathbb{R}^{d,1}$ as the image of $\mathrm{O}(d,1) \ltimes \mathbb{R}^{d,1}$ into $\mathrm{PGL}(d+2)$, by a projective quotient of the representation given by (16.1).

The map $\pi : co\mathcal{M}in_{+}^{d+1} \to \mathcal{H}^d$ induces an $\mathbb{R}$-fibration $*\pi : *\mathbb{R}^{d,1} \to \mathcal{H}^d$, which is an isometry, and is equivariant for the action of the projectivization of $\mathrm{O}(d,1) \ltimes \mathbb{R}^{d,1}$.

In will be interesting to work in a particular affine model of co-Minkowski space. This will be the cylindrical coordinates introduced in the next section.

16.2.2 Cylindrical Model

16.2.2.1 Klein Ball Model of the Hyperbolic Space

We have seen that the subspace $\mathcal{H}^d$ of Minkowski space, endowed with the induced metric, is a model of the hyperbolic space. It is isometric to the subset $\{x \in \mathbb{R}^{d,1} | \langle x, x\rangle_{d,1} < 0\}$ of the projective space $\mathbb{P}(\mathbb{R}^{d,1})$ endowed with the push-forward metric.

The *Klein ball model* of the hyperbolic space is the image of the projective model of the hyperbolic space in the affine chart $\{x_{d+1} = 1\}$. As a set, it is the open Euclidean unit ball B^d. The push-forward of the hyperbolic metric on B^d is denoted by $g_{\mathbb{H}^d}$. We will sometimes use the notation $\mathbb{H}^d$ to designate the hyperbolic space $(B^d, g_{\mathbb{H}^d})$. In the remainder of this section, we give explicit formulas relating the hyperbolic geometry on B^d to the standard Euclidean geometry on B^d, that will be needed in the sequel.

If $x \in B^d$, then the vector $\binom{x}{1}$ of $\mathbb{R}^{d,1}$ rescaled by the factor $L^{-1}(x)$ belongs to $\mathcal{H}^d$, where

$$L(x) = \sqrt{1 - \|x\|^2}$$

and $\|\cdot\|$ is the Euclidean norm on B^d:

$$\|(x_1, \ldots, x_d)\| = \sqrt{x_1^2 + \cdots + x_d^2}\,,$$

see Fig. 16.3.

The expression of the hyperbolic metric $g_{\mathbb{H}^d}$ in the Klein ball model is:

$$g_{\mathbb{H}^d}(x)(X, Y) = L(x)^{-2}\langle X, Y\rangle_d + L(x)^{-4}\langle x, X\rangle_d \langle x, Y\rangle_d \tag{16.3}$$

where $\langle\cdot,\cdot\rangle_d$ is the standard Euclidean metric on $\mathbb{R}^d \supset B^d$, $x \in B^d$, $X, Y \in T_x B^d \cong \mathbb{R}^d$. In order to help computations, one may note that

$$D_X L^{-1}(x) = L^{-3}(x)\langle x, X\rangle_d \tag{16.4}$$

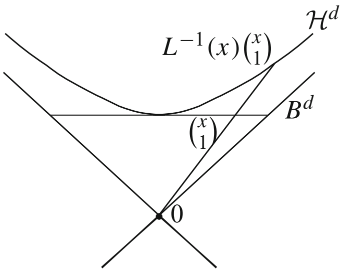


Fig. 16.3 The hyperboloid $\mathcal{H}^d$ and the Klein ball model of the hyperbolic space

and

$$\text{Hess}\, L = -L g_{\mathbb{H}^d} \,, \tag{16.5}$$

where Hess is the usual Hessian on $\mathbb{R}^d$.

If ω_{B^d} if the restriction to B^d of the Euclidean volume form, and $\omega_{\mathbb{H}^d}$ is the volume form on B^d associated to the hyperbolic metric $g_{\mathbb{H}^d}$, from (16.3) one obtains

$$\omega_{B^d} = L^{d+1}\omega_{\mathbb{H}^d} \,. \tag{16.6}$$

The main feature of the Klein ball model of the hyperbolic space is that the (unparameterized) geodesics of $g_{\mathbb{H}^d}$ are exactly the affine segments in B^d. This is straightforward, as the geodesics of $\mathcal{H}^d$ are the intersections of $\mathcal{H}^d$ with linear timelike planes of $\mathbb{R}^{d,1}$. This gives the following correspondence between the connections, see [28, Lemma 4.17].

Proposition 16.2.6 (Weyl Formula) *If $\nabla^{\mathbb{H}^d}$ is the Levi-Civita connection of $g_{\mathbb{H}^d}$ and D is the canonical connection on B^d, then*

$$\nabla^{\mathbb{H}^d}_X Y = D_X Y + L^{-2}(x)(\langle x, X\rangle_d Y + \langle x, Y\rangle_d X) \,. \tag{16.7}$$

Corollary 16.2.7 *If* $\text{Hess}^{\mathbb{H}^d}$ *is the Hessian given by* $\nabla^{\mathbb{H}^d}$*, then, for a smooth map* $f : B^d \to \mathbb{R}$*,*

$$\text{Hess}^{\mathbb{H}^d} f(x)(X,Y) = \text{Hess}\, f(x)(X,Y) - L^{-2}(x)(\langle x, X\rangle_d\, \mathrm{d}f(x)(Y) + \langle x, Y\rangle_d\, \mathrm{d}f(x)(Y)) \,. \tag{16.8}$$

Also,

$$L^{-1}(x)\text{Hess}\, f(x)(X,Y) = \left(\text{Hess}^{\mathbb{H}^d}(L^{-1}f)(x)(X,Y) - (L^{-1}f)(x) g_{\mathbb{H}^d}(x)(X,Y)\right) \,. \tag{16.9}$$

Proof (16.8) follows from (16.7) and

$$\mathrm{Hess}^{\mathbb{H}^d} f(x)(X, Y) = X.Y.f(x) - \mathrm{d} f(x)(\nabla^{\mathbb{H}^d}_X Y)\ . \tag{16.10}$$

Finally, (16.9) comes from (16.8), (16.3) and

$$\mathrm{Hess}\ fg = f\mathrm{Hess}\ g + g\mathrm{Hess}\ f + \mathrm{d} f \otimes \mathrm{d} g + \mathrm{d} g \otimes \mathrm{d} f\ . \tag{16.11}$$

□

Fact 16.2.8 *If Δ is the Euclidean Laplacian on B^d, then*

$$\mathrm{tr}_{g_{\mathbb{H}^d}} \mathrm{Hess}\ f(x) = L^2(x)(\Delta f - \mathrm{Hess}\ f(x)(x, x))\ . \tag{16.12}$$

If $\Delta^{\mathbb{H}^d}$ is the Laplacian on B^d given by $g_{\mathbb{H}^d}$, then

$$\mathrm{tr}_{g_{\mathbb{H}^d}} L^{-1}\mathrm{Hess}\ f = \Delta^{\mathbb{H}^d}(L^{-1} f) - d(L^{-1} f)\ . \tag{16.13}$$

Proof Let A be the linear operator such that $\mathrm{Hess}\ f(x)(X, Y) = g_{\mathbb{H}^d}(x)(AX, Y)$. For $x \neq 0$, let $(e_i)_{1,\cdots,d}$ be an orthonormal Euclidean basis of $T_x B^d$, such that $e_1 = x/\|x\|$. The definition of A and (16.3) give, for $i > 1$,

$$\langle Ae_i, e_i\rangle_d = L^2(x) g_{\mathbb{H}^d}(x)(Ae_i, e_i) = L^2(x)\,\mathrm{Hess}\ f(x)(e_i, e_i)\ ,$$

and

$$\langle Ae_1, e_1\rangle_d = L^2(x)\,\mathrm{Hess}\, h(x)(e_1, e_1) + L^{-2}(x)\langle x, Ax\rangle_d\ .$$

Also from the definition of A and (16.3),

$$L^{-2}(x)\langle x, Ax\rangle_d = L^2(x) g_{\mathbb{H}^d}(x)(x, Ax) = L^2(x)\,\mathrm{Hess}\ f(x)(x, x)\ .$$

(16.12) follows from $\mathrm{tr}_{g_{\mathbb{H}^d}} \mathrm{Hess}\ f(x) = \sum_{i=1}^d \langle Ae_i, e_i\rangle_d$. Also, (16.13) is immediate from (16.9). □

Let us end this section with some basic facts about (smooth) hyperbolic Codazzi tensors.

Definition 16.2.9 A $(0, 2)$-tensor C on $\mathbb{H}^d$ is a *(hyperbolic) Codazzi tensor* if it satisfies the the Codazzi equation on $\mathbb{H}^d$:

$$\nabla^{\mathbb{H}^d}_X C(Y, Z) = \nabla^{\mathbb{H}^d}_Y C(X, Z)\ .$$

Lemma 16.2.10 *Let C be a $(0,2)$-tensor on B^d. Then C is a hyperbolic Codazzi tensor if and only if*

$$D_X(LC)(Y,Z) = D_Y(LC)(X,Z)\ .$$

Proof The definition of Codazzi tensor means that

$$X.C(Y,Z) - C(\nabla^{\mathbb{H}^d}_X Y, Z) - C(Y, \nabla^{\mathbb{H}^d}_X Z) = Y.C(X,Z) - C(\nabla^{\mathbb{H}^d}_Y X, Z) - C(X, \nabla^{\mathbb{H}^d}_Y Z)\ .$$

Developing this expression using (16.7), one obtains, at a point x,

$$D_X C(x)(Y,Z) - L^{-2}(x)\langle x, X\rangle_d C(Y,Z) = D_Y C(x)(X,Z) - L^{-2}(x)\langle x, Y\rangle_d C(X,Z)\ .$$

Writing $C = L^{-1}LC$, developing the above expression and using (16.4) leads to the result. □

Fact 16.2.11 *Let S be a $(0,2)$-tensor on B^d. If $D_X S(Y,Z) = D_Y S(X,Z)$, then there exists a function $F = (F_1, \dots, F_n)$ with $F_i : B^d \to \mathbb{R}$ such that S is the Jacobian matrix of F.*

Proof Let $\Omega_j = \sum_{i=1}^d S_{ij} \mathrm{d}x^i$. As $\frac{\partial S_{ij}}{\partial x_k} = \frac{\partial S_{kj}}{\partial x_i}$, $\mathrm{d}\Omega_j = 0$, so by the Poincaré Lemma, there exists a function $F_j : B^d \to \mathbb{R}$ such that $\mathrm{d}F_j = \Omega_j$. □

Fact 16.2.12 *Let $F = (F_1, \dots, F_d)$ with $F_j : B^d \to \mathbb{R}$. Then there exists $f : B^d \to \mathbb{R}$ with $\frac{\partial f}{\partial x_i} = F_i$ if and only if $\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$.*

In other term, the Jacobian matrix of F is a Hessian matrix (namely the one of f) if and only if it is a symmetric matrix.

Proof One implication is Schwarz's theorem. On the other direction, the one-form $\omega = \sum_{i=1}^d F_i \mathrm{d}x^i$ is closed by hypothesis, hence exact by the Poincaré Lemma, and it suffices to set $\omega = \mathrm{d}f$. □

We finally obtain the following classical result [9, 15, 53].

Lemma 16.2.13 *Let C be a $(0,2)$-tensor on B^d. Then C is a symmetric hyperbolic Codazzi tensor if and only if there exists $f : B^d \to \mathbb{R}$ such that*

$$C = L^{-1}\mathrm{Hess}\ f\ .$$

16.2.2.2 Affine Representation of Co-Minkowski Space

To keep track of some relevant affine notions such as convexity, we will work in an affine model of co-Minkowski space. Namely, we will consider the affine model of co-Minkowski space given by the central projection of $co\mathcal{M}in^{d+1}_+$ onto the hyperplane $\{x_{d+1} = 1\}$ of $\mathbb{R}^{d+2}$. Observe that in doing so, we favor the coordinate x_{d+1}, i.e. we distinguish the future timelike vector $(0, \ldots, 0, 1)$ of $\mathbb{R}^{d,1}$. We will come back to this remark in Sect. 16.4. In the hyperplane $\{x_{d+1} = 1\}$, the image of $co\mathcal{M}in^{d+1}$ is the cylinder $B^d \times \mathbb{R}$, where B^d is the open unit ball centered at the origin of $\mathbb{R}^d$.

We denote by $\pi : B^d \times \mathbb{R} \to B^d$ the projection on the first factor. It corresponds to the fibration $\pi : co\mathcal{M}in_+^{d+1} \to \mathcal{H}^d$. We will call *vertical lines* the fibers of π. They correspond to parallel spacelike hyperplanes in Minkowski space.

Remark 16.2.14 In these coordinates $B^d \times \mathbb{R} \subset \mathbb{R}^{d+1}$, the degenerate metric $g_{*\mathbb{R}^{d,1}}$ of co-Minkowski space is $g_{\mathbb{H}^d} + 0\,\mathrm{d}x_t^2$. The degenerate metric $g_{*\mathbb{R}^{d,1}}$ defines a "distance" between points of co-Minkowski space. Actually this distance is nothing but the Klein projective metric: if $x, y \in B^d \times \mathbb{R}$, then they are on a line meeting $\partial B^d \times \mathbb{R} \cup \{\infty\}$ either at two distinct points I, J, or at $I = J = \infty$. Then the Klein projective distance is $d(x, y) = \frac{1}{2}|\ln[x, y, I, J]|$, where $[\cdot, \cdot, \cdot, \cdot]$ is the cross-ratio, see [28].

Remark 16.2.15 The *boundary at infinity* of co-Minkowski space is $\partial B^d \times \mathbb{R}$. It parameterizes the set of lightlike affine hyperplanes of Minkowski space, and it is called *Penrose boundary* in [3]. Note that $(\mathbb{R}^d \setminus \bar{B}^d) \times \mathbb{R}$ parameterize the set of affine timelike hyperplanes of Minkowski space, but we don't need to consider it.

The interest of an affine model is essentially given by the following facts. The first one is an immediate consequence of the last point of Fact 16.2.3.

Fact 16.2.16 (Unparameterized) *geodesics of* $^*\mathbb{R}^{d,1}$ *in the cylindrical model* $B^d \times \mathbb{R}$ *are (affine) geodesic segments.*

The second fact follows from Fact 16.2.4 and by construction.

Fact 16.2.17 *The intersection of* $B^d \times \mathbb{R}$ *with any affine* k*-plane not containing a vertical line, with the metric induced by* $g_{*\mathbb{R}^{d,1}}$*, is isometric to the hyperbolic space of dimension* k.

In particular, $B^d \times \{0\} \cong B^d$ *is the Klein ball model of the* d*-dimensional hyperbolic space.*

When $k = d$, we will call the intersection of $B^d \times \mathbb{R}$ with a d-plane not containing a vertical line a *hyperbolic hyperplane*.

Remark 16.2.18 As every non-degenerate tangent plane of co-Minkowski space is isometric to the tangent plane of a hyperbolic space, the sectional curvature of co-Minkowski space is -1.

16.2.2.3 Duality

This cylindrical affine model can be directly described from Minkowski space as follows. Let P be an affine spacelike hyperplane of $\mathbb{R}^{d,1}$, and let $(x, 1)$ be a normal

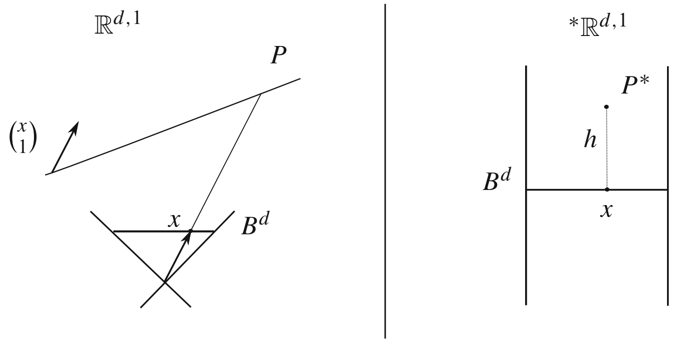


Fig. 16.4 The dual P^* of the hyperplane $P = \{y \in \mathbb{R}^{d,1} | \langle \binom{x}{1}, y \rangle_{d,1} = h\}$

vector, with $x \in B^d$. Then there exists a number h such that

$$P = \{y \in \mathbb{R}^{d,1} | \langle \begin{pmatrix} x \\ 1 \end{pmatrix}, y \rangle_{d,1} = h\}$$

and P defines a point $P^* = (x, h) \in B^d \times \mathbb{R}$, see Fig. 16.4.

Let us give more precisions about the "duality" between Minkowski space and co-Minkowski space. We already know that if P is a spacelike hyperplane of Minkowski space, then P^* is a point in $^*\mathbb{R}^{d,1}$. Conversely, if P is a hyperbolic hyperplane of $^*\mathbb{R}^{d,1}$, let P^* be the intersection of all the hyperplanes of Minkowski space whose duals are points in P. For future reference, let us express this fact in terms of the cylindrical coordinates $B^d \times \mathbb{R}$.

Fact 16.2.19 *Let P be a hyperbolic hyperplane of co-Minkowski space, which is the graph of the affine function $h : B^d \to \mathbb{R}$, $h(x) = \langle x, v \rangle_d + c$. Then the point P^* dual to P has coordinates $P^* = (v, -c) \in \mathbb{R}^d \times \mathbb{R} = \mathbb{R}^{d,1}$.*

In other terms, if Q is a point of Minkowski space, then the hyperplane Q^* in co-Minkowski space is the graph of the affine map $h : B^d \to \mathbb{R}$, $h(x) = \langle Q, \binom{x}{1} \rangle_{d,1}$.

Proof Let us fix $x \in B^d$. Then the point $X = (x, \langle v, x \rangle_d + c) \in B^d \times \mathbb{R}$ of co-Minkowski space belongs to P. Its dual is the spacelike hyperplane of Minkowski space defined as

$$X^* = \{(y, y_{d+1}) \in \mathbb{R}^d \times \mathbb{R} | \langle \begin{pmatrix} x \\ 1 \end{pmatrix}, \begin{pmatrix} y \\ y_{d+1} \end{pmatrix} \rangle_{d,1} = \langle v, x \rangle_d + c\} ,$$

i.e., $X^* = \{(y, y_{d+1}) \in \mathbb{R}^d \times \mathbb{R} | \langle \binom{x}{1}, \binom{y}{y_{d+1}} \rangle_{d,1} = \langle \binom{x}{1}, \binom{v}{-c} \rangle_{d,1}\}$ and obviously $(v, -c)$ belongs to this hyperplane. As x was arbitrary, $(v, -c)$ belongs to all the hyperplanes dual to the points of P, which is the definition of P^*. □

The proof of the following facts are left to the reader.

Fact 16.2.20

1. *If P is a hyperbolic hyperplane in co-Minkowski space ${}^*\mathbb{R}^{d,1}$, then P^* is a point in Minkowski space $\mathbb{R}^{d,1}$ and $(P^*)^* = P$.*
2. *Let P and Q be two hyperbolic hyperplanes in ${}^*\mathbb{R}^{d,1}$.*
   (a) *if P and Q meet in ${}^*\mathbb{R}^{d,1}$ then P^* and Q^* are joined by a spacelike segment in $\mathbb{R}^{d,1}$.*
   (b) *if P is strictly above Q in $\bar{B}^d \times \mathbb{R}$, then $Q^* - P^*$ is a future directed timelike segment in $\mathbb{R}^{d,1}$.*
   (c) *if P and Q have a common point in $\partial B^d \times \mathbb{R}$, then P^* and Q^* are joined by a lightlike segment.*

The vector space structure of Minkowski space corresponds via duality to the vector space structure on the space of restrictions to B^d of affine maps.

Fact 16.2.21 *Let h_Q and h_P be the restrictions to B^d of affine maps, such that their graphs are the hyperbolic hyperplanes P, Q of co-Minkowski space, and let $\lambda \in \mathbb{R}$. Then the graph of $h_P + \lambda h_Q$ is dual to the point $P^* + \lambda Q^*$ of Minkowski space.*

Remark 16.2.22 A convex spacelike hypersurface S of Minkowski space is the boundary of the intersection of half-spaces bounded by spacelike hyperplanes. A hypersurface is *F-convex* if it is the boundary of a spacelike convex hypersurface such that any spacelike vector hyperplane is the direction of a support plane, and if the surface is in the future side of its support planes. Each support plane P has a normal vector of the form $\binom{x}{1}$ for $x \in B^d$, so there is $h(x) \in \mathbb{R}$ such that

$$P = \{y | \langle y, \begin{pmatrix} x \\ 1 \end{pmatrix} \rangle_{d,1} = h(x)\} .$$

The graph S^* of the function h in $B^d \times \mathbb{R}$ is actually a convex hypersurface, see [13, 30]. In more classical terms, h is the *support function* of the convex set K bounded by S:

$$h(x) = \max_{k \in K} \langle \begin{pmatrix} x \\ 1 \end{pmatrix}, k \rangle_{d,1} . \tag{16.14}$$

Let us suppose furthermore that S is the graph of a function $f : \mathbb{R}^d \to \mathbb{R}$. Then if $k \in K$ there is $y \in \mathbb{R}^d$ such that $k = \binom{y}{f(y)}$, and from (16.14),

$$h(x) = \max_{y \in \mathbb{R}^d} \langle \begin{pmatrix} x \\ 1 \end{pmatrix}, \begin{pmatrix} y \\ f(y) \end{pmatrix} \rangle_{d,1} = \max_{y \in \mathbb{R}^d} \{\langle x, y \rangle_d - f(y)\} ,$$

i.e. h is nothing but the conjugate (Legendre–Fenchel dual) of f.

In the same way, convex hypersurfaces of Minkowski space which are in the past side of their support planes have dual hypersurfaces in the cylindrical model of co-Minkowski space, which are graphs of concave function $h : B^d \to \mathbb{R}$.

Example 16.2.23 *The dual surface of the hyperboloid* $\{y | \langle y, y\rangle_{d,1} = -t^2, y_{d+1} > 0\}$ *is the graph of the function* $B^d \to \mathbb{R}$, $x \mapsto -tL(x)$. *Note that this function is convex (see (16.5)). In the same way, the dual surface of the hyperboloid* $\{y | \langle y, y\rangle_{d,1} = -t^2, y_{d+1} < 0\}$ *is the graph of the concave function* $h(x) = tL(x)$.

Remark 16.2.24 Any hypersurface in Minkowski space which is an envelope of spacelike hyperplanes has a dual hypersurface in co-Minkowski space. This is more easily seen in the other way. For any C^2 function $h : B^d \to \mathbb{R}$, there exists a map $\chi : B^d \to \mathbb{R}^{d,1}$, the *normal representation*, such that $P = \{y | \langle y, \binom{x}{1}\rangle_{d,1} = h(x)\}$ is tangent to $\chi(B^d)$ at the point $\chi(x)$, see [30, 2.12]. Notice that χ is in general not a regular map, and that the concept of tangent hyperplane has to be understood in a generalized sense. The simplest example is when h is the restriction to B^d of an affine map: its graph is a hyperplane P in the cylindrical model $B^d \times \mathbb{R}$ of co-Minkowski space, and $\chi(B^d)$ is reduced to a point, the dual point of P in Minkowski space.

Remark 16.2.25 The duality between $\mathbb{R}^{d,1}$ and $^*\mathbb{R}^{d,1}$ can also be seen in $\mathbb{R}^{d+2}$, looking at $\mathbb{R}^{d,1}$ as a degenerate quadric in $\mathbb{R}^{d+2}$. See [28, Section 2.5] for more details.

16.2.2.4 Isometries in Cylindrical Coordinates

Let us write the action of the isometry group of co-Minkowski space in the cylindrical coordinates $B^d \times \mathbb{R}$. First let us state some facts about the action of hyperbolic isometries on B^d. The group $O_+(d, 1)$ acts by isometries on the hyperbolic space $\mathcal{H}^d$, and hence on the Klein ball model. More precisely, let $x \in B^d$ and $A \in O_+(d, 1)$. We will denote by $A \cdot x$ the image of x by the isometry of the Klein ball model defined by A. We have

$$\frac{1}{\left(A\binom{x}{1}\right)_{d+1}} A\begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} A \cdot x \\ 1 \end{pmatrix}. \tag{16.15}$$

Note that as A is a linear isometry of Minkowski space $\mathbb{R}^{d,1}$, we have

$$|\left(A\binom{x}{1}\right)_{d+1}|^2(\|A \cdot x\|^2 - 1) = \|x\|^2 - 1 ,$$

i.e.,

$$\left(A\binom{x}{1}\right)_{d+1} = \frac{L(x)}{L(A \cdot x)} , \tag{16.16}$$

so, together with (16.15), one obtains

$$A\begin{pmatrix} x \\ 1 \end{pmatrix} = \frac{L(x)}{L(A\cdot x)}\begin{pmatrix} A\cdot x \\ 1 \end{pmatrix}. \tag{16.17}$$

For simplicity, let us fix also the following coordinate system; every element $(x_1, \dots, x_{d+1})$ of $\mathbb{R}^{d+1}$ has a horizontal component $\bar{x} = (x_1, \dots, x_d)$ and a vertical component x_{d+1}. If $\langle \bar{x}, \bar{y}\rangle_d$ is the scalar product of horizontal elements, we have, for $x, y \in \mathbb{R}^{d,1}$, $\langle x, y\rangle_{d,1} = \langle \bar{x}, \bar{y}_d\rangle_d - x_{d+1}y_{d+1}$.

Lemma 16.2.26 *Let $(x, h) \in B^d \times \mathbb{R}$ and $(A, v) \in O_+(d, 1) \ltimes \mathbb{R}^{d,1}$. Then the isometry of co-Minkowski space defined by (A, v) acts on the cylindrical coordinates as follows:*

$$(A, v)(x, h) = \left(A\cdot x, \frac{L(A\cdot x)}{L(x)}h + \langle A\cdot x, \bar{v}\rangle_d - v_{d+1}\right). \tag{16.18}$$

Proof When the isometry is linear, i.e. when $v = 0$, the elements of the image of (x, h) by (A, v) are elements of $\mathbb{R}^{d,1}$ satisfying:

$$\begin{aligned} h &= \langle \begin{pmatrix} x \\ 1 \end{pmatrix}, A^{-1}\begin{pmatrix} y \\ y_{d+1} \end{pmatrix}\rangle_{d,1} \\ &= \langle A\begin{pmatrix} x \\ 1 \end{pmatrix}, \begin{pmatrix} y \\ y_{d+1} \end{pmatrix}\rangle_{d,1} \\ &\overset{(16.17)}{=} \langle \frac{L(x)}{L(A\cdot x)}\begin{pmatrix} A.x \\ 1 \end{pmatrix}, \begin{pmatrix} y \\ y_{d+1} \end{pmatrix}\rangle_{d,1}. \end{aligned}$$

Therefore, the image of (x, h) by $(A, 0)$ is $(A\cdot x, \frac{L(A\cdot x)}{L(x)}h)$.

In the case of a translation by a vector $v = \binom{\bar{v}}{v_d}$ we have:

$$\begin{aligned} h &= \langle \begin{pmatrix} x \\ 1 \end{pmatrix}, \begin{pmatrix} y \\ y_{d+1} \end{pmatrix} - \begin{pmatrix} \bar{v} \\ v_{d+1} \end{pmatrix}\rangle_{d,1} \\ &= \langle \begin{pmatrix} x \\ 1 \end{pmatrix}, \begin{pmatrix} y \\ y_{d+1} \end{pmatrix}\rangle_{d,1} - \langle x, \bar{v}\rangle_d + v_{d+1}. \end{aligned}$$

Hence the image of (x, h) by the translation is

$$(x, h + \langle x, \bar{v}\rangle_d - v_{d+1}).$$

The Lemma follows because from (16.2), $(A, v) = (\mathrm{Id}, v)(A, 0)$. □

Remark 16.2.27 There is an easy way to see the action of $O_+(d, 1)$ in the coordinates $B^d \times \mathbb{R}$. Actually, $B^d \times \mathbb{R}$ is foliated by the graphs of the functions tL, $t \in \mathbb{R}$. Note that these graphs are, for $t \neq 0$, the duals of the two-sheeted

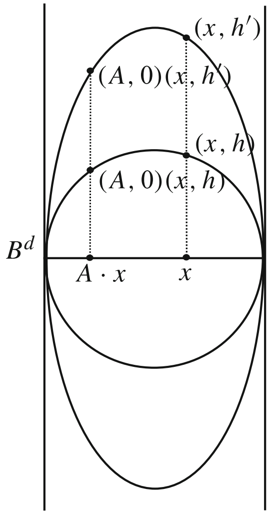


Fig. 16.5 Action of $(A, 0)$ on $B^d \times \mathbb{R}$

hyperboloids centered at the origin in Minkowski space, see Example 16.2.23. Observe that for the sheet with positive (respectively negative) x_{d+1}, the parameter t is negative (respectively positive). Hence if $(x, h) \in B^d \times \mathbb{R}$ belongs to the graph of tL for some t, then for any $A \in O_+(d, 1)$, $(A, 0)(x, h)$ still belongs to the graph of tL, and of course its projection onto $B^d \times \{0\}$ is $(A \cdot x, 0)$, see Fig. 16.5.

Remark 16.2.28 In order to fully understand the action of $O(d, 1)$ on co-Minkowski space, we have to describe the action of $-\operatorname{Id} \in O(d, 1)$ on $B^d \times \mathbb{R}$. It is actually straightforward that

$$(-\operatorname{Id}, 0)(x, h) = (x, -h) \ . \tag{16.19}$$

We now describe the action of the isometries of co-Minkowski space on functions. Let S be a hypersurface in Minkowski space which is the graph of a map $h : B^d \to \mathbb{R}$. Then, for $(A, v) \in O_+(d, 1) \ltimes \mathbb{R}^{d,1}$, due to (16.18), the hypersurface $(A, v)S$ is the graph of the map $(A, v)h : B^d \to \mathbb{R}$ defined as

$$(A, v)h(x) := \frac{L(x)}{L(A^{-1} \cdot x)} h(A^{-1} \cdot x) + \langle x, \bar{v} \rangle_d - v_{d+1} \ . \tag{16.20}$$

Lemma 16.2.29 *Let $h : B^d \to \mathbb{R}$ be a C^2 map and $(A, v) \in O_+(d, 1) \ltimes \mathbb{R}^{d,1}$. Then*

$$\operatorname{Hess}[(A, v)h](x)(X, Y) = \frac{L(x)}{L(A^{-1} \cdot x)} \operatorname{Hess} h(A^{-1} \cdot x)(DA^{-1}(x)X, DA^{-1}(x)Y) \ .$$

Proof As $(\mathrm{Id}, v)h$ is the sum of h with an affine function, we clearly have $\mathrm{Hess}[(\mathrm{Id}, v)h](x) = \mathrm{Hess}\, h(x)$. So we need to check the result only for $(A, 0)$. As

$$\mathrm{Hess}[(A,0)h] = \mathrm{Hess}\left(\frac{L}{L \circ A^{-1}}(h \circ A^{-1})\right) ,$$

the result follows from the rules (16.11) and

$$\mathrm{Hess}(f \circ g)(x)(X, Y) = \mathrm{Hess}\, f(g(x))(\mathrm{d}\, g(x)(X), \mathrm{d}\, g(x)(Y)) + \mathrm{d}\, f(g(x))(\mathrm{Hess}\, g(x)(X, Y)) , \tag{16.21}$$

using the two following facts during the computations:

- $\frac{L}{L \circ A}$ is an affine map by (16.16), so has null Hessian;
- Differentiating two times (16.15) we obtain

$$A\binom{X}{0}_{d+1} DA(x)(Y) + A\binom{Y}{0}_{d+1} DA(x)(X) + A\binom{x}{1}_{d+1} \mathrm{Hess}\, A(x)(X, Y) = 0 ,$$

  so using (16.16) again,

$$\mathrm{d}\,\frac{L}{L \circ A} \otimes \mathrm{d}\, A + \mathrm{d}\, A \otimes \mathrm{d}\,\frac{L}{L \circ A} + \frac{L}{L \circ A}\mathrm{Hess}\, A = 0 .$$

□

Lemma 16.2.30 *Let $h : B^d \to \mathbb{R}$ be a convex map. Then for $(A, v) \in O_+(d, 1) \ltimes \mathbb{R}^{d,1}$, $(A, v)h$ is a convex map.*

Note that from (16.19), $(-\mathrm{Id}, 0)h$ is concave if h is convex.

Proof The simplest way to see this is to argue that the dual of the epigraph of h is a future convex set in Minkowski space, see Remark 16.2.22. The isometry (A, v) will send this future convex set to a future convex set (because $A \in O_+(d, 1)$), whose support function is exactly $(A, v)h$, hence convex. □

16.2.2.5 Connection in Cylindrical Coordinates

Clearly, the restriction of the vector field $\frac{\partial}{\partial x_t} = (0, \dots, 0, 1)$ of $\mathbb{R}^{d+2}$ to $co\mathcal{M}in^{d+1}$ is invariant under the action of the isometries of $co\mathcal{M}in^{d+1}$. It is also immediate to see that $\frac{\partial}{\partial x_t}$ is parallel: $\nabla^{co\mathcal{M}^{d+1}} \frac{\partial}{\partial x_t} = 0$. We will denote by T the image of $\frac{\partial}{\partial x_t}$ in co-Minkowski space. An elementary computation (see Fig. 16.6) shows that in the cylindrical coordinates $B^d \times \mathbb{R}$,

$$\mathrm{T} = L\frac{\partial}{\partial x_t} . \tag{16.22}$$

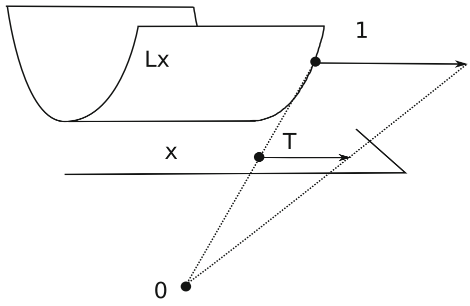


Fig. 16.6 By Thales theorem, $\mathrm{T} = L\frac{\partial}{\partial x_t}$

In particular, T is invariant under the action of $O_+(d,1) \ltimes \mathbb{R}^{d,1}$, and T is parallel: $\nabla^{*\mathbb{R}^{d,1}}\mathrm{T} = 0$. Observe that the trajectories of the flow generated by T are the vertical lines. In Minkowski space, the flow generated by T corresponds to parallel displacement of spacelike hyperplanes.

With the help of T, one can express the connection $\nabla^{*\mathbb{R}^{d,1}}$ in the cylindrical coordinates. Namely, at each point $(x,h) \in B^d \times \mathbb{R}$, we set $\mathrm{T}(x,h)$ as the vector basis for the $\mathbb{R}$-component of the tangent space. Hence a vector field X of $B^d \times \mathbb{R}$ can be written as $X = X_h + X_\mathrm{T}\mathrm{T}$, with $X_h \in T_x B^d$ and $X_\mathrm{T} \in \mathbb{R}$. If Y is another vector field of $B^d \times \mathbb{R}$, then

$$\nabla_Y^{*\mathbb{R}^{d,1}} X = \nabla_{Y_h} X_h + Y_h(X_\mathrm{T})\mathrm{T} + Y_\mathrm{T}[\mathrm{T}, X]\,. \tag{16.23}$$

This is easily checked using the definition of the connection $\nabla^{*\mathbb{R}^{d,1}}$ and the fact that T is parallel.

16.2.2.6 Volume Form

For future reference, let us mention that a volume form $\omega_{co\mathcal{M}^{d+1}}$ is also given on $co\mathcal{M}in^{d+1}$. For $v_1,\dots,v_{d+1}$ vectors of $\mathbb{R}^{d+2}$ tangent to $co\mathcal{M}in^{d+1}$, set

$$\omega_{co\mathcal{M}^{d+1}}(v_1,\dots,v_{d+1}) := \omega_{\mathbb{R}^{d+2}}(v_1,\dots,v_{d+1},\mathrm{N})$$

(recall that N is the vector field $\mathrm{N}(x) = x$ on $co\mathcal{M}in^{d+1}$). This form is invariant under orientation-preserving isometries and parallel for $\nabla^{co\mathcal{M}^{d+1}}$. It induces a parallel form $\omega_{*\mathbb{R}^{d,1}}$ on co-Minkowski space, invariant under orientation preserving isometries, and called the *volume form* of co-Minkowski space.

In the cylindrical coordinates, $\omega_{*\mathbb{R}^{d,1}}$ is defined as follows. At a point of $B^d \times \mathbb{R}$, let $v_1,\dots,v_d$ be an oriented free family of non-vertical tangent vectors. In particular, $v_1,\dots,v_d$ are tangent to a hyperbolic hyperplane, so, keeping the same notation, we can consider a family $v_1,\dots,v_d$ of oriented orthonormal vectors fields, such that $v_1,\dots,v_d,\mathrm{T}$ is positively oriented. Then $\omega_{*\mathbb{R}^{d,1}}$ is the unique $(d+1)$-form which is equal to 1 when evaluated at such a family of vectors.

Hence, if $\omega_{\mathbb{H}^d}$ is the volume form of the hyperbolic space and $\mathcal{L}_{\mathbb{R}}$ is the Lebesgue measure on the real line, then, in the cylindrical coordinates,

$$\omega = \omega_{\mathbb{H}^d} \times L^{-1}\mathcal{L}_R$$

16.2.3 Extrinsic Geometry of Graphs

Let $h : B^d \to \mathbb{R}$ be a C^2 map. Its graph S is a hypersurface in $B^d \times \mathbb{R}$, hence in co-Minkowski space if one uses the cylindrical coordinates. Note that the graph is always transverse to the vertical vector field T defined by (16.22), so the metric induced on S by the ambient degenerate metric $g_{*\mathbb{R}^{d,1}}$ of co-Minkowski space is always a hyperbolic metric. This does not give too much information. But still, some information can be obtained from the extrinsic geometry of S. To do so, we will consider the vector field T as the normal vector to S.

16.2.3.1 Second Fundamental Form and Mean Curvature

Let $h : B^d \to \mathbb{R}$ be a C^2 map and let S be its graph. Any vector field of S can be written as $X + \mathrm{d}\,h(X)L^{-1}\mathrm{T}$, where X is a vector field of B^d.

Fact 16.2.31 *For any smooth vector field X on B^d and C^2 map $h : B^d \to \mathbb{R}$,*

$$\nabla^{*\mathbb{R}^{d,1}}_{(Y+L^{-1}\,\mathrm{d}\,h(Y)\mathrm{T})}(X + L^{-1}\,\mathrm{d}\,h(X)\mathrm{T}) = \nabla^{\mathbb{H}^d}_Y X + L^{-1}\,\mathrm{d}\,h(\nabla^{\mathbb{H}^d}_Y X)\mathrm{T} + L^{-1}\mathrm{Hess}\,h(X,Y)\mathrm{T}\,. \tag{16.24}$$

Proof First let $k \in \{1, \dots, d\}$. As X does not depend on the $\frac{\partial}{\partial x_t}$ direction, and as $\mathrm{T}^k = 0$, we have $[\mathrm{T}, X]^k = \mathrm{T}^i \frac{\partial X^k}{\partial x_i} - X^i \frac{\partial \mathrm{T}^k}{\partial x_i} = 0$, and $[\mathrm{T}, X]^t = -X^i \frac{\partial \mathrm{T}^t}{\partial x_i} = -X(L)$. Also, as $L^{-1}\,\mathrm{d}\,h(X)$ does not depend on the vertical coordinate, $[\mathrm{T}, L^{-1}\,\mathrm{d}\,h(X)\mathrm{T}] = 0$. At the end of the day, if we are at a point $x \in B^d$,

$$[\mathrm{T}, X + L^{-1}\,\mathrm{d}\,h(X)\mathrm{T}] = -X(L)\frac{\partial}{\partial x_t} = -X(L)L^{-1}\mathrm{T} = L^{-2}\langle x, X\rangle\mathrm{T}\,.$$

So from (16.23),

$$\nabla^{*\mathbb{R}^{d,1}}_{(Y+L^{-1}\,\mathrm{d}\,h(Y)\mathrm{T})}(X+L^{-1}\,\mathrm{d}\,h(X)\mathrm{T}) = \nabla_Y X + (Y(L^{-1}\,\mathrm{d}\,h(X)) + L^{-3}\,\mathrm{d}\,h(Y)\langle x, X\rangle_d)\mathrm{T}\,.$$

We have $Y(L^{-1}\,\mathrm{d}\,h(X)) = L^{-1}(Y(X(h)) + Y(L^{-1})\,\mathrm{d}\,h(X)$, and from (16.10), $L^{-1}(Y(X(h)) = L^{-1}\,\mathrm{Hess}^{\mathbb{H}^d}\,h(X,Y) + L^{-1}\,\mathrm{d}\,h(\nabla_Y X)$. Also, if we are at the point x, $Y(L^{-1}) = \langle x, Y\rangle_d L^{-3}$:

$$\nabla^{*\mathbb{R}^{d,1}}_{(Y+L^{-1}\,\mathrm{d}\,h(Y)\mathrm{T})}(X + L^{-1}\,\mathrm{d}\,h(X)\mathrm{T}) = \nabla_Y X + L^{-1}\,\mathrm{d}\,h(\nabla_Y X)\mathrm{T} + L^{-1}\mathcal{X}\mathrm{T}$$

with $\mathcal{X} = \mathrm{Hess}^{\mathbb{H}^d} h(X, Y) + L^{-2}\langle x, Y\rangle_d \, \mathrm{d}h(X) + L^{-2}\langle x, X\rangle_d \, \mathrm{d}h(Y)$, and by (16.8), $\mathcal{X} = \mathrm{Hess}\, h(X, Y)$. □

Given two vector fields tangent to S, the graph of h, then their co-Minkowski connection decomposes as a part tangent to S, and a part colinear to T, where T may be thought of as a unit normal vector field to S. Mimicking the classical theory of surfaces, we define the *second fundamental form* II_h of S as the collinearity factor. More precisely, Eq. (16.24) says that for $x \in B^d$ and $X, Y \in T_x B^d$,

$$\mathrm{II}_h(x)(X, Y) = L^{-1}(x)\mathrm{Hess}\; h(x)(X, Y)\;. \tag{16.25}$$

Remark 16.2.32 From Lemma 16.2.13, the second fundamental form is a symmetric Codazzi tensor on $\mathbb{H}^d$, and any symmetric Codazzi tensor on hyperbolic space is the second fundamental form of a unique hypersurface in co-Minkowski space. This is a kind of "fundamental theorem for hypersurfaces" in co-Minkowski space, with the condition about the first fundamental form reduced to the hypothesis that the metric is hyperbolic. Note that here there is no Gauss condition, i.e. for $d = 2$ there is no relation between the curvature of the induced metric and the determinant of the second fundamental form. This result was probably first proved in [60], see Proposition 9.1.10 in this reference.

The *shape operator* $\mathbf{shape}(h)$ of S is the symmetric linear mapping associated to the second fundamental form by the hyperbolic metric: $\mathrm{II}_h(X, Y) = g_{\mathbb{H}^d}(\mathbf{shape}(h)(X), Y)$. From (16.9), if $\mathrm{grad}^{\mathbb{H}^d}$ is the gradient for $g_{\mathbb{H}^d}$, we have

$$\mathbf{shape}(h)(X) = \nabla_X^{\mathbb{H}^d} \mathrm{grad}^{\mathbb{H}^d}(L^{-1}h) - (L^{-1}h)X\;.$$

The *mean curvature* $\mathbf{Mean}(h)$ of the graph of h is the trace for the hyperbolic metric of the shape operator times $1/d$. From the definition or Fact 16.2.8, it can be written in different ways: with the help of the the Euclidean Laplacian Δ

$$\mathbf{Mean}(h)(x) = \frac{1}{d}\mathrm{Tr}_{g_{\mathbb{H}^d}}\left(L^{-1}\mathrm{Hess}\; h\right)(x) = \frac{1}{d}L(x)(\Delta h(x) - \mathrm{Hess}\; h(x)(x, x))\;, \tag{16.26}$$

or with the help of the hyperbolic Laplacian $\Delta^{\mathbb{H}^d}$

$$\mathbf{Mean}(h)(x) = \frac{1}{d}\Delta^{\mathbb{H}^d}(L^{-1}h)(x) - (L^{-1}h)(x)\;. \tag{16.27}$$

Remark 16.2.33 Let us suppose that $h : B^d \to \mathbb{R}$ is C^2 and convex. Using a basis of eigenvectors, it follows from (16.26) that $\mathbf{Mean}(h)$ is non-negative, and that if $\mathbf{Mean}(h) = 0$ then h is affine.

Proposition 16.2.34 ([49]) *If the graph of a C^2 convex function $h : B^d \to \mathbb{R}$ has its mean curvature bounded from above, then h has a continuous extension to $\bar{B}^d$.*

Proof Suppose that there is C such that for any $x \in B^d$, $\mathbf{Mean}(h)(x) < C$. Let $\theta \in \partial B^d$, and let h_θ be the restriction of h to the segment parameterized by $r \in [0, 1[$ from the origin to θ. Let us also denote $l(r) = \sqrt{1-r^2}$. By (16.26),

$$h''_\theta(r) < Cl(r)^{-3} .$$

For $1/2 < r < 1$, we write

$$h'_\theta(r) \leq h'_\theta(1/2) + C \int_{1/2}^{r} l^{-3}$$

and as, for $1/2 < t < 1$, $(1-t^2)^{-1} < (1-t)^{-1}$, we have

$$h'_\theta(r) < h'_\theta(1/2) + 2C(1-r)^{-1/2} . \tag{16.28}$$

Also, as h is convex, h_θ is convex, hence for $1/2 < r < 1$,

$$h'_\theta(1/2) \leq h'_\theta(r) . \tag{16.29}$$

Let us define

$$g(\theta) = \int_{1/2}^{1} h'_\theta - h_\theta(1/2) .$$

As $\int_{1/2}^{1}(1-r)^{-1/2}\mathrm{d}r$ is finite, by (16.28) and (16.29), $g(\theta)$ is well defined. Also, together with (16.28), (16.29) and the Dominated convergence theorem, g is continuous. □

16.2.3.2 Mean Surfaces

Definition 16.2.35 A hypersurface S of co-Minkowski space is called *mean* if it is the graph of a C^2 function $h : B^d \to \mathbb{R}$ with $\mathbf{Mean}(h) = 0$.

Abusing terminology, the function h itself may be also called mean.

Note that when $d = 2$, the mean surface is not critical for the area functional, as all the graphs of functions $B^2 \to \mathbb{R}$ in co-Minkowski space have the same area form (because they are all isometric to the hyperbolic plane).

Due to (16.26), h is mean if and only if for any $x \in B^d$, $\Delta h(x) - \mathrm{Hess}\, h(x)(x, x) = 0$. This is an elliptic equation with only second-order terms, that allows to apply strong results of PDE theory. For this, we have to consider boundary conditions.

Definition 16.2.36 Let $b : \partial B^d \to \mathbb{R}$ be a continuous map. A continuous function $h : B^d \to \mathbb{R}$ is called a *b-map* if it extends continuously as b on ∂B^d.

Proposition 16.2.37 *For any continuous function $b : \partial B^d \to \mathbb{R}$, there is a unique C^∞ smooth mean b-map, denoted by h_b^{mean}.*

Proof Uniqueness is classical from the ellipticity of L^{-1}**Mean** [33, Theorem 3.3]. Existence follows from the fact that the elliptic equation $\Delta f(x) - \text{Hess}\ f(x)(x,x) = 0$ has only second-order terms and that the domain is a ball, see [33, Corollary 6.24']. Dividing the equation by L^2, we obtain a strictly elliptic equation, and regularity theorems apply, e.g. [33, Corollary 8.11]. □

Lemma 16.2.38 *If $b_n : \partial B^d \to \mathbb{R}$ are continuous functions uniformly converging to $b : \partial B^d \to \mathbb{R}$, then $h_{b_n}^{\text{mean}}$ is converging to h_b^{mean}.*

Proof Let b_n be such that the supremum of $|b_n - b|$ is arbitrarily small. Then $\mathbf{Mean}(h_{b_n}^{\text{mean}} - h_b^{\text{mean}}) = 0$, with boundary data $b_n - b$. By the maximum principle [33, Theorem 3.1], $h_{b_n}^{\text{mean}} - h_b^{\text{mean}}$ is arbitrarily small. The same conclusion holds for $h_b^{\text{mean}} - h_{b_n}^{\text{mean}}$. □

Remark 16.2.39 For a continuous map $b : \partial B \to \mathbb{R}$, it is possible to associate to h_b^{mean} a (non-regular and non convex) dual hypersurface in Minkowski space, see Remark 16.2.24. For $d = 2$, at points of regularity, this surface has zero mean curvature. We refer to [30] for more details.

16.2.3.3 Convex Hull

Let $b : \partial B^d \to \mathbb{R}$ be a continuous map. Let

$$\mathcal{A}_b = \{a | a : \mathbb{R}^d \to \mathbb{R} \text{ is an affine function and } a|_{\partial B^d} \leq b\}$$

and for $x \in B^d$, let us define

$$h_b^-(x) := \sup\{a(x) | a \in \mathcal{A}_b\}\ , \tag{16.30}$$

and

$$h_b^+(x) := -h_{-b}^-(x)\ . \tag{16.31}$$

Proposition 16.2.40 *For any $x \in B^d$, $h_b^-(x)$ defines a convex b-map $h_b^- : B^d \to \mathbb{R}$. Moreover, if $h : B^d \to \mathbb{R}$ is a convex b-map, then $h_b^- \geq h$.*

For any $x \in B^d$, $h_b^+(x)$ defines a concave b-map $h_b^+ : B^d \to \mathbb{R}$. Moreover, if $h : B^d \to \mathbb{R}$ is a concave b-map, then $h_b^+ \leq h$.

In general, we have $h_b^+ \geq h_b^-$. If $h_b^+(x) = h_b^-(x)$ for some $x \in B^d$, then b is the restriction to ∂B^d of an affine map of $\mathbb{R}^d$.

Proof The properties of h_b^- are proved in the proof of Theorem 1.5.2 in [38]. The properties of h_b^+ then follows immediately from (16.31). The last property is then obvious, since $h_b^+ - h_b^-$ is then a non-negative concave map vanishing somewhere, therefore, vanishing everywhere, and since affine maps are the only ones being at the same time convex and concave. □

Let $\Lambda(b)$ be the graph of $b : \partial B^d \to \mathbb{R}$ in $\partial B^d \times \mathbb{R}$, and let $\mathrm{CH}(b)$ be the affine convex hull of $\Lambda(b)$ in $\mathbb{R}^{d+1}$, that is, the smallest convex set of $\mathbb{R}^{d+1}$ containing $\Lambda(b)$. Note that as $\bar{B}^d \times \mathbb{R}$ is a convex set containing $\Lambda(b)$, then $\mathrm{CH}(b) \subset \bar{B}^d \times \mathbb{R}$.

Lemma 16.2.41 *The boundary of* $\mathrm{CH}(b)$ *is the union of the graphs of* h_b^+ *and* h_b^-.

Proof This follows from the definitions of h_b^+ and h_b^-, because $\mathrm{CH}(b)$ is the intersection of all the half-spaces containing $\Lambda(b)$. □

The set $\mathrm{CH}(b)$ satisfies the local geodesic property: for any $x \in \mathrm{CH}(b) \setminus \Lambda(b)$, x lies in an open segment contained in $\mathrm{CH}(b) \setminus \Lambda(b)$ [62, Theorem 4.19].

Lemma 16.2.42 *The mean surface given by the boundary condition* $b : \partial B^d \to \mathbb{R}$ *is contained in the convex hull* $\mathrm{CH}(b)$*:*

$$h_b^- \le h_b^{\mathrm{mean}} \le h_b^+ .$$

Proof Let $a \in \mathcal{A}_b$. By the maximum principle [33, Theorem 3.1], $a - h_b^{\mathrm{mean}}$ attains its maximal value on ∂B. But on ∂B^d, $a \ge b$, so on B^d, $a - h_b^{\mathrm{mean}} \le a|_{\partial B^d} - b = b - b = 0$, i.e. $a \le h_b^{\mathrm{mean}}$. Then by definition of h_b^-, $h_b^- \le h_b^{\mathrm{mean}}$. Similarly, one proves that $h_{-b}^- \le h_{-b}^{\mathrm{mean}} = -h_b^{\mathrm{mean}}$ i.e. $h_b^+ = -h_{-b}^- \ge h_b^{\mathrm{mean}}$. □

Lemma 16.2.43 *If* $(b_n)_{n\in\mathbb{N}}$ *is a sequence of continuous functions from* ∂B^d *into* $\mathbb{R}$ *converging uniformly to* $b : \partial B^d \to \mathbb{R}$*, then* $(h_{b_n}^-)_{n\in\mathbb{N}}$ *(resp.* $(h_{b_n}^+)_{n\in\mathbb{N}}$*) converges to* h_b^- *(resp.* h_b^+*).*

Proof Let $\epsilon > 0$ and $x \in B^d$. Then there exists an affine function a such that $h_b^-(x) \ge a(x)$, $a(x) + \epsilon \ge h_b^-(x)$ and $a|_{\partial B^d} \le b$. In particular, for n large enough, $a|_{\partial B^d} - \epsilon \le b_n$. As $a|_{\partial B^d} - \epsilon$ is an affine function, then $h_{b_n}^-(x) \ge a(x) - \epsilon$. As a was chosen such that $a(x) + \epsilon \ge h_b^-(x)$, then $h_{b_n}^-(x) + 2\epsilon \ge h_b^-(x)$. A similar conclusion holds, exchanging the roles of b and b_n. □

Remark 16.2.44 The dual in Minkowski space of the epigraph of a convex b-map is a convex set. Its *domain of dependence*, or *Cauchy domain*, denoted by Ω_b^-, is the interior of the intersection of the future side of all the lightlike hyperplanes containing it. This intersection is nothing but the dual of the epigraph of h_b^-. The domain of dependence Ω_b^- is future complete. Considering h_b^+ instead of h_b^-, and concave figures instead of convex ones, we obtain the domain of dependence Ω_b^+. See Fig. 16.8 and [3, 12] for more details.

Remark 16.2.45 The function h_b^{mean} is the solution of the Dirichlet problem for an elliptic linear equation. The convex function h_b^- is the solution of the Dirichlet problem for the Monge–Ampère equation, see [38].

16.2.3.4 The Mean Curvature Measure

For a C^2 function $h : B^d \to \mathbb{R}$, we have defined in Sect. 16.2.3.1 the mean curvature function, which by Remark 16.2.33 is non-negative if h is convex. For a convex C^2 function $h : B^d \to \mathbb{R}$, let us define the *mean curvature measure*

$$\mathbf{MM}(h) = d \times \mathbf{Mean}(h)\omega_{\mathbb{H}^d} ,$$

where $\omega_{\mathbb{H}^d}$ is the volume form given by the hyperbolic metric on B^d. By (16.27), for any $\varphi \in C^0_0(B^d)$ (here the subscript 0 means "with compact support"),

$$\mathbf{MM}(h)(\varphi) = \int_{B^d} \varphi \left(\frac{1}{d} \Delta^{\mathbb{H}^d}(L^{-1}h) - (L^{-1}h) \right) \mathrm{d}\,\omega_{\mathbb{H}^d} .$$

If moreover $\varphi \in C^\infty_0(B^d)$, by integration by part:

$$\mathbf{MM}(h)(\varphi) = \int_{B^d} L^{-1}h \left(\frac{1}{d} \Delta^{\mathbb{H}^d}\varphi - \varphi \right) \mathrm{d}\,\omega_{\mathbb{H}^d} . \tag{16.32}$$

Using (16.6) and (16.26), alternatives formulas are, for $\varphi \in C^0_0(B^d)$

$$\mathbf{MM}(h)(\varphi) = \int_{B^d} (\Delta h(x) - \mathrm{Hess}\, h(x)(x, x))\, L^{-d}(x)\varphi(x)\,\mathrm{d}\,x ,$$

and for $\varphi \in C^\infty_0(B^d)$,

$$\mathbf{MM}(h)(\varphi) = \int_{B^d} (\Delta\varphi(x) - \mathrm{Hess}\,\varphi(x)(x, x))\, h(x)L^{-d}(x)\,\mathrm{d}\,x . \tag{16.33}$$

For any convex function $h : B^d \to \mathbb{R}$, let us define $\mathbf{MM}(h)$ as the linear form on $C^\infty_0(B^d)$ defined by (16.33). On any compact ball K contained in B^d, by standard convolution, one can find a sequence $(h_i)_{i\in\mathbb{N}}$ of C^∞ convex functions uniformly approximating h. For any C^∞ function φ whose support is included in K, we clearly have $\mathbf{MM}(h_j)(\varphi) \to \mathbf{MM}(h)(\varphi)$. As $\mathbf{MM}(h_j)$ is a measure, it is also a distribution, and the preceding limit says that $\mathbf{MM}(h)$ is also a distribution on K [39, Theorem 2.1.8]. Actually, as the $\mathbf{MM}(h_j)$ are measures, then $\mathbf{MM}(h)$ is a measure on K [39, Theorem 2.1.9, Theorem 2.1.7]. Changing K and using the localization property of distribution [39, Theorem 2.2.4], it follows that $\mathbf{MM}(h)$ is a measure on B^d. More precisely, $\mathbf{MM}(h)$ is a Radon measure on B^d.

The following result is given by [39, Theorem 2.1.9, Theorem 2.1.7].

Lemma 16.2.46 *Let $(h_n)_{n\in\mathbb{N}}$ be a sequence of convex functions from B^d into $\mathbb{R}$ converging to a convex function $h : B^d \to \mathbb{R}$. Then the sequence of measures $(\mathbf{MM}(h_n))_{n\in\mathbb{N}}$ weakly converges to $\mathbf{MM}(h)$.*

Recall the action of isometries on functions defined by (16.20). Recall also from Lemma 16.2.30 that if h is convex, then $(A, v)h$ is convex for $(A, v) \in O_+(d, 1) \ltimes \mathbb{R}^{d,1}$.

Lemma 16.2.47 *Let $\varphi \in C^0_0(B^d)$ and $(A, v) \in O_+(d, 1) \ltimes \mathbb{R}^{d,1}$. Then:*

$$\mathbf{MM}((A, v)h)(\varphi) = \mathbf{MM}(h)(\varphi \circ A) .$$

Proof We will prove the result for a C^2 function h; the general result follows by approximation. In the C^2 case, the result follows because by definition

$$\mathbf{MM}(h)(\varphi \circ A) = \int_{B^d} (\varphi \circ A)(x)(\mathrm{Tr}_{g_{\mathbb{H}^d}} L^{-1}\mathrm{Hess}\ h)(x)\mathrm{d}\omega_{\mathbb{H}^d}(x)$$

so by a change of variable, as A is a hyperbolic isometry,

$$\mathbf{MM}(h)(\varphi \circ A) = \int_{B^d} \varphi(x)(\mathrm{Tr}_{g_{\mathbb{H}^d}} L^{-1}\mathrm{Hess}\ h)(A^{-1} \cdot x)\mathrm{d}\omega_{\mathbb{H}^d}(x) ,$$

and by Lemma 16.2.29,

$$(\mathrm{Tr}_{g_{\mathbb{H}^d}} L^{-1}\mathrm{Hess}\ h)(A^{-1} \cdot x) = (\mathrm{Tr}_{g_{\mathbb{H}^d}} L^{-1}\mathrm{Hess}\ [(A, v)h])(x) .$$

□

16.2.3.5 The Fundamental Example of a Wedge

Let us consider an elementary example to give a geometric insight on the mean curvature measure introduced in the previous section. This example will make clear that, for well-chosen convex functions, this measure is a kind of "pleating measure", similar to the notion developed by Thurston for isometric pleated embeddings of hyperbolic surfaces in the 3-dimensional hyperbolic space, see Sects. 16.3.3.4 and 16.3.4.

Let l be the intersection of B^d with an affine hyperplane of $\mathbb{R}^d$, which separates B^d into two connected components l^- and l^+, where l^- is the component containing the origin 0 of the coordinates of $\mathbb{R}^d$. Let p_l be the (Euclidean) orthogonal projection of 0 onto l, and let $n_l = p_l/\|p_l\|$. If l is a vector hyperplane, then l^- is chosen arbitrarily, and n_l is the (Euclidean) unit normal vector pointing to l^+.

Definition 16.2.48 The *canonical map* $h_l : B^d \to \mathbb{R}$ associated to l is defined as $h_l(x) = \frac{1}{L(p_l)}\langle x - p_l, n_l\rangle$.

Observe that h_l is an affine map vanishing on l. Let $\mathbf{1}_A$ be the indicator function of a set A.

Definition 16.2.49 A *wedge* on a hyperplane l is a continuous map $h : B^d \to \mathbb{R}$ of the form $h = h_- + (h_+ - h_-)\mathbf{1}_{l^+}$ where h_-, h_+ are two affine maps.

The *angle* of a wedge (in the co-Minkowski sense) is the unique real number α such that, with the above notation,

$$h_+ - h_- = \alpha h_l \ . \quad (16.34)$$

The wedge is therefore a piecewise affine map, admitting l as a locus of non-differentiability (if the angle is nonzero).

Fact 16.2.50 *A wedge is convex and different from an affine map if and only if its angle is positive.*

Proof By definition, h_l is positive on $l^+ \setminus l$. And h is strictly convex if and only if on $l^+ \setminus l$, $h_+ = h_- + \alpha h_l > h_-$, which is true if and only if $\alpha > 0$. □

Remark 16.2.51 The hyperplane l in B^d defines a timelike vector hyperplane in Minkowski space, namely, if B^d is identified with the Klein ball model of the hyperbolic space in $\mathbb{R}^{d,1}$, the vector hyperplane passing through $l \times \{1\}$. Let v_l be its unit spacelike normal vector pointing to the side containing l^+. Then it is easy to see that

$$v_l = \frac{1}{p_l}\begin{pmatrix} n_l \\ \|p_l\| \end{pmatrix} , \quad (16.35)$$

and so the canonical map h_l is the restriction to $B^d \times \{1\}$ of the linear map $(x, x_{d+1}) \mapsto \langle \binom{x}{x_{d+1}}, v_l \rangle_{d,1}$. If l is a vector hyperplane, then $v_l = \binom{n_l}{0}$. Moreover, if P_+ and P_- are the duals of the graphs of h_+ and h_-, then $P_+ - P_-$ is colinear to v, that expresses the definition (16.34) (compare also with Fact 16.2.21). The absolute value of α is the Minkowski length of the spacelike segment $P_+ - P_-$. See Fig. 16.7.

Fact 16.2.52 *Let* $A \in O_+(d, 1)$. *Then* $h_{A \cdot l} = \frac{L}{L \circ A^{-1}} h_l \circ A^{-1}$.

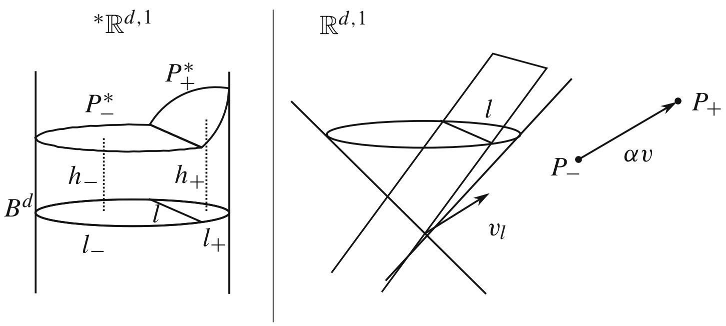


Fig. 16.7 The spacelike segment in Minkowski space dual to a convex wedge in co-Minkowski space

Proof With the notation of Remark 16.2.51, we clearly have $v_{A\cdot l} = A(v_l)$, hence $h_{A\cdot l}(x) = \langle \binom{x}{1}, A(v_l)\rangle_{d,1} = \langle A^{-1}\binom{x}{1}, v_l\rangle_{d,1}$, and by (16.17), $A^{-1}(\binom{x}{1}) = \frac{L(x)}{L(A^{-1}\cdot x)}\binom{A^{-1}\cdot x}{1}$. □

Fact 16.2.53 *The image of the graph of a wedge by an orientation-preserving co-Minkowski isometry is the graph of a wedge of the same angle.*

Proof The result is obvious from Remark 16.2.51, as a co-Minkowski isometry acts as a Minkowski isometry on the dual objects, and hence sends a spacelike segment to a spacelike segment of the same length. □

The choice of the normal n_l gives an orientation on the vector hyperplane l, which is also isometric to $\mathbb{H}^{d-1}$. We denote by $\omega_l^{\mathbb{H}}$ its volume form for the hyperbolic metric.

Lemma 16.2.54 *Let h be a convex wedge of angle α on a hyperplane l. Then the following identity holds:*

$$\mathbf{MM}(h) = \alpha\omega_l^{\mathbb{H}} .$$

The simplest illustration of the lemma is for $d = 1$, $l = \{0\}$ and $h(x) = |x| = -x + 2x\mathbf{1}_{\mathbb{R}_+}$. Then the angle is equal to 2, and h'' in the sense of distributions is equal to $2\delta(0)$.

Proof From (16.34), $h = h_- + \alpha h_l \mathbf{1}_{l^+}$, so as h_- is affine, in the sense of distributions, $\partial_{ij}h = \alpha\partial_{ij}(h_l\mathbf{1}_{l^+})$. By successive integrations by part, for $\phi \in C_0^\infty$, using that $h_l = 0$ on l and that h_l is affine, we obtain, in the sense of distributions, $\partial_{ij}h = \alpha\partial_i h_l(n_l)_j \mathrm{d}S$, where $\mathrm{d}S$ is the (Euclidean) area form on l (n_l is an inward normal vector for l^+).

Hence by (16.33) and (16.34), the measure $\mathbf{MM}(h)$ is given for $x \in l$ by

$$\alpha(\langle n_l, \operatorname{grad} h_l\rangle + \langle n_l, x\rangle\langle \operatorname{grad} h_l, x\rangle)L^{-d}(x)\,\mathrm{d}\,S(x) . \tag{16.36}$$

Let us first consider that l is the intersection of B^d with a vector hyperplane. Then, $\langle n_l, x\rangle = 0$, $\langle n_l, \operatorname{grad} h_l\rangle = 1$, and from (16.6), $L^{-d}(x)\,\mathrm{d}\,S = \mathrm{d}\,\omega_l^{\mathbb{H}}$. At the end of the day, (16.36) becomes $\alpha\,\mathrm{d}\,\omega_l^{\mathbb{H}}$, that is the desired result when l is defined by a vector hyperplane. The general case follows by performing an orientation-preserving isometry sending l to a vector hyperplane, and using Lemma 16.2.47 and Fact 16.2.53. □

Remark 16.2.55 Given a hyperplane l of B^d weighted by a positive number α, it is almost clear how to construct a convex wedge in co-Minkowski space with angle α. This construction can be easily extended to non-intersecting weighted hyperplanes (see Sect. 16.3.3.4), or to a "polyhedral case", i.e., weighted hyperplanes are allowed to meet to form a convex cellulation of $\mathbb{H}^d$, together with a natural compatibility conditions at the weights, see [30, 4.4] and [27] for the $d = 2$ case. This is a polyhedral version of the *Christoffel problem*, whose aim is to find a convex hypersurface in Minkowski space prescribing the dual Mean curvature measure. In this setting the dual Mean curvature measure is called the *area measure of order one*. The Christoffel problem in Minkowski space is the subject of [30].

The polyhedral construction is also a version of the classical *Maxwell-Cremona correspondence* or *Maxwell lift*, see [40].

16.3 Action of Cocompact Hyperbolic Isometry Groups

16.3.1 Translation Parts as Cocycles

Let Γ be a subgroup of $O_+(d, 1)$ such that $\mathcal{H}^d/\Gamma$ is a compact oriented hyperbolic manifold. A *cocycle* $\tau \in Z^1(\Gamma, \mathbb{R}^{d,1})$ is a map $\tau : \Gamma \to \mathbb{R}^{d,1}$ satisfying, for $A, B \in \Gamma$,

$$\tau(AB) = \tau(A) + A(\tau(B)) .$$

Let us denote

$$\Gamma_\tau = \{(A, \tau(A))|A \in \Gamma\} .$$

From (16.2), Γ_τ is a subgroup of the isometry group of Minkowski space. In turn, it defines a group of isometries of co-Minkowski space, that we will also denote by Γ_τ.

In the cylindrical coordinates $B^d \times \mathbb{R}$ of co-Minkowski space, Γ acts freely and properly discontinuously on $B^d \times \{0\}$. As co-Minkowski space is the product manifold $B^d \times \mathbb{R}$, due to (16.18), the following result is trivial, but worth to notice.

Lemma 16.3.1 *The action of* Γ_τ *on* $^*\mathbb{R}^{d,1}$ *is free and properly discontinuous.*

A *coboundary* is a particular cocycle of the form

$$\tau(A) = Av - v$$

for a given $v \in \mathbb{R}^{d,1}$. The group $H^1(\Gamma, \mathbb{R}^{d,1})$ is the quotient of the space of cocycles by the space of coboundaries: two cocycles are in relation if and only if they differ by a coboundary.

In the following, we make the implicit assumption that we are looking at Γ such that $H^1(\Gamma, \mathbb{R}^{d,1})$ is not reduced to zero.

Let us give a criterion of non-triviality. Let us suppose that the compact hyperbolic manifold $\mathcal{H}^d/\Gamma$ contains n disjoints embedded totally geodesic hypersurfaces $H_1, \dots, H_n$. Also, let us set some positive weights ω_i to each H_i. This is actually a *simplicial measured geodesic lamination* λ on $\mathcal{H}^d/\Gamma$.

A lift to B^d of an H_i is a hyperplane l. Recall from (16.35) that a vector v_l of $\mathbb{R}^{d,1}$ is assigned to any such l. Let us denote by $\tilde{L}$ the set of the lifts the H_i. Let us fix an arbitrary base point $\tilde{x} \in B^d \setminus \tilde{L}$. Then define, for $A \in \Gamma$, and for any path $c : [0, 1] \to B^d$, transverse to $\tilde{L}$ and joining $\tilde{x}$ to $A \cdot \tilde{x}$:

$$\tau_\lambda(A) = \sum_{j \in c([0,1]) \cap \tilde{L}} \omega_j v_j \ , \tag{16.37}$$

where v_j is plus or minus the vector v associated to the jth element of $c([0, 1]) \cap \tilde{L}$. The sign is chosen so that v_j has the same direction than the path c at the corresponding point. It follows that the definition of τ_λ is independent from the choice of the path c among paths transverse to $\tilde{L}$ joining the same endpoints.

Fact 16.3.2 *With the above notation $\tau_\lambda \in Z^1(\Gamma, \mathbb{R}^{d,1})$.*

Proof Let $A, B \in \Gamma$. Let $c_A, c_B : [0, 1] \to B^d$ be paths transverse to $\tilde{L}$, and joining $\tilde{x}$ to $A \cdot \tilde{x}$ and $B \cdot \tilde{x}$ respectively. Let c_{AB} be the concatenation of c_A with $A \cdot c_B$. This is a path joining $\tilde{x}$ to $(AB) \cdot \tilde{x}$ and transverse to L, so

$$\tau_\lambda(AB) = \sum_{j \in (c_{AB}([0,1]) \cap \tilde{L})} \omega_j v_j = \sum_{j \in (c_A([0,1]) \cap \tilde{L})} \omega_j v_j + \sum_{j \in (A \cdot c_B([0,1]) \cap \tilde{L})} \omega_j v_j \ .$$

By definition of v_l, we clearly have $v_{A \cdot l} = A(v_l)$, and A acts linearly on $\mathbb{R}^{d,1}$, so

$$\tau_\lambda(AB) = \sum_{j \in (c_A([0,1]) \cap \tilde{L})} \omega_j v_j + A \left(\sum_{j \in (c_B([0,1]) \cap \tilde{L})} \omega_j v_j \right) = \tau(A) + A(\tau(B)) \ .$$

□

Fact 16.3.3 *Let τ'_λ be the cocycle defined by* (16.37)*, but choosing another base-point $\tilde{x}'$. Then $\tau'_\lambda - \tau_\lambda$ is a coboundary.*

Proof For any $A \in \Gamma$, let $c : [0, 1] \to B^d$ be a path transverse to $\tilde{L}$ joining $\tilde{x}$ to $A \cdot \tilde{x}$, and let $c' : [0, 1] \to B^d$ be a path transverse to $\tilde{L}$ joining $\tilde{x}'$ to $A \cdot \tilde{x}'$. Let $\bar{c} : [0, 1] \to B^d$ be any path transverse to $\tilde{L}$ joining $\tilde{x}$ to $\tilde{x}'$. Then the concatenation c^* of $\bar{c}$ with c' and $-A \cdot \bar{c}$ is a transverse path joining $\tilde{x}$ to $A \cdot \tilde{x}$, so

$$\tau_\lambda(A) = \sum_{j \in (c^*([0,1]) \cap \tilde{L})} \omega_j v_j = \sum_{j \in (\bar{c}([0,1]) \cap \tilde{L})} \omega_j v_j + \sum_{j \in (c'([0,1]) \cap \tilde{L})} \omega_j v_j - \sum_{j \in (A \cdot \bar{c}([0,1]) \cap \tilde{L})} \omega_j v_j$$

$$= \sum_{j \in (\bar{c}([0,1]) \cap \tilde{L})} \omega_j v_j + \tau'_\lambda(A) - A \left(\sum_{j \in (\bar{c}([0,1]) \cap \tilde{L})} \omega_j v_j \right)$$

so if v is the vector $-\sum_{j \in (\bar{c}([0,1]) \cap \tilde{L})} \omega_j v_j$ we have $\tau_\gamma(A) - \tau'_\gamma(A) = Av - v$. □

So for each choice of positive weights, we have constructed an element of $H^1(\Gamma, \mathbb{R}^{d,1})$. Clearly, a linearly independent change in the weights will produce a different element in $H^1(\Gamma, \mathbb{R}^{d,1})$, hence we have a simple geometric proof of the following classical result (see Sect. 16.1).

Theorem 16.3.4 *If $\mathbb{H}^d/\Gamma$ contains n disjointly embedded totally geodesic hypersurfaces, then the dimension of $H^1(\Gamma, \mathbb{R}^{d,1})$ is $\geq n$.*

16.3.2 Equivariant Maps

Let $\tau \in Z^1(\Gamma, \mathbb{R}^{d,1})$. We will give more details on the action of Γ_τ by looking at particular functions. The analysis is simplified using the cylindrical coordinates of co-Minkowski space. We say that a continuous map $h : B^d \to \mathbb{R}$ is Γ-*invariant* if its graph is invariant for the action of Γ, i.e. for all $A \in \Gamma$, $(A, 0)h = h$ (recall (16.20)):

$$\forall x \in B^d,\ (L^{-1}h)(A \cdot x) = (L^{-1}h)(x)\ ,$$

in other terms, h is Γ-invariant if and only if $L^{-1}h$ is invariant for the action of Γ. In particular, if h is Γ-invariant, as the action of Γ is cocompact on B^d, $L^{-1}h$ is bounded. Note that the function L is obviously Γ-invariant (see Remark 16.2.27 for a geometric viewpoint).

Fact 16.3.5 *Let h be a Γ-invariant function. Then h extends continuously as the constant zero function on ∂B^d.*

Proof There exist two constants c_1, c_2 such that $c_1 \leq L^{-1}h \leq c_2$, so $c_1 L \leq h \leq c_2 L$, and the result follows. □

Definition 16.3.6 A continuous map $h : B^d \to \mathbb{R}$ is Γ_τ-*equivariant* if its graph is invariant for the action of Γ_τ, i.e. for all $A \in \Gamma$, $(A, \tau(A))h = h$, using the notation introduced in (16.20).

The vector space structure of $Z^1(\Gamma, \mathbb{R}^{d,1})$ fits well with the vector space structure of maps, as the following lemma shows. Its proof is trivial from Definition 16.3.6.

Fact 16.3.7 *Let $\tau_1, \tau_2 \in Z^1(\Gamma, \mathbb{R}^{d,1})$ and let h_1 and h_2 be Γ_{τ_1} and Γ_{τ_2}-equivariant maps respectively, and $\alpha \in \mathbb{R}$. Then $h_1 + \alpha h_2$ is $\Gamma_{\tau_1+\alpha\tau_2}$-equivariant. In particular, the difference between two Γ_τ-equivariant map is a Γ-invariant map.*

Fact 16.3.8 *Let h be a Γ_τ and $\Gamma_{\tau'}$-equivariant map. Then $\tau = \tau'$.*

Proof Because $0 = h - h$ is $\Gamma_{\tau-\tau'}$-equivariant, it suffices to check that if 0 is Γ_τ-equivariant, then $\tau = 0$. By definition, this implies that for any $A \in \Gamma$ and any $x \in B^d$, $\langle A \cdot x, \overline{\tau(A)}\rangle_d) = \tau(A)_{d+1}$. If $\overline{\tau(A)}_d \neq 0$, then for x this equation can be satisfied only by a dimension one affine space, which contradicts our assumption that it is satisfied by any $x \in B^d$, $d > 1$.

□

The following fact is clear from the definition of Γ_τ-equivariant map and Lemma 16.2.29.

Fact 16.3.9 *Let $h : B^d \to \mathbb{R}$ be a C^2 Γ_τ-equivariant function. Then $L^{-1}\,\mathrm{Hess}\,h$ is Γ-invariant:*

$$(L^{-1}\,\mathrm{Hess}\,h)(x)(X, Y) = (L^{-1}\,\mathrm{Hess}\,h)(A\cdot x)\,(DA(x)(X), DA(x)(Y))\ .$$

Remark 16.3.10 Fact 16.3.9 says that the second fundamental form of the hypersurface which is the graph of h (see Sect. 16.2.3.1) defines a symmetric $(0, 2)$-tensor on $\mathbb{H}^d/\Gamma$. Moreover this tensor is a symmetric Codazzi tensor, see Remark 16.2.32.

It can be useful to note the following converse to Fact 16.3.9.

Lemma 16.3.11 *Let $h : B^d \to \mathbb{R}$ be a C^2 map such that $L^{-1}\,\mathrm{Hess}\,h$ is Γ-invariant. Then there exists a unique $\tau \in Z^1(\Gamma, \mathbb{R}^{d,1})$ such that h is Γ_τ-equivariant.*

Proof Let $A \in \Gamma$. As A acts as an affine map on $B^d \times \mathbb{R}$, by the rule of the Hessian of a composition (16.21) and the invariance of the Hessian, we obtain

$$\mathrm{Hess}(h\circ A)(x)(X, Y) = \mathrm{Hess}\ h(A\cdot x)(DA(x)(X), DA(x)(Y)) = \mathrm{Hess}h(x)(X, Y)\,,$$

hence h and $h\circ A$ differ by an affine map, which in turn gives a vector $\tau(A^{-1}) \in \mathbb{R}^{d,1}$:

$$h(x) - h(A\cdot x) = \langle \tau(A^{-1}), \begin{pmatrix} x \\ 1 \end{pmatrix}\rangle_{d,1}\ .$$

Writing $h(x) - h(A\cdot(B\cdot x))$ as $h(x) - h(B\cdot x) + h(B\cdot x) - h(A\cdot(B\cdot x))$, it follows that τ satisfies the cocycle relation. Uniqueness is given by Fact 16.3.8. □

Now let us check that the discussion is not void. First there are easy examples in the coboundary case.

Fact 16.3.12 *Let τ_v be a coboundary, i.e. there is $v \in \mathbb{R}^{d,1}$ such that $\tau_v(A) = Av - v$. Then $h_v(x) = -\langle x, \bar{v}\rangle_d + v_{d+1}$ is a τ_v-invariant map.*

In full generality, if the cocycle is equal to zero, we know the function $-L$ which is a C^∞ Γ-invariant function with positive definite Hessian. By the very general "Ehresmann–Weil–Thurston holonomy principle" [35], for cocycles close to 0 enough, there exist Γ_τ-equivariant maps which depend continuously on the cocycle. For convenience we recall the argument in our very simplified case, which follows the lines from [18, Lemma I.1.7.2]. We need to take care about convexity, which is also classical [31].

Proposition 16.3.13 *For any cocycle τ there exists a C^∞ convex (resp. concave) Γ_τ-equivariant function $h(\tau)$.*

Moreover, if $\tau_n \to \tau$, then there exist C^∞ convex (resp. concave) Γ_{τ_n}-equivariant functions $h(\tau_n)$ such that $(h(\tau_n))_{n\in\mathbb{N}}$ converges to $h(\tau)$, and the second partial derivatives of $h(\tau_n)$ converge to the second partial derivatives of $h(\tau)$.

Proof Clearly it suffices to prove the statement for the convex case. Also by Fact 16.3.7, it suffices to prove it for any cocycle close to 0.

Let $\{B_i(r_i)\}_{i=1,\dots,k}$ be disjoint open balls of $\mathbb{H}^d$, such that $\Gamma \cdot \cup_i B_i(r_i)$ is a covering of $\mathbb{H}^d$. On $B_1(r_1)$, let us set $h_1 = -L$. For $A \in \Gamma$ and $y \in A \cdot B_1(r_1)$, let us set $h_1(y) = -L(y) + \langle \binom{y}{1}, \tau(A)\rangle_{d,1}$. Such a function h_1 is C^∞ and Γ_τ-equivariant on $\Gamma \cdot B_1(r_1)$. The function h_1 converges to $-L$ uniformly on each orbit of $B_1(r_1)$ if τ goes to 0. Also the first partial derivatives of h_1 converge to the ones of $-L$ uniformly on each orbit of $B_1(r_1)$ if τ goes to 0. Moreover, the Hessian of h_1 is equal to the one of $-L$ on $\Gamma \cdot B_1(r_1)$, in particular it is positive definite.

Let $r_i' < r_i$ for all i, such that $\Gamma \cdot \cup_i B_i(r_i')$ is still a covering of $\mathbb{H}^d$. Up to change the indices, suppose that $B_2(r_2)$ has non empty intersection with the orbit of $B_1(r_1)$. Let W be an open neighborhood of $B_2(r_2') \cap \Gamma \cdot B_1(r_1')$ such that its closure is contained in $B_2(r_2) \cap \Gamma \cdot B_1(r_1)$. Let ϕ be a bump function which is equal to 1 on $B_2(r_2') \cap \Gamma \cdot B_1(r_1')$ and whose support is contained in W. Note that the function ϕh_1 is well-defined and C^∞ on $\mathbb{H}^d$, by setting the zero value out of W.

Let us define $f = \phi h_1 + (1-\phi)(-L)$ on $B_2(r_2)$. The function f is C^∞, and equal to h_1 on $\Gamma \cdot B_1(r_1') \cap B_2(r_2')$. When the cocycle goes to 0, f and its first and second derivatives go to $-L$ and to its respective derivatives, uniformly on $B_2(r_2')$. In particular, we suppose that the cocycle is sufficiently small, so that the Hessian of f is positive definite.

Then we define $h_2 = f$ on $B_2(r_2')$, and by equivariance we define h_2 on $\Gamma \cdot B_2(r_2')$. Also we set $h_2 = h_1$ on $\Gamma \cdot B_1(r_1')$. By construction, h_2 is well defined on the non-empty intersections between orbits of $B_1(r_1')$ and orbits of $B_2(r_2')$. Clearly, h_2 converges to $-L$ when the cocycle goes to 0. As the Hessian of h_2 converges to the one of $-L$ uniformly on $B_2(r_2')$, by Fact 16.3.9, this is true on each element in the orbit of $B_2(r_2')$, in particular the Hessian of h_2 is positive definite.

In the same way, if $r_i'' < r_i'$ is such that $\Gamma \cdot \cup_i B_i(r_i'')$ is still a covering of $\mathbb{H}^d$, then we can construct a function h_3, equivariant on the orbit of $B_1(r_1'') \cup B_2(r_2'') \cup B_3(r_3'')$ and satisfying the statement of the proposition. After a finite number of steps, we have constructed the wanted functions.

□

Corollary 16.3.14 *For any $\tau \in Z^1(\Gamma, \mathbb{R}^{d,1})$, there exists a continuous map $b_\tau : \partial B \to \mathbb{R}$ such that any Γ_τ-equivariant map extends continuously as b_τ on ∂B.*

Moreover if $\tau_1, \tau_2 \in Z^1(\Gamma, \mathbb{R}^{d,1})$ and $\alpha \in \mathbb{R}$, then $b_{\alpha\tau_1+\tau_2} = \alpha b_{\tau_1} + b_{\tau_2}$. And τ is a coboundary if and only if b_τ is the restriction to ∂B^d of an affine map of $\mathbb{R}^d$.

Proof From Proposition 16.3.13, there exists a C^∞ convex Γ_τ-equivariant map. From Fact 16.3.9, **Mean**(h) is a Γ-invariant function, hence bounded, so by Proposition 16.2.34, there exists a continuous function $b_\tau : \partial B \to \mathbb{R}$ which extends continuously h. As the difference of two Γ_τ-equivariant map is a Γ-invariant function, and as a Γ-invariant function extends continuously as the zero function on

∂B (Fact 16.3.5), it follows that b_τ is the continuous extension of any Γ_τ-equivariant map.

The second property is obvious from the definition of b_τ and Fact 16.3.7. The last property follows from Fact 16.3.12 □

From the existence of b_τ we deduce easily the existence of a unique Γ_τ-equivariant mean map in the following lemma. The maps whose graphs are the boundary of the convex hull of the graph of b_τ will be introduced in Sect. 16.3.3.

Corollary 16.3.15 *Let $\tau \in Z^1(\Gamma, \mathbb{R}^{d,1})$. There exists a unique C^∞ Γ_τ-equivariant map, denoted by h_τ^{mean}, satisfying* $\mathbf{Mean}(h_\tau^{\text{mean}}) = 0$. *Moreover, for $\alpha \in \mathbb{R}$ and $\tau' \in Z^1(\Gamma, \mathbb{R}^{d,1})$, $h_{\tau+\alpha\tau'}^{\text{mean}} = h_\tau^{\text{mean}} + \alpha h_{\tau'}^{\text{mean}}$, and h_τ^{mean} is the restriction to B^d of an affine map if and only if τ is a coboundary.*

Proof By Corollary 16.3.14 and Proposition 16.2.37, we know that there exists a unique C^∞ map, denoted by h_τ^{mean}, having b_τ as values on ∂B, and such that $\mathbf{Mean}(h_\tau^{\text{mean}}) = 0$. This map is Γ_τ-equivariant. Indeed, apply an element of Γ_τ to the graph of h_τ^{mean}. Then we obtain the graph of a map with vanishing **Mean** and boundary value b_τ, so it has to be h_τ^{mean} by uniqueness.

The second point is clear from Fact 16.3.7 and the fact that h_τ^{mean} satisfies a linear equation, namely $(\frac{1}{d}\Delta^{\mathbb{H}^d} - 1)h_\tau^{\text{mean}} = 0$. The last point is immediate from Corollary 16.3.14. □

Remark 16.3.16 For any $t \in \mathbb{R}$, the map $h_\tau^{\text{mean}} - tL$ is Γ_τ-equivariant, with mean curvature equal to t. Hence the graphs of these maps gives a smooth foliation of $^*\mathbb{R}^{d,1}/\Gamma_\tau$ by hypersurfaces of constant mean curvature.

Corollary 16.3.15 allows to recover a classical relation between cocycles and traceless Codazzi tensors [15, 48, 53]. Let $\mathcal{C}od_0^\Gamma$ be the vector space of traceless symmetric Codazzi tensors on $\mathbb{H}^d/\Gamma$. Let $\tau \in Z^1(\Gamma, \mathbb{R}^{d,1})$. By Corollary 16.3.15, there is a map h_τ^{mean} whose second fundamental form is a Γ-invariant traceless Codazzi tensor (see Remark 16.3.10), hence it defines an element of $\mathcal{C}od_0^\Gamma$, denoted by $\mathbf{Cod}(\tau)$. By Corollary 16.3.15, the map $\mathbf{Cod} : Z^1(\Gamma, \mathbb{R}^{d,1}) \to \mathcal{C}od_0^\Gamma$ is linear. The kernel of this map corresponds to the τ such that h_τ^{mean} is affine, hence to the coboundaries by Corollary 16.3.15. We thus obtain an injective morphism from $H^1(\Gamma, \mathbb{R}^{d,1})$ to $\mathcal{C}od_0^\Gamma$, still denoted by **Cod**.

Proposition 16.3.17 *The map* $\mathbf{Cod} : H^1(\Gamma, \mathbb{R}^{d,1}) \to \mathcal{C}od_0^\Gamma$ *is an isomorphism.*

Proof Let $C \in \mathcal{C}od_0^\Gamma$, which defines a Γ-invariant symmetric traceless Codazzi tensor $\tilde{C}$ on $\mathbb{H}^d$. By Lemma 16.2.13, there exists $h : B^d \to \mathbb{R}$ such that $\tilde{C} = \mathrm{I\!I}_h$. From Lemma 16.3.11, there exists a cocycle τ such that h is Γ_τ-equivariant, hence as $\mathrm{I\!I}_h$ is traceless, by the uniqueness part of Corollary 16.3.15, we will have $h = h_\tau^{\text{mean}}$.

□

16.3.3 Volume of the Convex Core and Asymmetric Norm

16.3.3.1 Convex Core

Let $\tau \in Z^1(\Gamma, \mathbb{R}^{d,1})$. There is an associated map $b_\tau : \partial B^d \to \mathbb{R}$ given by Corollary 16.3.14. This map has a graph $\Lambda(b_\tau)$, and we will look at its convex hull $\mathrm{CH}(\tau)$ in the affine space $\mathbb{R}^{d+1}$, as well as the functions $h^-_{b_\tau}$ and $h^+_{b_\tau}$ (see Sect. 16.2.3.3) whose graphs are the boundary of $\mathrm{CH}(\tau)$. We will denote these two last maps by h^-_τ and h^+_τ respectively.

The argument to check the following fact is analogous to the one used in the proof of Corollary 16.3.15.

Fact 16.3.18 *The map h^-_τ and h^+_τ are Γ_τ-equivariant, in particular $\mathrm{CH}(\tau)$ is globally invariant for the action of Γ_τ.*

Lemma 16.3.19 *Let $\tau \in Z^1(\Gamma, \mathbb{R}^{d,1})$. Then:*

1. *for any convex (resp. concave) Γ_τ-equivariant map h, we have $h \leq h^-_\tau$ (resp. $h \geq h^+_\tau$),*
2. $h^+_\tau = -h^-_{-\tau}$,
3. *For $\alpha > 0$, $h^-_{\alpha\tau} = \alpha h^-_\tau$,*
4. $h^-_\tau + h^-_{\tau'} \leq h^-_{\tau+\tau'}$ *and* $h^+_\tau + h^+_{\tau'} \geq h^+_{\tau+\tau'}$.

Proof The two first points are from the definitions of h^+_τ and h^-_τ, Proposition 16.2.40 and Corollary 16.3.14. The third point follows from (16.30) and the fact that $b_{\alpha\tau} = \alpha b_\tau$. The fourth point follows from the first point, as $h^-_\tau + h^-_{\tau'}$ is a convex $\Gamma_{\tau+\tau'}$-equivariant function. □

Lemma 16.3.20 *Let τ_v be a coboundary, then $h^-_{\tau_v} = h^{\mathrm{mean}}_{\tau_v}$ is an affine map and $h^-_{\tau+\tau_v} = h^-_\tau + h^-_{\tau_v}$. Conversely, if $h^-_\tau = h^{\mathrm{mean}}_\tau$, then h^{mean}_τ is affine and τ is a coboundary.*

Proof If τ is a coboundary, we know that there exists a Γ_τ-equivariant affine map (Fact 16.3.12). Hence the convex hull of $\Lambda(b_\tau)$ is a piece of a hyperplane, and this hyperplane is also the τ_v-mean hypersurface. Then $h^-_{\tau+\tau_v} = h^-_\tau + h^-_{\tau_v}$ follows from (16.30) because h_{τ_v} is an affine map. For the second part, on the one hand, $\mathbf{Mean}(h^{\mathrm{mean}}_\tau) = 0$. On the other hand, if $h^{\mathrm{mean}}_\tau = h^-_\tau$, then h^{mean}_τ is convex, hence affine (Remark 16.2.33), so b_τ is the restriction to ∂B^d of an affine map. By Corollary 16.3.14, τ is a coboundary.

□

Definition 16.3.21 The *convex core* of ${}^*\mathbb{R}^{d,1}/\Gamma_\tau$, denoted by $\mathrm{CC}(\tau)$, is the smallest non-empty convex set of ${}^*\mathbb{R}^{d,1}/\Gamma_\tau$.

In the above definition, "convex" has to be understood in the strong sense of geodesically convex: C is convex if for $x, y \in C$, any geodesic between x and y belongs to C. So for example, a single point or a small open ball may not be convex. In the cylindrical model of the universal cover, this notion of convexity coincides with the affine one.

Clearly, $\mathrm{CC}(\tau) = \mathrm{CH}(\tau)/\Gamma_\tau$. Hence ${}^*\mathbb{R}^{d,1}/\Gamma_\tau$ has a compact convex core, so the action of Γ_τ on ${}^*\mathbb{R}^{d,1}$ is *convex cocompact*, in the sense of [22, 23].

Recall the volume form on co-Minkowski space, Sect. 16.2.2.6. Let us denote by Vol the induced volume on ${}^*\mathbb{R}^{d,1}$. It is then immediate than for any $\tau \in Z^1(\Gamma, \mathbb{R}^{d,1})$,

$$\mathrm{Vol}(\mathrm{CC}(\tau)) = \int_{\mathbb{H}^d/\Gamma} L^{-1}(h_\tau^+ - h_\tau^-)\ . \tag{16.38}$$

Here by abuse of notation, we denote in the same way the Γ-invariant function $h_\tau^+ - h_\tau^-$ and the corresponding function on the compact hyperbolic manifold $\mathbb{H}^d/\Gamma$. (Recall that we defined a function h to be Γ-invariant if $L^{-1}h$ is invariant under the action of Γ.) The integration is implicitly with respect to the volume form given by the hyperbolic metric.

Definition 16.3.22 The function $\mathbf{vol} : H^1(\Gamma, \mathbb{R}^{d,1}) \to \mathbb{R}$ associates $\mathrm{Vol}(\mathrm{CC}(\tau))$ to any representative τ of an element of $H^1(\Gamma, \mathbb{R}^{d,1})$.

By Lemma 16.3.20, **vol** is well-defined. Actually, the following result is straightforward to check from Lemmas 16.3.20 and 16.3.19.

Proposition 16.3.23 **vol** *is a norm on* $H^1(\Gamma, \mathbb{R}^{d,1})$.

Remark 16.3.24 The volume of the convex core can be geometrically interpreted in Minkowski space as a mean width defined as follows. From a cocycle τ, we have the boundary map b_τ which defines two convex sets $\Omega^+_{b_\tau}$ and $\Omega^-_{b_\tau}$ in Minkowski space, see Remark 16.2.44. It follows from the previous section that those two sets (here denoted by Ω^+_τ and Ω^-_τ) are invariant under the action of Γ_τ on Minkowski space. Actually the action is free and properly discontinuous on $\Omega^+_\tau \cup \Omega^-_\tau$, and the quotient of Ω^-_τ (resp. Ω^+_τ) is a *future complete flat* (resp. *past complete*) Globally Hyperbolic Maximal Cauchy Compact (in short, GHMC) spacetime. As the addition of a coboundary to the cocycle τ will only change the origin in Minkowski space, then $H^1(\Gamma, \mathbb{R}^{d,1})$ parameterizes the space of future complete (or past complete) flat GHMC spacetimes with a given linear holonomy, up to conjugacy. See [3, 12] for more details.

Moreover, for any cocycle τ, we have that $-\Omega^-_\tau = \Omega^+_{-\tau}$. Now, for any $x \in B^d$, let us denote by $\mathbf{width}(x)$ the Lorentzian distance between the support plane of Ω^+_τ with outward unit normal $\binom{x}{1}$, and the support plane of Ω^-_τ with inward unit normal $\binom{x}{1}$. Note that the map $x \mapsto \mathbf{width}(x)$ is Γ-invariant. Then the mean width, defined as $\int_{\mathbb{H}^d/\Gamma} \mathbf{width}(\cdot)$, is given by (16.38), see Fig. 16.8.

16.3.3.2 Asymmetric Norm

In the previous section we showed that the volume of the convex core is a norm on $H^1(\Gamma, \mathbb{R})$. We now see that it is actually the symmetrization of an asymmetric norm on $H^1(\Gamma, \mathbb{R})$.

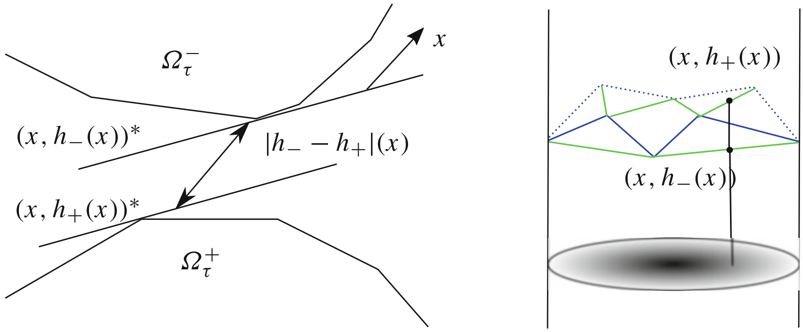


Fig. 16.8 The volume of the convex core $CC(\tau)$ is a "mean distance" between a past complete flat GHMC spacetime and a future complete flat GHMC spacetime with the same holonomy

For a cocycle τ, the S_1 norm is defined as follows

$$\|\tau\|_{S^1} = \int_{\mathbb{H}^d/\Gamma} L^{-1}(h_\tau^{\text{mean}} - h_\tau^-) \ . \tag{16.39}$$

(The denomination will be motivated in Remark 16.3.26.)

By Lemma 16.3.20, if τ_v is a coboundary, then $\|\tau + \tau_v\| = \|\tau\| + \|\tau_v\| = \|\tau\|$. Hence $\|\cdot\|$ is well defined on $H^1(\Gamma, \mathbb{R}^{d+1})$.

Proposition 16.3.25 *The S_1 norm $\|\cdot\|_{S_1}$ defines an asymmetric norm on $H^1(\Gamma, \mathbb{R}^{d+1})$, i.e.* $\forall[\tau], [\tau'] \in H^1(\Gamma, \mathbb{R}^{d+1})$

1. $\|[\tau]\|_{S_1} \geq 0$;
2. $\|[\tau]\|_{S_1} = 0$ *if and only if* $[\tau] = 0$;
3. $\|[\tau] + [\tau']\|_{S_1} \leq \|[\tau]\|_{S_1} + \|[\tau']\|_{S_1}$;
4. $\forall \alpha \geq 0$, $\|\alpha[\tau]\|_{S_1} = \alpha\|[\tau]\|_{S_1}$.

Proof The first property comes from Lemma 16.2.42. The second point is Lemma 16.3.20. The third and forth points are immediate consequence of Lemma 16.3.19 and Corollary 16.3.15. □

It is obvious from (16.39) and (16.38) that **vol** is the symmetrization of $\|\cdot\|_{S_1}$:

$$\mathbf{vol}([\tau]) = \frac{1}{2}\left(\|[\tau]\|_{S_1} + \|-[\tau]\|_{S_1}\right) \ .$$

16.3.3.3 Mean Curvature Measure

We now explain how $\|\cdot\|_{S^1}$ is related to the mean curvature measure introduced in Sect. 16.2.3.4. From Lemma 16.2.47, we have that for any convex Γ_τ-equivariant map h, the measure $\mathbf{MM}(h)$ is Γ-invariant, and then defines a Radon measure $\mathbf{MM}^{\Gamma}(h)$ on $\mathbb{H}^d/\Gamma$. Actually, there is a nice expression for this measure. Let h be a convex Γ_τ-equivariant function. By definition of h_τ^{mean}, $\mathbf{MM}(h) = \mathbf{MM}(h) - \mathbf{MM}(h_\tau^{\text{mean}}) = \mathbf{MM}(h - h_\tau^{\text{mean}})$. On the other hand, $h - h_\tau^{\text{mean}}$ is Γ-invariant, so from (16.32), using a partition of unity, one obtains that for any C^∞ function φ on $\mathbb{H}^d/\Gamma$,

$$\mathbf{MM}^{\Gamma}(h)(\varphi) = \int_{\mathbb{H}^d/\Gamma} L^{-1}(h - h_\tau^{\text{mean}})(\frac{1}{d}\Delta^{\mathbb{H}^d} - 1)\varphi \,.$$

Taking $\varphi = 1$,

$$\mathbf{MM}^{\Gamma}(h)(\mathbb{H}^d/\Gamma) = \int_{\mathbb{H}^d/\Gamma} L^{-1}(h_\tau^{\text{mean}} - h) \,,$$

in particular, if $h = h_\tau^-$, by definition of the S_1 norm,

$$\|\tau\|_{S^1} = \mathbf{MM}^{\Gamma}(h_\tau^-)(\mathbb{H}^d/\Gamma) \,. \tag{16.40}$$

Remark 16.3.26 Consider the convex set Ω_τ^- in Minkowski space, as well as the ϵ-equidistant convex set $\Omega_\tau^-(\epsilon)$ (this is the dual convex set in Minkowski space of the epigraph of $h_\tau^- - \epsilon$ in $B^d \times \mathbb{R}$). By a Lorentzian version of the Steiner formula proved in [30], the volume of $(\Omega_\tau^- \setminus \Omega_\tau^-(\epsilon))/\Gamma_\tau$ is a polynomial in ϵ of degree $d+1$. Up to a dimensional constant, the coefficient in front of ϵ^d is nothing but $\mathbf{MM}^{\Gamma}(h_\tau^-)(\mathbb{H}^d/\Gamma)$. The analogous quantity in the classical theory of convex bodies is called the (total) area measure of order one [59], and usually denoted by S_1, which explains our terminology (see also Remark 16.2.55).

Lemma 16.3.27 *Let $\tau_n \to \tau$. Then b_{τ_n} (resp. $h_{\tau_n}^-$) pointwise converges to b_τ (resp. h_τ^-).*

Proof By Proposition 16.3.13, we have convex (resp. concave) Γ_{τ_n}-equivariant functions converging to a Γ_τ-equivariant convex (resp. concave) function. For any n, as the concave and the convex Γ_{τ_n}-equivariant functions coincide on ∂B^d with b_{τ_n} given by Corollary 16.3.14, they bound a convex body K_n of $\mathbb{R}^{d+1}$. Let us denote by K the convex body bounded by the Γ_τ-equivariant convex and concave functions.

Let us denote by C^{d+1} the space of non-empty compact sets of $\mathbb{R}^{d+1}$, endowed with the Hausdorff topology. Suppose that there is a subsequence (K_{n_i}) of (K_n) which converges to K' in C^{d+1}. Then K' is a convex body [59, Theorem 1.8.6]. Moreover, each point of K' is the limit of a sequence of points (x_{n_i}) with $x_{n_i} \in K_{n_i}$ [59, Theorem 1.8.8]. From this it is easy to deduce that $K' = K$.

Now as the Γ_{τ_n}-equivariant functions are converging, they are bounded, and in turn the sequence of convex bodies $(K_n)_n$ is bounded in $B^d \times \mathbb{R} \subset \mathbb{R}^{d+1}$. By the Blaschke selection theorem [59, Theorem 1.8.7], there is a subsequence K_{n_i} converging to a convex body K'. Moreover, the sequence K_n is contained in a compact subspace of C^{d+1} [59, Theorem 1.8.4]. As we saw that any convergent subsequence of (K_n) converges to K, it follows that (K_n) converges to K.

As the limit of any convergent sequence (x_{n_i}) with $x_{n_i} \in K_{n_i}$ must belong to K [59, Theorem 1.8.8], it is easy to deduce that $b_{\tau_n} \to b_\tau$. This easily implies the Hausdorff convergence of $\mathrm{CH}(\tau_n)$ to $\mathrm{CH}(\tau)$, see e.g. [61, Lemma 2.1], which in turn gives the convergence of $h^-_{\tau_n}$ to h^-_τ, as the Hausdorff convergence of convex bodies implies the Hausdorff convergence of the boundaries [59, Lemma 1.8.1]. □

Remark 16.3.28 By standard properties of convex functions, it follows from Lemma 16.3.27 that $h^-_{\tau_n}$ converges to h^-_τ uniformly on any compact set of B^d. But there is no general argument that would allow to deduce the uniform convergence of $(b_{\tau_n})_n$ from the pointwise convergence, see [50] for a counter-example.

However, in our situation, the uniform convergence of $(h^-_{\tau_n})_n$ holds. It will be obtained as a byproduct of the considerations of Sect. 16.4. More precisely, the uniform converge will be proved in Lemma 16.4.11. Anticipating this result, we obtain the following proposition.

Proposition 16.3.29 *The S_1 norm $\|\cdot\|_{S_1} : H^1(\Gamma, \mathbb{R}^{d,1}) \to \mathbb{R}$ is continuous.*

Proof Let $\tau_n \to \tau$. From Lemma 16.3.27, $h^-_{\tau_n}$ converges to h^-_τ, and, using a partition of unity, it is not hard to deduce from Lemma 16.2.46 that $\mathbf{MM}^\Gamma(h^-_{\tau_n})$ weakly converges to $\mathbf{MM}^\Gamma(h^-_\tau)$, so that the result follows from (16.40).

□

16.3.3.4 Simplicial Measured Geodesic Laminations

We use the notation and definitions of Sect. 16.3.1, where we have considered a simplicial measured geodesic lamination λ on the compact hyperbolic manifold $\mathbb{H}^d/\Gamma$. Namely we have supposed that $\mathbb{H}^d/\Gamma$ contains n disjoints embedded totally geodesic hypersurfaces $H_1, \ldots, H_n$ with positive weights ω_i.

Let us push the construction a step forward. For any $y \in B^d$, let $c : [0, 1] \to B^d$ be any curve transverse to $\tilde{L}$ joining the base point $\tilde{x}$ to y, and define

$$h_\lambda(y) = \sum_{j \in c[0,1] \cap \tilde{L}} \omega_j h_{l_j}(y)$$

where h_{l_j} is the canonical map associated with l_j (Definition 16.2.48), where we consider that the outward unit normal n_l has the same direction than the curve at the corresponding point of $c[0, 1] \cap \tilde{L}$.

Fact 16.3.30 *If τ_λ is the cocycle given by* (16.37)*, then $h_\lambda = h^-_{\tau_\lambda}$.*

Proof As the weights are positive, by Fact 16.2.50, h_λ is a convex map.

Let us check that h_λ is Γ_{τ_λ}-equivariant. Let $\tilde{c} : [0, 1] \to B^d$ be a path joining $\tilde{x}$ to $A \cdot \tilde{x}$, and let $c' : [0, 1] \to B^d$ be a path joining $\tilde{x}$ to y, both assumed to be transverse to $\tilde{L}$. Then the concatenation of c with $A \cdot c'$ is a path joining $\tilde{x}$ to $A \cdot y$, hence

$$h_\lambda(A \cdot y) = \sum_{j \in (c([0,1]) \cap \tilde{L})} \omega_j h_{l_j}(y) + \sum_{j \in (A \cdot c'([0,1]) \cap \tilde{L})} \omega_j h_{l_j}(y) ,$$

and as by definition $h_l(y) = \langle \binom{y}{1}, v_l \rangle_{d,1}$, then $\sum_{j \in (c([0,1]) \cap \tilde{L})} \omega_j h_{l_j}(y) = \langle \binom{y}{1}, \tau_\lambda(A) \rangle_{d,1}$. Also, $\sum_{j \in (A \cdot c'([0,1]) \cap \tilde{L})} \omega_j h_{l_j}(y) = \sum_{j \in (c'([0,1]) \cap \tilde{L})} \omega_j h_{A \cdot l_j}(A \cdot y)$, and by Fact 16.2.52, $h_{A \cdot l_j}(A \cdot y) = \frac{L(A \cdot y)}{L(y)} h_l(y)$. The equivariance is proved.

A h_λ is a convex Γ_{τ_γ}-equivariant map, then $h_\lambda \leq h^-_{\tau_\lambda}$. By construction, the graph of h_λ is made of segments joining points of graph of b_{τ_λ}, hence it is contained in $\mathrm{CH}(b_{\tau_\lambda})$, so $h_\lambda \geq h^-_{\tau_\lambda}$.

□

The *length* **length**(λ) of a simplicial measured geodesic lamination λ on $\mathbb{H}^d/\Gamma$ is defined as the sum of the weights times the total volume of the corresponding totally geodesic hypersurfaces. By Lemma 16.2.54 and (16.40), we obtain the following.

Proposition 16.3.31 *Let λ be a simplicial measured geodesic lamination on $\mathbb{H}^d/\Gamma$. Then*

$$\mathbf{length}(\lambda) = \|\tau_\lambda\|_{S^1} .$$

Remark 16.3.32 There is no reason why for $d \geq 3$ any cocycle should arise from a (simplicial) measured geodesic lamination on $\mathbb{H}^d/\Gamma$. So for $d \geq 3$, the concept of measured geodesic lamination is not sufficient. A more suitable concept is the one of *measured geodesic stratification*, introduced in [12]. In contrast, we will see in the next section that for $d = 2$, any cocycle arises from a measured geodesic lamination.

16.3.4 The Case of Dimension $2+1$

In this part we study the particularities of the $d = 2$ case. We will denote by Teich_S the Teichmüller space of a compact surface homeomorphic to $\mathbb{H}^2/\Gamma$. We will denote by g the genus of S.

16.3.4.1 Goldman Isomorphism

The Teichmüller space Teich_S can be defined as the space of faithful and discrete representations of $\pi_1 S$ into $\mathrm{Isom}_0(\mathbb{H}^2)$ up to conjugacy. Let ρ be such a representation such that $\Gamma = \rho(\pi_1 S)$. Then the tangent space of Teich_S at ρ is naturally identified with $H^1(\pi_1(S), \mathfrak{isom}(\mathbb{H}^2))$, where $\pi_1 S$ acts on the Lie group $\mathfrak{isom}(\mathbb{H}^2)$

via Ad ρ [34]. Using the hyperboloid model $\mathcal{H}^2$ for $\mathbb{H}^2$, $\mathfrak{isom}(\mathbb{H}^2)$ can be identified with $\mathfrak{o}(2, 1)$. Let us write it as follows.

Theorem 16.3.33 ([34]) *There is a vector space isomorphism*

$$\mathbf{Gold} : H^1(\Gamma, \mathfrak{o}(2, 1)) \to T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S\ .$$

There is also a one-to-one correspondence between vectors of $\mathbb{R}^{2,1}$ and infinitesimal Minkowski isometries of $\mathfrak{o}(2, 1)$. This may be seen for example using the Minkowski cross product, see e.g. [25]. This identification gives a vector space isomorphism

$$\mathbf{C} : H^1(\Gamma, \mathbb{R}^{2,1}) \to H^1(\Gamma, \mathfrak{o}(2, 1))\ ,$$

and in turn we have the following vector space isomorphism

$$\xi = \mathbf{Gold} \circ \mathbf{C} : H^1(\Gamma, \mathbb{R}^{2,1}) \to T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S\ .$$

In particular, we obtain the following.

Corollary 16.3.34 *The vector space* $H^1(\Gamma, \mathbb{R}^{2,1})$ *has dimension* $6g - 6$.

16.3.4.2 Mess Homeomorphism

Let us call an *entire segment* of B^2 a segment of B^2 whose endpoints are in ∂B^2. A *geodesic lamination* $\tilde{L}$ of B^2 is a non-empty closed union of disjoint entire segments of B^2. Let $\tilde{L}$ be a geodesic lamination on B^2 which is invariant under the action of Γ. Then the image L of $\tilde{L}$ under the projection is a geodesic lamination on the compact hyperbolic surface B^2/Γ. A *measured geodesic lamination* $\lambda = (L, \mu)$ on B^2/Γ is the data of a geodesic lamination L together with a *transverse measure* μ, that is, the data of a Radon measure on each compact rectifiable curve transverse to L, such that

- the support of the measure is the intersection of the arc with L,
- if two arcs are homotopic through arcs transverse to L, then the homotopy sends the measure on one segment to the measure on the other one.

A simplicial measured geodesic lamination on B^2/Γ is a set of non-intersecting closed simple geodesics weighted by positive numbers. Note that the action of Γ onto B^2 is via the identification of the disc with the Klein model of the hyperbolic plane, but the notation B^2 stands for reminding the affine nature of the measured geodesic lamination on the universal cover.

Let $\mathcal{ML}_\Gamma$ be the set of measured geodesic laminations on the compact hyperbolic surface $\mathbb{H}^2/\Gamma$. $\mathcal{ML}_\Gamma$ is endowed with the following topology. We say that λ_n converges to λ if, for any compact segment c transverse to L we have

- c is transverse to L_n for n big,
- μ_n weakly converges to μ on c.

We have the following classical result of Thurston, see e.g. [11] and the references therein.

Theorem 16.3.35 (Thurston) *For the topology defined above, $\mathcal{ML}_\Gamma$ is a manifold of dimension $6g - 6$.*

Recall from (16.35) that a vector v_l of $\mathbb{R}^{2,1}$ is assigned to any entire segment l of B^2. Let $\mathbf{e}$ be a continuous function such that, for any path $c : [0, 1] \to B^2$ transverse to L, $\mathbf{e}_L(c(t)) = v_l$ if $c(t) \in l$, where v_l has the same direction as the curve at $c(t)$, and $l \in \tilde{L}$. Let us fix an arbitrary base point $\tilde{x} \in B^2$. Then define, for $A \in \Gamma$, and for any path $c : [0, 1] \to B^2$ transverse to L joining $\tilde{x}$ and $A \cdot \tilde{x}$:

$$\tau_\lambda(A) = \int_0^1 \mathbf{e}_L(c(t))\mathrm{d}\mu(t) \ . \tag{16.41}$$

As the measure is transverse, the definition of τ_λ is independent from the choice of the path c and the function $\mathbf{e}_L$. The following fact is proved formally in the same way as Facts 16.3.2 and 16.3.3.

Fact 16.3.36 *We have $\tau_\lambda \in Z^1(\Gamma, \mathbb{R}^{2,1})$. Moreover, if the basepoint is changed, the new cocycle differs from the preceding one by a coboundary.*

Hence we have constructed a well-defined map

$$\mathbf{Mess} : \mathcal{ML}_\Gamma \to H^1(\Gamma, \mathbb{R}^{d,1}) \ .$$

Theorem 16.3.37 ([51]) *The map* **Mess** *defined above is a homeomorphism.*

Proof The map is clearly injective and continuous. By Theorem 16.3.35 and Corollary 16.3.34, both $\mathcal{ML}_\Gamma$ and $H^1(\Gamma, \mathbb{R}^{d,1})$ are manifolds of same dimension. Hence by the invariance of domain theorem **Mess** is a local homeomorphism. Now for $\lambda \in \mathcal{ML}_\Gamma$ and $t \geq 0$, let us define $t\lambda$ as the measured geodesic lamination obtained from λ by simply multiplying the transverse measure by t. By (16.37) we clearly have $\mathbf{Mess}(t\lambda) = t\mathbf{Mess}(\lambda)$. As $H^1(\Gamma, \mathbb{R}^{2,1})$ is a vector space and **Mess** a local homeomorphism, it follows that **Mess** is surjective. □

Remark 16.3.38 It is possible to describe the inverse map to **Mess**, by defining a "bending measure" belonging to $\mathcal{ML}_\Gamma$ from the graph of h^-_τ, for any cocycle τ. There are at least three ways to define such a bending measure. The first one is to mimic the construction of the bending measure given by the upper boundary component of the convex core of a hyperbolic quasi-Fuchsian manifolds [26]. The second one is to define, as in [51], the induced distance on the spacelike part of the boundary of Ω_τ^-, the dual of the epigraph of h_τ^- in Minkowski space. The last one is to consider the mean curvature measure given by h_τ^-.

16.3.4.3 Length of Measured Geodesic Laminations

We have encountered the length of simplicial measured geodesic laminations in Sect. 16.3.3.4. For $d = 2$, the length of a measured geodesic lamination is defined as the total mass on the surface of the measure which is the product of the hyperbolic measure along the leaves of the lamination and the measure transverse to the leaves. We refer to [11] for more details. Actually, the simplicial case suffices, as the following results shows. One may see for example Lemma 2.4 in [44] for the first one, and Theorem 3.1.3 in [56] or Section 3.4.3 in [8] for the second one.

Proposition 16.3.39 *The map* **length** : $\mathcal{ML}_\Gamma \to \mathbb{R}$ *is continuous.*

Proposition 16.3.40 *Simplicial measured geodesic laminations are dense in* $\mathcal{ML}_\Gamma$.

So from the above results, Proposition 16.3.31 generalizes as follows.

Proposition 16.3.41 *Let* $\lambda \in \mathcal{ML}_\Gamma$. *Then*

$$\mathbf{length}(\lambda) = \|\mathbf{Mess}(\lambda)\|_{S^1} \ .$$

16.3.4.4 Thurston Earthquake Norm

From a measured geodesic lamination λ on $\mathbb{H}^2/\Gamma$, one obtains another hyperbolic metric on S by performing a (left) earthquake along the lamination. We refer to [44] and the reference therein for more details about earthquakes. Actually for t near 0, earthquakes along $t\lambda$ define a path in Teich_S starting at $\mathbb{H}^2/\Gamma$. This path has a well defined derivative at 0, which gives an element in $T_{\mathbb{H}^2/\Gamma}\text{Teich}_S$, the tangent space of Teichmüller space at the point $\mathbb{H}^2/\Gamma$. In turn, we have an *infinitesimal earthquake map*:

$$\mathbf{InfEarth} : \mathcal{ML}_\Gamma \to T_{\mathbb{H}^2/\Gamma}\text{Teich}_S \ .$$

Theorem 16.3.42 ([44, Proposition 2.6]) *The map* **InfEarth** *is a homeomorphism.*

So the map $\mathbf{InfEarth} \circ \mathbf{Mess}^{-1}$ provides a homeomorphism between $H^1(\Gamma, \mathbb{R}^{2,1})$ and $T_{\mathbb{H}^2/\Gamma}\text{Teich}_S$. Although there is no natural vector space structure on $\mathcal{ML}_\Gamma$, we have the following.

Proposition 16.3.43 ([14, Proposition B.3]) *We have* $\mathbf{InfEarth} \circ \mathbf{Mess}^{-1} = \xi$. *In particular,* $\mathbf{InfEarth} \circ \mathbf{Mess}^{-1}$ *is a vector space isomorphism.*

In other terms, as $\xi = \mathbf{Gold} \circ \mathbf{C}$, the following diagram commutes:

$$\begin{array}{ccc} \mathcal{ML}_\Gamma & \xrightarrow{\mathbf{InfEarth}} & T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S \\ \mathbf{Mess}\downarrow & & \uparrow\mathbf{Gold} \\ H^1(\Gamma, \mathbb{R}^{2,1}) & \xrightarrow{\mathbf{C}} & H^1(\Gamma, \mathfrak{o}(2,1)) \end{array} .$$

Definition 16.3.44 Let $X \in T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$. The *earthquake norm* of X is

$$\|X\|_{\mathrm{earth}} = \mathbf{length}(\mathbf{InfEarth}^{-1}(X)) .$$

From Propositions 16.3.41 and 16.3.43, one has in fact

$$\|X\|_{\mathrm{earth}} = \|\xi^{-1}(X)\|_{S_1}$$

and as ξ is a vector space isomorphism, from Proposition 16.3.25, one finally obtains the following result.

Theorem 16.3.45 ([63, Theorem 5.2]) *The earthquake norm is an asymmetric norm on* $T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$.

Remark 16.3.46 There is a smooth analogue of Proposition 16.3.43 proved in [15]. Namely, Proposition 16.3.17 gives a map **Cod** from $H^1(\Gamma, \mathbb{R}^{d,1})$ to $\mathcal{C}od_0^\Gamma$, the space of traceless symmetric Codazzi tensors on $\mathbb{H}^2/\Gamma$. In dimension 2, there is also an isomorphism **InfDef** from $\mathcal{C}od_0^\Gamma$ to $T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$, where a (0, 2)-tensor is seen as an infinitesimal deformation of the hyperbolic metric (see [64], where such tensors are called TT, for *transverse traceless*). Then, if J is the almost complex structure of $T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$, the following diagram commutes:

$$\begin{array}{ccc} \mathcal{C}od_0^\Gamma & \xrightarrow{\mathbf{InfDef}} & T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S \\ \mathbf{Cod}^{-1}\downarrow & & \downarrow J \\ H^1(\Gamma, \mathbb{R}^{2,1}) & \xrightarrow{\xi} & T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S \end{array} .$$

16.3.4.5 Thurston Length Norm

Following [63], we note that two successive identifications of the tangent space $T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$ of Teichmüller space with the cotangent space $T^*_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$ will permit to define another asymmetric norm on $T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$, that is actually the Thurston length norm, that induces the Thurston asymmetric distance (see Sect. 16.1).

A first identification between $T^*_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$ and $T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$ is given by the Weil–Petersson form of Teichmüller space, that is a symplectic form on $T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$. For $\alpha \in T^*_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$, let $\alpha^\sharp$ be the dual element in $T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$ of α for the symplectic form. Then define naturally

$$\|\alpha\|^*_{\mathrm{length}} := \|\alpha^\sharp\|_{\mathrm{earth}} \,.$$

On the other hand, for any vector space E endowed with an asymmetric norm N, then its dual E^* is endowed with the asymmetric norm N^* defined, for $v \in E^*$, by

$$N^*(v) := \sup\left\{\frac{v(x)}{N(x)} | x \in E \setminus \{0\}\right\} \,.$$

Applying this to the cotangent space of Teichmüller space endowed with $\|\cdot\|^*_{\mathrm{length}}$, we obtain a new asymmetric norm on the tangent space of Teichmüller space.

Definition 16.3.47 Let $X \in T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$. The *length norm* of X is

$$\|X\|_{\mathrm{length}} = \sup\left\{\frac{\alpha(X)}{\|\alpha^\sharp\|_{\mathrm{earth}}} | \alpha \in T^*_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S \setminus \{0\}\right\} \,.$$

If it possible to describe more precisely the length norm, using a famous result of Wolpert. Let λ be a measured lamination on the surface S. The function $\mathbf{length}(\lambda)$ on the Teichmüller space of S is defined as follows: for each choice of a hyperbolic metric on S, $\mathbf{length}(\lambda)$ is the length of the corresponding measured geodesic lamination. Due to a formula of Wolpert [66, Lemma 4.1], the tangent vector $\mathbf{InfEarth}(\lambda)$ of the Teichmüller space of S at a point $\mathbb{H}^2/\Gamma$ is the symplectic gradient of the function $\mathbf{length}(\lambda)$ at the same point, with respect to the Weil–Petersson form of Teichmüller space:

$$d\mathbf{length}(\lambda)^\sharp = \mathbf{InfEarth}(\lambda) \,,$$

in particular,

$$\|d\mathbf{length}(\lambda)\|^*_{\mathrm{length}} = \mathbf{length}(\lambda) \,.$$

Theorem 16.3.42 together with Wolpert's result gives an identification between the cotangent space of Teichmüller space and $\mathcal{ML}_\Gamma$ [46, Lemma 2.3]. In consequence we obtain the following.

Theorem 16.3.48 ([63, Theorem 5.1]) *Let* $\alpha \in T^*_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$*, and* λ *such that* $\alpha = d\,\mathbf{length}(\lambda)$*. Then*

$$\|\alpha\|^*_{\mathrm{length}} = \mathbf{length}(\lambda)$$

defines an asymmetric norm on $T^*_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$.

And finally, the length norm on $T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$ can be written as follows: for $X \in T_{\mathbb{H}^2/\Gamma}\,\mathrm{Teich}_S$,

$$\|X\|_{\mathrm{length}} = \sup\left\{\frac{\mathrm{d}\,\mathbf{length}(\lambda)(X)}{\mathbf{length}(\lambda)} | \lambda \in \mathcal{ML}_\Gamma \setminus \{0\}\right\} . \tag{16.42}$$

16.4 Anosov Representations

In all this section, Γ is a cocompact lattice of $O_+(d,1)$, and τ an element of $Z^1(\Gamma, \mathbb{R}^{d,1})$. We will consider the associated group Γ_τ of isometries of co-Minkowski space. The aim of this section is to provide an alternative proof (Proposition 16.4.10) of the existence and uniqueness of the τ-invariant map b_τ already exhibited in Lemma 16.3.14. This proof involves the *Anosov character* of Γ_τ as a representation of the hyperbolic group Γ into the group of isometries of the Minkowski space. As a by-product, we will see that the convergence in Lemma 16.3.27 is not only pointwise, but uniform (Lemma 16.4.11).

We start by the following fundamental observation: since stabilizers of points are non-compact, there is no $O_+(d,1) \ltimes \mathbb{R}^{d,1}$ invariant metric on the boundary of co-Minkowski space. However, if one fixes an element x_0 in $\mathcal{H}^d$, then $x \mapsto \langle x, x\rangle_{d,1} + 2\langle x, x_0\rangle_{d,1}^2$ is a positive definite form, hence a Euclidean metric on $\mathbb{R}^{d,1}$, which we denote by $\langle\cdot,\cdot\rangle_{x_0}$.

The choice of x_0 also induces a splitting $\mathbb{R}^{d,1} \approx \mathbb{R}^d \times \mathbb{R}$: here, $\mathbb{R}$ is the linear subspace spanned by x_0, and $\mathbb{R}^d$ is the orthogonal of x_0 for the Minkowskian scalar product. Until now, when writing $\partial^*\mathbb{R}^{d,1} \approx \partial B^d \times \mathbb{R}$, we were always implicitly doing the choice $x_0 = (0, \ldots, 0, 1)$, but in this section we will also consider other choices. What is relevant for us now, is that the choice of x_0 induces Riemannian metrics on $\partial^*\mathbb{R}^{d,1}$: for example, the one making ∂B^d and $\mathbb{R}$ orthogonal, and whose restrictions to ∂B^d and $\mathbb{R}$ are the ones induced by the Euclidean metric $\langle\cdot,\cdot\rangle_{x_0}$. But it is not precisely the one we will actually use, and we now describe a distance function d_{x_0} on $\partial^*\mathbb{R}^{d,1}$.

Let us be more precise: we can define $\partial^*\mathbb{R}^{d,1}$ as the space of lightlike affine hyperplanes of Minkowski space. Once we have fixed the unit timelike vector x_0, we can parametrize $\partial^*\mathbb{R}^{d,1}$ by pairs (w, h) where:

- w is a future lightlike vector in $\mathbb{R}^{d,1}$ in the affine spacelike hyperplane H_{x_0} of equation $\langle x_0, w\rangle_{d,1} = -1$ (therefore, H_{x_0} is the hyperplane tangent to $\mathcal{H}^d$ at x_0, and the set of future lightlike vectors lying in H_{x_0} is the unit sphere in this Euclidean space),
- h any real number.

The associated lightlike affine hyperplane is then the one given by the equation:

$$\langle w, \cdot\rangle_{d,1} = -h .$$

The distance function we will actually use is the following one:

$$d_{x_0}((w_1,h_1),(w_2,h_2)) = \sqrt{\langle w_1-w_2, w_1-w_2\rangle_{d,1} + (h_1-h_2)^2}$$
$$= \sqrt{-2\langle w_1, w_2\rangle_{d,1} + (h_1-h_2)^2}\ .$$

We will also consider the closed hyperbolic manifold $N = \Gamma\backslash\mathcal{H}^d$, and the geodesic flow ϕ^t on the unitary tangent bundle $M = T^1N$. Recall that, for any element v of M, the image $\phi^t(v)$ is the unique vector tangent to the geodesic starting from v and at distance t along this geodesic.

Actually, M is the quotient of the unitary tangent bundle $T^1\mathcal{H}^d$ by the natural action of Γ. The unitary tangent bundle $T^1\mathcal{H}^d$ is also naturally identified with pairs (x, v), where the base point x is an element of $\mathcal{H}^d$, and v a unit spacelike vector in Minkowski space orthogonal to x. The geodesic flow $\widetilde{\phi}^t$ on $T^1\mathcal{H}^d$ is then:

$$\widetilde{\phi}^t(x,v) = (\cosh(t)x + \sinh(t)v, \sinh(t)x + \cosh(t)v)\ .$$

Definition 16.4.1 (Foliated Bundle Over M) Let E_τ be the quotient of the product $T^1\mathcal{H}^d \times \partial^*\mathbb{R}^{d,1}$ by the diagonal action of Γ_τ, where Γ_τ acts on $\mathcal{H}^d$ through its linear part. Let $\pi_\tau : E_\tau \to M$ be the map induced by the projection on the first factor. This map is a fibration, of fiber $\partial^*\mathbb{R}^{d,1}$. It is called the foliated bundle of the holonomy group Γ_τ over M.

Definition 16.4.2 (Lifted Geodesic Flow) Let $\widetilde{\phi}^t_\tau$ be the flow on $T^1\mathcal{H}^d \times \partial^*\mathbb{R}^{d,1}$ defined by:

$$\widetilde{\phi}^t((x,v),\xi) = (\widetilde{\phi}^t(x,v),\xi)\ .$$

This flow commutes with the Γ_τ action, and induces a flow on E_τ, denoted by ϕ^t_τ.

We clearly have:

$$\forall t \in \mathbb{R}\ \ \phi^t_\tau \circ \pi_\tau = \pi_\tau \circ \widetilde{\phi}^t\ .$$

We also can distinguish two subbundles $\Delta^\pm_\tau$ of $\pi_\tau : E_\tau \to M$. More precisely:

Lemma 16.4.3 *Let (x, v) in $T^1\mathcal{H}^d$. Let (w, h) be an element of $\partial^*\mathbb{R}^{d,1}$ parametrized by the pair (w, h) under the identification defined above associated to x. Then:*

$$-1 \le \langle w, v\rangle_{d,1} \le 1\ .$$

Moreover, the equality $\langle w, v\rangle_{d,1} = 1$ holds if and only if $w = x + v$, and the equality $\langle w, v\rangle_{d,1} = -1$ holds if and only if $w = x - v$.

Proof For every (x, v) in $T^1\mathcal{H}^d$, and every lightlike element w of H_x, v, $-v$ and $w-x$ are unit elements in the Euclidean hyperplane $x^\perp$. The Lemma follows easily since $\langle w, v\rangle_{d,1} = \langle w-x, v\rangle_{d,1}$. □

Definition 16.4.4 We denote by $\widetilde{\Delta}^+$ (respectively $\widetilde{\Delta}^-$) the closed subset of $T^1\mathcal{H}^d \times \partial^*\mathbb{R}^{d,1}$ comprising elements (x, v, ξ) such that the orthogonal of the lightlike hyperplane ξ is $x+v$ (respectively $x-v$).

The complement $T^1\mathcal{H}^d \times \partial^*\mathbb{R}^{d,1} \setminus \widetilde{\Delta}^\pm$ is an open subset that we denote by $\widetilde{\aleph}^\pm$.

It is straightforward to check that $\widetilde{\Delta}^\pm$ and $\widetilde{\aleph}^\pm$ are Γ_τ-invariant and define closed subsets $\Delta^\pm_\tau$ and open subsets $\aleph^\pm_\tau$ of E_τ. Moreover:

Lemma 16.4.5 *$\widetilde{\Delta}^\pm$ and $\widetilde{\aleph}^\pm$ are $\widetilde{\phi}^t$-invariant.*

Proof We just have to prove that $\widetilde{\Delta}^\pm$ is $\widetilde{\phi}^t$-invariant. We just treat the case of $\widetilde{\Delta}^+$, the case of $\widetilde{\Delta}^-$ is similar.

Let (x, v, ξ) be an element of $\widetilde{\Delta}^+$: this means that, for the parametrization defined by x, the lightlike hyperplane ξ is parametrized by (w, h), where $w = x+v$—or, equivalently, $\langle w, v\rangle_{d,1} = 1$ (see Lemma 16.4.3). Denote by (x_t, v_t) the iterate $\widetilde{\phi}^t(x, v)$. Let (w_t, h_t) be the pair parameterizing ξ for the identification defined by x_t. Then, $w_t = \lambda_t(x+v)$ for some positive real number λ_t. We must have:

$$\begin{aligned} -1 &= \langle w_t, x_t\rangle_{d,1} \\ &= \langle \lambda_t(x+v), \cosh(t)x + \sinh(t)v\rangle_{d,1} \\ &= -\lambda_t\cosh(t) + \lambda_t\sinh(t) \\ &= -\lambda_t\exp(-t)\ . \end{aligned}$$

Therefore $\lambda_t = \exp(t)$, and:

$$\begin{aligned} w_t &= \exp(t)(x+v) \\ &= (\cosh(t)+\sinh(t))x + (\cosh(t)+\sinh(t))v \\ &= (\cosh(t)x + \sinh(t)v) + (\sinh(t)x + \cosh(t)v))_{d,1} \\ &= x_t + v_t\ . \end{aligned}$$

The lemma follows.

□

Therefore, $\Delta^\pm_\tau$ are ϕ^t_τ-invariant. The restriction of π_τ to $\Delta^\pm_\tau$ is a fibration, with 1-dimensional fibers. The restriction $\pi^\pm_\tau$ to $\aleph^\pm_\tau$ is a fibration with contractible fibers. Indeed, every fiber is the complement in $\partial^*\mathbb{R}^{d,1}$ of a degenerate vertical line removed, i.e. the product of a 1-punctured sphere by the real line.

Definition 16.4.6 Let $\mathcal{F}_\tau$ be the space of continuous sections of the fibration $\pi_\tau : E_\tau \to M$. We denote by $\mathcal{F}_\tau^\pm$ the open subset comprising sections of $\pi_\tau : \aleph_\tau^\pm \to M$, and by $\mathcal{F}(\Delta^\pm)_\tau$ the space of sections of $\pi_\tau^\pm : \Delta_\tau^\pm \to M$.

Let σ be an element of $\mathcal{F}_\tau$. It lifts uniquely to a Γ_τ-equivariant section of the fibration $T^1\mathcal{H}^d \times \partial^*\mathbb{R}^{d,1} \to T^1\mathcal{H}^d$ and therefore provides a Γ_τ-equivariant map $F : T^1\mathcal{H}^d \to \partial^*\mathbb{R}^{d,1}$. Actually, $\mathcal{F}_\tau$ is in 1-1 correspondence with the space of Γ_τ-equivariant maps from $T^1\mathcal{H}^d$ into $\partial^*\mathbb{R}^{d,1}$.

Given two elements σ_1, σ_2 of $\mathcal{F}_\tau$, let F_1, F_2 their associated Γ_τ-equivariant maps from $T^1\mathcal{H}^d$ into $\partial^*\mathbb{R}^{d,1}$. We define:

$$D(\sigma_1, \sigma_2) = \sup_{(x,v)\in T^1\mathcal{H}^d} d_x(F_1(x, v), F_2(x, v)) .$$

Since M is compact, the Γ_τ-equivariance implies that this upper bound is always attained.

It defines a metric D on $\mathcal{F}_\tau$. Observe that the metric space $(\mathcal{F}_\tau, D)$ is complete.

The flow ϕ_τ induces a 1-parameter group of transformations on $(\mathcal{F}_\tau, D)$: for every t in $\mathbb{R}$, and any σ in $\mathcal{F}_\tau$, define:

$$\Phi_\tau^t(\sigma)(x, v) = \phi_\tau^t(\sigma(\phi^{-t}(x, v))) .$$

According to Lemma 16.4.5, the subbundles $\mathcal{F}_\tau^\pm$ are Φ_τ^t-invariant.

We can now prove the fundamental fact:

Lemma 16.4.7 *The flow Φ_τ^t on $\mathcal{F}_\tau^+$ is exponentially contracting: there exist positive real numbers T, a and $0 < C < 1$ such that, for every $t > T$ and for every σ_1, σ_2 in $\mathcal{F}_\tau^+$ we have:*

$$D(\Phi_\tau^t(\sigma_1), \Phi_\tau^t(\sigma_2)) < Ce^{-at} D(\sigma_1, \sigma_2) .$$

Proof Let $F : T^1\mathcal{H}^d \to \partial^*\mathbb{R}^{d,1}$ be a Γ_τ-equivariant map corresponding to elements of $\mathcal{F}_\tau^+$. We denote by F_t the iterate $\Phi_\tau^t(F)$. Let (x, v) be an element of $T^1\mathcal{H}^d$. Let ξ be the image $F(x, v)$. This is an affine lightlike hyperplane. By definition of Φ_τ^t, ξ is the image under F_t of $\widetilde{\phi}^t(x, v) = (x_t, v_t) = (\cosh(t)x + \sinh(t)v, \sinh(t)x + \cosh(t)v)$. Let (w_t, h_t) be the pair corresponding to ξ satisfying $\langle x_t, w_t\rangle_{d,1} = -1$ and such that ξ is the hyperplane of equation:

$$\langle w_t, .\rangle_{d,1} = -h_t .$$

In particular, we see that $-hx$ belongs to ξ, and therefore, for every t we have:

$$h_t = -h\langle w_t, x\rangle_{d,1} . \tag{16.43}$$

Since the lightlike vectors w_t are all orthogonal to ξ, they are proportional: for every t, there is a real number $\lambda_t > 0$ such that $w_t = \lambda_t w_0$. From Eq. (16.43) we see:

$$h_t = h\lambda_t \ .$$

A straightforward computation shows:

$$\lambda_t = \frac{1}{\cosh(t) - \sinh(t)\langle v, w_0\rangle_{d,1}} \ .$$

Let now F_1, F_2 be two Γ_τ-equivariant maps from $T^1\mathcal{H}^d$ into $\partial^*\mathbb{R}^{d,1}$ corresponding to sections of $\pi_\tau : \aleph_\tau^+ \to M$. The distance in $\digamma_\tau$ between the corresponding sections is then the supremum of $d_x(F_1(x, v), F_2(x, v))$ where (x, v) describes $T^1\mathcal{H}^d$. Applying Φ_τ^t simply means that we replace F_1 and F_2 by $F_1 \circ \widetilde{\phi}_\tau^{-t}$ and $F_2 \circ \widetilde{\phi}_\tau^{-t}$. It follows that the distance after applying Φ^τ is the supremum of $d_{x_t}(F_1(x, v), F_2(x, v))$ where (x, v) describes $T^1\mathcal{H}^d$ and *where x_t denotes as above the x component of* $\widetilde{\phi}^t(x, v)$, i.e. $\cosh(t)x + \sinh(t)v$.

The computation above shows that, for $i = 1, 2$, the pair (w_t^i, h_t^i) representing $F_i(x, v)$ satisfies:

$$w_t^i = \frac{w_0^i}{\cosh(t) - \sinh(t)\langle v, w_0\rangle_{d,1}} \ ,$$

$$h_t^i = \frac{h_0^i}{\cosh(t) - \sinh(t)\langle v, w_0\rangle_{d,1}} \ .$$

Therefore, $d_{x_t}(F_1(x, v), F_2(x, v)) = \frac{d_{x_0}(F_1(x,v), F_2(x,v))}{\cosh(t)-\sinh(t)\langle v, w_0\rangle_{d,1}}$. Since the F_1 and F_2 correspond to sections in $\digamma_\tau^+$, we have:

$$-1 \leq \langle v, w_0\rangle_{d,1} < 1 \ .$$

It follows that for big t, the quantity $\cosh(t) - \sinh(t)\langle v, w_0\rangle_{d,1}$ is equivalent to $e^t(1 - \langle v, w_0\rangle_{d,1})/2$. The lemma follows. □

Corollary 16.4.8 *There exists one and only one Φ_τ^t-invariant section σ_τ^+ of $\pi_\tau : \aleph_\tau^+ \to M$. This invariant section actually takes value in Δ_τ^-.*

Proof Let $T > 0$ be a real number big enough so that Φ_τ^T is contracting. Since Δ_τ^- is a subbundle of $\aleph_\tau^+$, $\digamma(\Delta^-)_\tau$ is a closed subset of $\digamma_\tau^+$. The restriction D to $\digamma(\Delta^-)_\tau$ is therefore complete. Hence, as any contracting map acting on a complete metric space, Φ_τ^T admits a unique fixed point σ_τ^+ in $\digamma(\Delta^-)_\tau$. Since its action on $\digamma_\tau^+$ is contracting too, σ_τ^+ is the unique fixed point in $\digamma_\tau^+$. Since Φ_τ^T commutes with Φ_τ^t for every real number t, σ_τ^+ is fixed by every Φ_τ^t. □

Let $F_\tau : T^1\mathcal{H}^d \to \partial^*\mathbb{R}^{d,1}$ be the Γ_τ-equivariant lift of the Φ^t_τ-invariant section σ^+_τ exhibited in Corollary 16.4.8. The Φ^t_τ-invariance means that F_τ is constant along the orbits of the geodesic flow $\widetilde{\phi}^t$ of $T^1\mathcal{H}^d$. The following Lemma shows that we have much more:

Lemma 16.4.9 *The map F_τ is constant along the leaves of the weak unstable foliation of the geodesic flow ϕ^t.*

Proof Let θ_1, θ_2 be two orbits of $\widetilde{\phi}^t$ in the same *unstable leaf*, i.e. such that for every (x_1, v_1) in θ_1 and every (x_2, v_2) in θ_2 the isotropic vectors $x_1 - v_1$ and $x_2 - v_2$ are proportional, i.e. represent the same element of $\partial\mathcal{H}^d$. On the other hand, since the invariant section takes its values in Δ^-_τ, $F_\tau(x_1, v_1)$ and $F_\tau(x_2, v_2)$ are lightlike hyperplanes orthogonal to respectively $x_1 - v_1$ and $x_2 - v_2$. Therefore, they are parallel.

Let p_1, p_2 be the projections of (x_1, v_1) and (x_2, v_2) in M. Then, by replacing p_2 by another element of its ϕ^t-orbit, one can assume that p_1 and p_2 lie in the same *strong unstable leaf*, i.e. that the hyperbolic distance between $\phi^t(p_1)$ and $\phi^t(p_2)$ converge exponentially to 0 when t goes to $-\infty$.

It follows that the hyperbolic distance between $\widetilde{\phi}^t(x_1, v_1)$ and $\widetilde{\phi}^t(x_2, v_2)$ converges to 0 when t tends to $-\infty$. Let $\xi_1 = F_\tau(x_1, v_1)$ and $\xi_2 = F_\tau(x_2, v_2)$. Since F_τ is (uniformly) continuous, it follows that $d_t(\xi_1, \xi_2)$ converges to 0, where d_t is the distance on $\partial^*\mathbb{R}^{d,1}$ defined by $\widetilde{\phi}^t(x_1, v_1)$. But this is almost a contradiction with Lemma 16.4.7, which shows that this distance should be exponentially increasing when t tends to $-\infty$. The only possibility is that this distance actually vanishes, i.e. $\xi_1 = \xi_2$. The lemma is proved.

□

In the sequel, we use the cylindrical affine model of the co-Minkowski space, i.e. write elements of $\partial^*\mathbb{R}^{d,1}$ as pairs (w, h) where w is a lightlike vector satisfying $\langle x_0, w\rangle_{d,1} = -1$, where x_0 denotes the element $(0, \dots, 0, 1)$ of $\mathbb{R}^{d,1}$.

Proposition 16.4.10 *There is a continuous map $b_\tau : \partial B^d \to \mathbb{R}$ such that the Γ_τ-equivariant map $F_\tau : T^1\mathcal{H}^d \to \partial^*\mathbb{R}^{d,1}$ is given by:*

$$(x, v) \mapsto \left(-\frac{x - v}{\langle x_0, x - v\rangle_{d,1}}, b_\tau\left(-\frac{x - v}{\langle x_0, x - v\rangle_{d,1}}\right)\right) .$$

Proof We still parameterize the unit tangent bundle of the hyperbolic space by pairs (x, v) where x is a unit timelike vector and v a unit spacelike vector orthogonal to x.

Since the invariant section takes its values in the subbundle Δ^-, the map F_τ is such that $F_\tau(x, v) = (w(x, v), h(x, v))$ where $w(x, v)$ is proportional to $x - v$, hence is equal to $-\frac{x-v}{\langle x_0, x-v\rangle_{d,1}}$. Moreover, according to Lemma 16.4.9, $h(x, v)$

depends only on $x - v$, hence, only on $-\frac{x-v}{\langle x_0, x-v\rangle_{d,1}}$. Therefore, F_τ is given by:

$$(x, v) \mapsto \left(-\frac{x-v}{\langle x_0, x-v\rangle_{d,1}}, b_\tau\left(-\frac{x-v}{\langle x_0, x-v\rangle_{d,1}}\right)\right)$$

for some map $b_\tau : \partial B^d \to \mathbb{R}$. □

As a corollary, we get the following improvement of Lemma 16.3.27:

Lemma 16.4.11 *Let $\tau_n \to \tau$. Then b_{τ_n} (resp. $h^\pm_{\tau_n}$, $h^{\text{mean}}_{\tau_n}$) converge uniformly to b_τ (resp. $h^\pm_\tau$, h^{mean}_τ).*

Proof We just give a sketch of proof. First, we observe that we just have to prove the statement for b_{τ_n}, since the uniform convergence of $h^\pm_{\tau_n}$ (resp. h^{mean}_τ) follows then from Lemma 16.2.43 (resp. Lemma 16.2.38). The key point is that when n is big enough, the fibration $\pi_{\tau_n} : E_{\tau_n} \to M$ is isomorphic to the fibration $\pi_\tau : E_\tau \to M$. More precisely, (the inverse of) this isomorphism of fibrations sends the graph of the section σ_τ to the graph of some section which is already an almost fixed point for $\Phi^t_{\tau_n}$. The bigger n is, the closer (for the metric D) is this almost fixed point to the eventual fixed point $\sigma^+_{\tau_n}$. In other words, the bigger is n, the closer to σ^+_τ is $\sigma^+_{\tau_n}$ for the compact-open topology. The Lemma clearly follows, due to the form of the lifts F_τ and F_{τ_n} given by Proposition 16.4.10. □

Remark 16.4.12 *Mutandi mutandis,* one can show that there is also a unique fixed point for Φ^t_τ in $\digamma^-_\tau$, which this time is an exponential repeller, and which is actually a section of the subbundle Δ^+. It provides, as in Proposition 16.4.10 a map from ∂B^d into $\mathbb{R}$, which is actually the map b_τ. Details are left to the reader.

Remark 16.4.13 Instead of considering the fiber bundle $\aleph^\pm_\tau$, one might have restricted the study to the subbundles $\Delta^\pm$, which are simpler since with one-dimensional fibers. However, the most efficient way to deal with these bundles is to consider them as subbundles of E_τ.

We conclude this section by an interpretation of its content in term of Anosov representations. Let G be a general Lie group acting on some space X, and let $\rho : \Gamma \to G$ be a representation. Consider as in Definition 16.4.1 the foliated bundle $\pi_\rho : E_\rho(X) \to M$ where $E_\rho(X)$ is the quotient of the product $T^1\mathcal{H}^d \times X$ by the diagonal action of Γ and where the action of Γ on X is given by ρ. As in Definition 16.4.2, the geodesic flow ϕ^t lifts to some horizontal flow ϕ^t_ρ on $E_\rho(X)$ so that the bundle map π_ρ is equivariant.

The representation ρ is said to be (G, X)*-Anosov* if the following holds: there is a section $\sigma : M \to E_\rho(X)$ which is equivariant for the flows, and such that the graph Λ of σ is a *closed hyperbolic subset* for the lifted flow ϕ^t_ρ: this means that the restriction $T_\Lambda E_\rho(X)$ of the tangent bundle of $E_\rho(X)$ to Λ splits as a Whitney sum of subbundles $E^+ \oplus E^- \oplus \Phi$, where:

- Φ is the one dimensional bundle tangent to the flow ϕ^t_ρ,
- E^+ is exponentially contracted by the flow,
- E^- is exponentially expanded by the flow.

For more details, see [47] or [4, 5].

In our case, the inclusion $\Gamma \approx \Gamma_\tau \subset SO_+(d,1) \ltimes \mathbb{R}^{d,1}$ is (G,X)-Anosov where X is the space of oriented $(d-1)$-dimensional spacelike affine subspaces of $\mathbb{R}^{d,1}$. Indeed, X is identified with the open domain in $\partial^*\mathbb{R}^{d,1} \times \partial^*\mathbb{R}^{d,1}$ made of pairs (ξ_1, ξ_2), where ξ_1 and ξ_2 are non-parallel affine lightlike hyperplanes. Therefore, the two equivariant sections $\sigma^\pm_\tau$ define altogether a section σ of $\pi_\rho : E_\rho(X) \to M$. Moreover, it follows from Lemma 16.4.7 and Remark 16.4.12 that the graph of σ is a closed hyperbolic subset for ϕ^t_ρ.

Acknowledgments The authors want to thank the referee for her/his careful reading of the manuscript. The present work is also part of the Math Amsud 2014 project n°38888QB-GDAR.

References

1. L. Andersson, T. Barbot, R. Benedetti, F. Bonsante, W.M. Goldman, F. Labourie, K.P. Scannell, J.-M. Schlenker, Notes on: "Lorentz spacetimes of constant curvature" [Geom. Dedicata **126** (2007), 3–45; mr2328921] by G. Mess. Geom. Dedicata **126**, 47–70 (2007)
2. B.N. Apanasov, Bending and stamping deformations of hyperbolic manifolds. Ann. Global Anal. Geom. **8**(1), 3–12 (1990)
3. T. Barbot, Globally hyperbolic flat space-times. J. Geom. Phys. **53**(2), 123–165 (2005)
4. T. Barbot, Three-dimensional Anosov flag manifolds. Geom. Topol. **14**(1), 153–191 (2010)
5. T. Barbot, Deformations of Fuchsian AdS representations are quasi-Fuchsian. J. Differ. Geom. **101**(1), 1–46 (2015)
6. T. Barbot, Lorentzian Kleinian groups, in *Handbook of Group Actions. Vol. III* Advanced Lectures in Mathematics, vol. 40 (International Press, Somerville, MA, 2018), pp. 311–358
7. A. Bart, K.P. Scannell, A note on stamping. Geom. Dedicata **126**, 283–291 (2007)
8. R. Benedetti, F. Bonsante, Canonical Wick rotations in 3-dimensional gravity. Mem. Am. Math. Soc. **198**(926), viii+164 (2009)
9. I. Bivens, J.-P. Bourguignon, A. Derdziński, D. Ferus, O. Kowalski, T. Klotz Milnor, V. Oliker, U. Simon, W. Strübing, K. Voss. Discussion on Codazzi-tensors, in *Global Differential Geometry and Global Analysis (Berlin, 1979)*. Lecture Notes in Mathematics, vol. 838 (Springer, Berlin/New York, 1981), pp. 243–299
10. M. Bridgeman, R. Canary, F. Labourie, A. Sambarino, The pressure metric for Anosov representations. Geom. Funct. Anal. **25**(4), 1089–1179 (2015)
11. F. Bonahon, Geodesic laminations on surfaces, in *Laminations and Foliations in Dynamics, Geometry and Topology (Stony Brook, NY, 1998)*. Contemporary Mathematics, vol. 269 (American Mathematical Society, Providence, RI, 2001), pp. 1–37
12. F. Bonsante, Flat spacetimes with compact hyperbolic Cauchy surfaces. J. Differ. Geom. **69**(3), 441–521 (2005)
13. F. Bonsante, F. Fillastre, The equivariant Minkowski problem in Minkowski space. Ann. Inst. Fourier (Grenoble) **67**(3), 1035–1113 (2017)
14. F. Bonsante, J.-M. Schlenker, Fixed points of compositions of earthquakes. Duke Math. J. **161**(6), 1011–1054 (2012)

15. F. Bonsante, A. Seppi, On Codazzi tensors on a hyperbolic surface and flat Lorentzian geometry. Int. Math. Res. Not. (2), 343–417 (2016)
16. F. Bonsante, A. Seppi, A. Tamburelli, On the volume of anti–de Sitter maximal globally hyperbolic three-manifolds. Geom. Funct. Anal. **27**(5), 1106–1160 (2017)
17. J.F. Brock, The Weil–Petersson metric and volumes of 3-dimensional hyperbolic convex cores. J. Am. Math. Soc. **16**(3), 495–535 (2003)
18. R.D. Canary, D.B.A. Epstein, P.L. Green, Notes on notes of Thurston [mr0903850], in *Fundamentals of Hyperbolic Geometry: Selected Expositions*. London Mathematical Society Lecture Note series, vol. 328, pp. 1–115 (Cambridge University Press, Cambridge, 2006) With a new foreword by Canary
19. R.D. Canary, M. Lee, M. Stover, Amalgam Anosov representations. Geom. Topol. **21**(1), 215–251 (2017). With an appendix by Canary, Lee, Andrés Sambarino and Stover
20. J. Danciger, Geometric transitions: From hyperbolic to AdS geometry. PhD thesis, Stanford University, 2011
21. J. Danciger, A geometric transition from hyperbolic to anti-de Sitter geometry. Geom. Topol. **17**(5), 3077–3134 (2013)
22. J. Danciger, F. Guéritaud, F. Kassel, Convex cocompact actions in real projective geometry. ArXiv e-prints, April 2017
23. J. Danciger, F. Guéritaud, F. Kassel, Convex cocompactness in pseudo-Riemannian hyperbolic spaces. Geom. Dedicata **192**, 87–126 (2018)
24. J. Danciger, S. Maloni, J.-M. Schlenker, Polyhedra inscribed in a quadric. Invent. Math. **221**(1), 237– 300 (2020)
25. T.A. Drumm, W.M. Goldman, The geometry of crooked planes. Topology **38**(2), 323–351 (1999)
26. D.B.A. Epstein, A. Marden, Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces [mr0903852], in *Fundamentals of Hyperbolic Geometry: Selected Expositions*. London Mathematical Society Lecture Note Series, vol. 328 (Cambridge University Press, Cambridge, 2006), pp. 117–266
27. F. Fillastre, A. Seppi, Generalization of a formula of Wolpert for balanced geodesic graphs on closed hyperbolic surfaces. Annales Henri Lebesgue **3**, 873–899 (2020)
28. F. Fillastre, A. Seppi, Spherical, hyperbolic, and other projective geometries: convexity, duality, transitions, in *Eighteen Essays in Non-Euclidean Geometry*. IRMA Lectures in Mathematics and Theoretical Physics, vol. 29 (European Mathematical Society, Zürich, 2019), pp. 321–409
29. F. Fillastre, G. Smith, Group actions and scattering problems in Teichmüller theory, in *Handbook of Group Actions. Vol. III*. Advanced Lectures in Mathematics, vol. 40 (International Press, Somerville, MA, 2018), pp. 359–417
30. F. Fillastre, G. Veronelli, Lorentzian area measures and the Christoffel problem. Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **16**(2), 383–467 (2016)
31. M. Ghomi, The problem of optimal smoothing for convex functions. Proc. Am. Math. Soc. **130**(8), 2255–2259 (2002)
32. S. Ghosh, Anosov structures on Margulis spacetimes. Groups Geom. Dyn. **11**(2), 739–775 (2017)
33. D. Gilbarg, N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics (Springer, Berlin, 2001). Reprint of the 1998 edition
34. W.M. Goldman, The symplectic nature of fundamental groups of surfaces. Adv. Math. **54**(2), 200–225 (1984)
35. W.M. Goldman, Flat affine, projective and conformal structures on manifolds: a historical perspective, in *Geometry in History*, ed. by S.G. Dani, A. Papadopoulos (Springer, Cham, 2019), pp. 515–522
36. F. Guéritaud, O. Guichard, F. Kassel, A. Wienhard, Anosov representations and proper actions. Geom. Topol. **21**(1), 485–584 (2017)
37. O. Guichard, A. Wienhard, Anosov representations: domains of discontinuity and applications. Invent. Math. **190**(2), 357–438 (2012)

38. C. Gutiérrez, *The Monge-Ampère Equation*. Progress in Nonlinear Differential Equations and their Applications, vol. 44 (Birkhäuser Boston Inc., Boston, MA, 2001)
39. L. Hörmander, *The Analysis of Linear Partial Differential Operators. I. Distribution Theory and Fourier Analysis*. Classics in Mathematics (Springer, Berlin, 2003). Reprint of the second (1990) edition [Springer, Berlin; MR1065993 (91m:35001a)]
40. I. Izmestiev, Statics and kinematics of frameworks in Euclidean and non-Euclidean geometry, in *Sixteen Essays on non-Euclidean Geometry*, ed. by A. Papadopoulos (European Mathematical Society Publishing House, Zürich, 2018)
41. D. Johnson, J.J. Millson, Deformation spaces associated to compact hyperbolic manifolds, in *Discrete Groups in Geometry and Analysis (New Haven, Conn., 1984)*. Progress in Mathematical, vol. 67 (Birkhäuser Boston, Boston, MA, 1987), pp. 48–106
42. M. Kapovich, *Hyperbolic Manifolds and Discrete Groups*. Modern Birkhäuser Classics (Birkhäuser Boston, Inc., Boston, MA, 2009). Reprint of the 2001 edition
43. M. Kapovich, B. Leeb, J. Porti, Some recent results on Anosov representations. Transform. Groups **21**(4), 1105–1121 (2016)
44. S.P. Kerckhoff, Earthquakes are analytic. Comment. Math. Helv. **60**(1), 17–30 (1985)
45. C. Kourouniotis, Deformations of hyperbolic structures. Math. Proc. Cambr. Philos. Soc. **98**(2), 247–261 (1985)
46. K. Krasnov, J.-M. Schlenker, Minimal surfaces and particles in 3-manifolds. Geom. Dedicata **126**, 187–254 (2007)
47. F. Labourie, Anosov flows, surface groups and curves in projective space. Invent. Math. **165**(1), 51–114 (2006)
48. J. Lafontaine, Modules de structures conformes plates et cohomologie de groupes discrets. C. R. Acad. Sci. Paris Sér. I Math. **297**(13), 655–658 (1983)
49. A.M. Li, Spacelike hypersurfaces with constant Gauss-Kronecker curvature in the Minkowski space. Arch. Math. (Basel) **64**(6), 534–551 (1995)
50. P. Majer (https://mathoverflow.net/users/6101/pietromajer), Uniform convergence of convex functions. MathOverflow. https://mathoverflow.net/q/196540 (version: 2015-02-15)
51. G. Mess, Lorentz spacetimes of constant curvature. Geom. Dedicata **126**, 3–45 (2007)
52. K. Nomizu, T. Sasaki, *Affine Differential Geometry. Geometry of Affine Immersions*. Cambridge Tracts in Mathematics, vol. 111 (Cambridge University Press, Cambridge, 1994)
53. V.I. Oliker, U. Simon, Codazzi tensors and equations of Monge-Ampère type on compact manifolds of constant sectional curvature. J. Reine Angew. Math. **342**, 35–65 (1983)
54. A. Papadopoulos, W. Su, On the Finsler structure of Teichmüller's metric and Thurston's metric. Expo. Math. **33**(1), 30–47 (2015)
55. A. Papadopoulos, G. Théret, On Teichmüller's metric and Thurston's asymmetric metric on Teichmüller space, in *Handbook of Teichmüller theory. Vol. I. IRMA Lectures in Mathematics and Theoretical Physics*, vol. 11 (European Mathematical Society, Zürich, 2007), pp. 111–204
56. R.C. Penner, J.L. Harer, *Combinatorics of Train Tracks*. Annals of Mathematics Studies, vol. 125 (Princeton University Press, Princeton, NJ, 1992)
57. S. Riolo, A. Seppi, Geometric transition from hyperbolic to anti-de sitter structures in dimension four (2019). arXiv:1908.05112v2 [math.GT]
58. J.-M. Schlenker, Variétés lorentziennes plates vues comme limites de variétés anti–de Sitter [d'après Danciger, Guéritaud et Kassel]. Astérisque (380, Séminaire Bourbaki. Vol. 2014/2015):Exp. No. 1103, 475–497 (2016)
59. R. Schneider, *Convex Bodies: The Brunn-Minkowski Theory*. Encyclopedia of Mathematics and its Applications, vol. 151 (Cambridge University Press, Cambridge, 2014), expanded edition
60. A. Seppi, Surfaces in constant curvature three-manifolds and the infinitesimal Teichmüller theory. PhD thesis, University of Pavia (2015)
61. G.A.C. Smith, A short proof of an assertion of thurston concerning convex hulls, ed. by K. Ohshika, A. Papadopoulos, in *In the Tradition of Thurston* (Springer, Cham, 2020), pp. 255–261

62. G.A.C. Smith, *Global Singularity Theory for the Gauss Curvature Equation*. Ensaios Matemáticos [Mathematical Surveys], vol. 28 (Sociedade Brasileira de Matemática, Rio de Janeiro, 2015)
63. W.P. Thurston, Minimal stretch maps between hyperbolic surfaces. ArXiv Mathematics e-prints, January 1998
64. A.J. Tromba, *Teichmüller theory in Riemannian geometry*. Lectures in Mathematics ETH Zürich (Birkhäuser Verlag, Basel, 1992). Lecture notes prepared by Jochen Denzler
65. C. Walsh, The horoboundary and isometry group of Thurston's Lipschitz metric, in *Handbook of Teichmüller theory. Vol. IV*. IRMA Lectures in Mathematics and Theoretical Physics, vol. 19 (European Mathematical Society, Zürich, 2014), pp. 327–353
66. S. Wolpert, On the symplectic geometry of deformations of a hyperbolic surface. Ann. Math. (2) **117**(2), 207–234 (1983)

Index

K. Ohshika, A. Papadopoulos (eds.), *In the Tradition of Thurston*,
https://doi.org/10.1007/978-3-030-55928-1